CHAPTER 13

CRM Applications and Data Mining

This chapter introduces concepts of customer relationship management (CRM), classifications, applications, and various data mining techniques used in CRM applications.

13.1 INTRODUCTION

Well-run small businesses naturally form learning relationships with their customers. These businesses learn more and more about their customers over the time and make use of that knowledge to serve them better. The result is happy, loyal customers and profitable businesses. In the case of large companies, with a large number of customers, actual personal relationships do not come in the picture. These large firms must rely on the abundant customer data and various techniques that turn customer data into customer knowledge to form learning relationships with their customers.

Customers are the most important asset of an organization. Customer satisfaction and their relationship with the organization are crucial in a business environment. Without satisfied customers who remain loyal, there cannot be any business prospects. For this reason, an organization should plan and employ a clear strategy for treating customers. CRM is the strategy for developing, managing, and strengthening loyal customer relationships [1]. CRM helps businesses, influences marketing concepts, and uses technology and human resources to gain insight into the behavior and value of customers. For effective CRM, an organization must first understand its customer needs and then determine ways to meet those needs in the best possible manner [2, 3]. It must also look into the different ways of gathering information about customers and find where and how this

information is stored and is currently used. CRM is an effective business strategy that can increase revenues by:

- Promising better services and products to valuable customers
- Cross-selling products more efficiently with effective marketing and sales
- Retaining existing customers and discovering new customers by a customer-focused business philosophy

The main objectives of CRM are customer retention through customer satisfaction and customer development through customer insight. These objectives can be made clearer with a real-life example of two stationary stores with different selling approaches. The first store sells everything to everyone, whereas the second store tries to identify each customer's needs and make appropriate suggestions. The second store brings in changes in the items for sale in accordance with the needs of valuable customers. The example shows that the second one seems more reliable and forms a long-lasting relationship since it aims for customer satisfaction by considering specific customer needs.

CRM should be a customer-centric approach based on customer insight. A brief description of customer life cycle and the business processes organized around the cycle is given before studying the CRM framework and data mining techniques in CRM.

13.2 CUSTOMER LIFE CYCLE

The customer life cycle helps in understanding customers as it progresses through different stages, focusing specifically on the business relationship that evolves over time. Although each business is different, customer relationship places customers into five major phases, as shown in Fig. 13.1.

1. *Prospects* are people in the target market who are not yet customers.
2. *Responders* are prospects who have shown some interest by filling out an application or enquiring the sales division or registering on a web site.
3. *New customers* are responders who have made a commitment; it can be signing a contract, registering on a site, etc.
4. *Established customers* are those new customers who return, enjoying special offers and service packages, thereby strengthening the relation.
5. *Former customers* are those who have left, as a result of voluntary or forced or expected attrition.

For an e-media site, anyone on the web may be considered a prospect, whereas a responder may be someone who has visited the site. A new customer is someone who has registered, and an established customer is a repeated visitor. The definitions might be quite different for other businesses. Figure 13.2 shows various business processes organized around the customer life cycle.

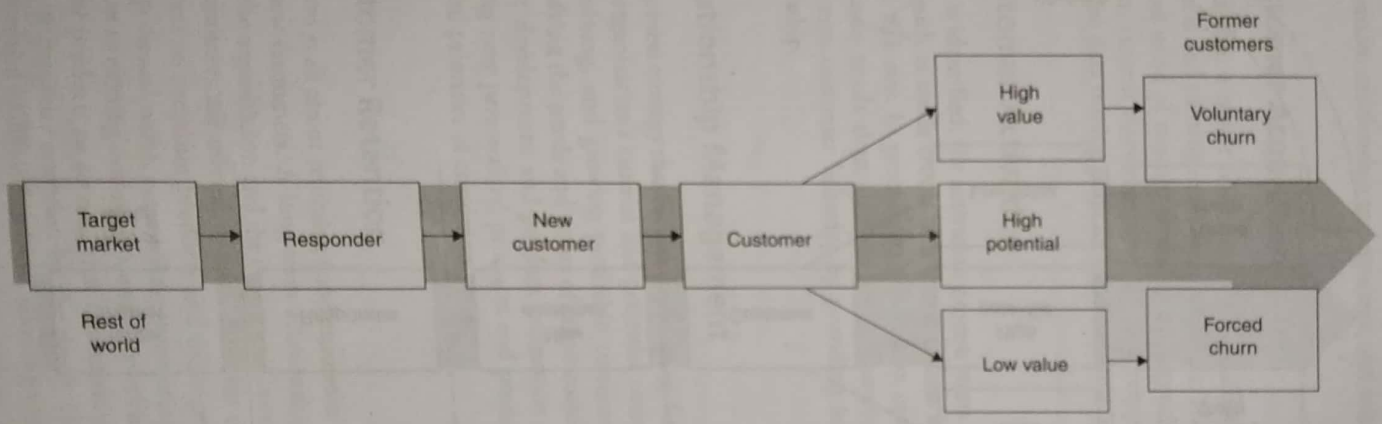

Fig. 13.1 Customer life cycle progresses through different stages

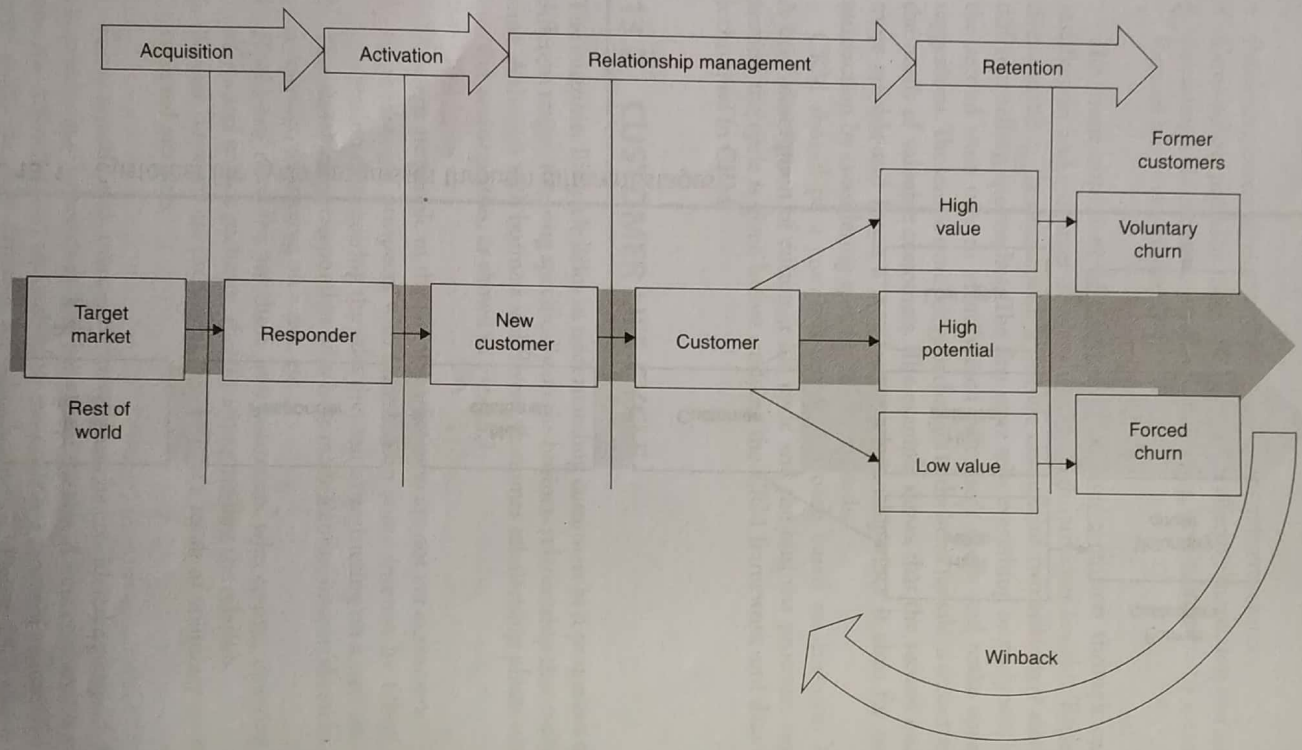

Fig. 13.2 Business processes organized around the customer life cycle

The business processes organized around the customer life cycle are customer acquisition, activation, relationship management, and retention.

13.2.1 Customer Acquisition

Customer acquisition, a major objective of all organizations, is the process of attracting prospects and turning them into customers. The process is often done by advertising and word of mouth, as well as by targeted marketing. Enquiries, former customers, current customers, suspects, and competitors serve as sources of acquisition. Also data mining plays an important role in acquisition.

13.2.2 Customer Activation

Once a prospect is identified, the activation process starts. This may be a lengthy approval process such as credit check, or a simple process such as filling a registration form on the web site. In general, activation is an operational process, more focused on business needs than analytic needs. Customer activation helps in predicting long-term customer behavior by providing initial conditions of the customer relationship.

13.2.3 Relationship Management

The CRM is a business strategy that focuses on grasping, foreseeing, and managing the needs of an organization's current and potential customers. It is the process of acquiring, retaining, and growing profitable customers. Marketing aims at meeting and satisfying the needs and wants of target customers. Data mining can support customer development and product promotion campaigns. It can also help in developing new personalized products and product offerings that cover the preferences and priorities of customers.

13.2.4 Customer Retention

Customer retention is all about retaining good customers and is more important than attracting new customers. A long-term relationship of mutual benefits is formed between the organization and the customer. Customer loyalty is the main concern. New customers are offered opportunities for up-selling (offering and switching customers to premium products and other products more profitable than the already possessed ones), cross-selling (promoting and selling additional products or services to existing customers), and deep-selling (usage stimulation—increasing usage of products or services that customers already have). Also data mining can help in customer retention by the timely identification of valuable customers with increased likelihood to leave, allowing time for targeted retention campaigns. In due course, all customers leave, making retention an important

data mining application both for marketing and forecasting. And once customers have left, they may be convinced to return through win-back strategies. These strategies begin as part of the retention efforts and depend more on operational strategies than on data analysis. Data analysis can sometimes help determine why customers are leaving, particularly when customer service complaints can be incorporated into operational data.

13.3 CRM FRAMEWORK AND ARCHITECTURE

From the viewpoint of architecture, the entire CRM framework can be classified into three key components: operational CRM, analytical CRM, and collaborative CRM.

13.3.1 Operational CRM

Operational CRM systems aim to enhance the services that take care of their customers especially in business processes such as sales, marketing, and customer services. It will automate the communications and interactions with customers, ensuring a consistent picture of the customer's relationship with the organization at all customer "touch" (interaction) points. It gives access to information about customers and their needs. Call centers, data aggregation systems, and web sites are examples of operational CRM. However, these systems are just tools that can be used to support the strategy of effectively managing customers.

13.3.2 Analytical CRM

Analytical CRM refers to the analysis of customer information gathered by acquiring customer knowledge. This knowledge can help an organization to allocate resources to the group of most profitable customers. It can also be used in campaign management as well as in especially targeted marketing campaigns that optimize marketing effectiveness. The use of data mining models in the framework helps to assess the value of customers and understand and predict their behavior. CRM analytics helps in understanding the best customers quickly and effectively. It is about analyzing data patterns to extract knowledge for optimizing customer relationships. Data collection and analysis is a continuous iterative process, which not only leads to productive customer relations but also to improvement in supply chain management.

13.3.3 Collaborative CRM

Collaborative CRM refers to the sharing of collected information among various sections of organizations to improve the quality of services offered, which in

turn will enhance customer satisfaction and loyalty. The interaction between customers and organizations can take place through web pages, e-mail, and other similar vehicles.

CRM is an enterprise approach to manage customer relationships in an organized manner. The enterprise can build a database of customers that will help the management, salespeople, service providers, and customers to interact well and achieve various needs. The process ensures that employees within an organization build effective relationships between the organization and its customers. This will help organizations to figure out answers to questions such as the following:

1. Which customers are most profitable and why?
2. What promotions are most effective and for which customers?
3. What kind of customers will be interested in a new product?
4. What customers are at risk to defect to my competitor?

The database will provide all necessary information so that each customer can be analyzed uniquely for answering these questions. CRM provides the framework for analyzing customer profitability and improving marketing effectiveness. Many organizations have collected and stored abundant data about their customers, suppliers, and business partners over a period of time. However, it was not possible to discover valuable information hidden in these data. This, in turn, prevents these companies from transforming this data into knowledge. Therefore, the business motive is to extract valid, previously unknown, and comprehensible information from large databases and use it for improvement. To accomplish these goals, organizations should first capture and integrate both internal and external data that give a complete view of the whole organization. It should then mine the integrated data for extracting valuable information. Finally, it should organize and present the extracted information with knowledge for decision-making.

Looking into the business benefits of CRM, we can learn that it enables an organization to provide better customer service, that products can be cross-sold more effectively, and that it helps sales staff to close deals faster, discover new customers, increase customer revenues, and simplify marketing and sales processes. The CRM basics employed in business organizations to find the most valuable customers are known as "RFM"—recency, frequency, and monetary value.

1. How recently a customer purchased items? (Recency)
2. How frequently a customer purchased items? (Frequency)
3. How much a customer spends on each purchase? (Monetary value)

As proposed by Ngai *et al.* [4], a graphical classification framework for data mining techniques in CRM is shown in Fig. 13.3. Four major CRM dimensions—customer identification, customer attraction, customer retention, and customer

development—are identified as the closed cycle of a CRM system. They help organizations in creating a deeper understanding of customers to maximize customer value to the organization in the long term. Application of data mining techniques in CRM helps in extracting or detecting hidden customer characteristics and behaviors from large databases.

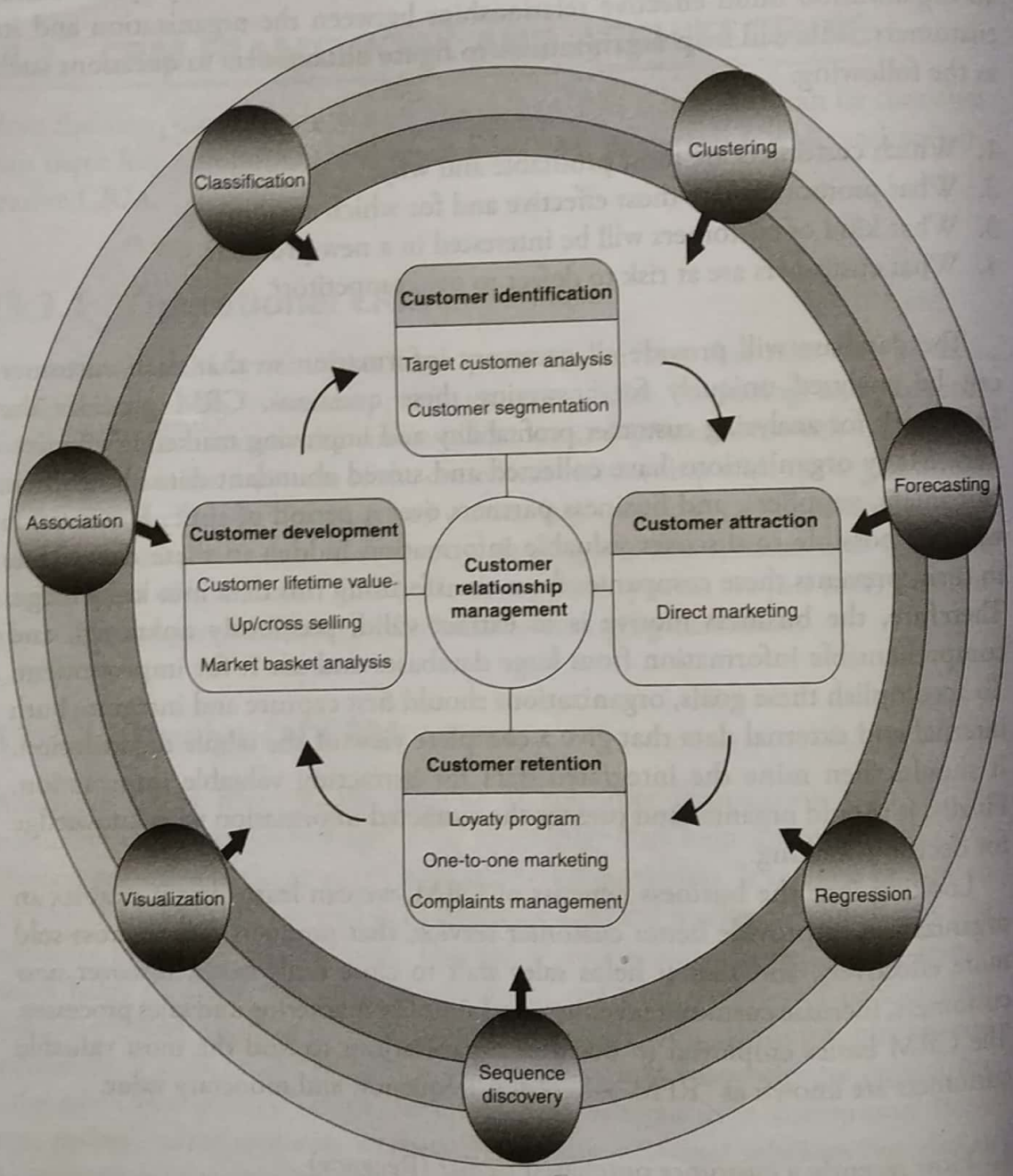

Fig. 13.3 Classification framework for data mining techniques in CRM

Data mining can be viewed as a business-driven process aimed at the discovery by consistent use of valuable knowledge from organizational data. It helps to

improve the decision-making process by giving a better insight into the domain concerned and hence making more accurate predictions.

13.4 CRM DIMENSIONS

Analyzing and understanding customer behaviors and characteristics are the major tasks in the development of a competitive CRM strategy. This strategy aids in acquiring and retaining prospective customers and maximizing customer value. To obtain a deeper understanding of each customer's behaviors, characteristics, and needs, detailed knowledge must be built up systematically. The four dimensions of the CRM [4, 5] cycle, described as follows, are essential efforts to gain customer insight.

13.4.1 Customer Identification

Customer identification, which is also referred to as customer acquisition, is the initial phase of the CRM framework. This phase targets people who are most likely to become customers or are most profitable to the company [6, 7]. It also targets customers who are being lost to the competition. The analysis focuses on how these customers can be won back. Detailed customer information will help in deciding the acquisition strategies to be adopted. The customer profiles build up as a result of analytics, and information mining will help in reaching the target set of customers.

13.4.2 Customer Attraction

The customer identification phase is followed by customer attraction. After identifying the segments of prospective customers, organizations can direct effort and resources into attracting the target customer segments. Direct marketing is an effective method for attracting customers in many areas. Networking, newsletters, and mails are some means by which more customers can be obtained.

13.4.3 Customer Retention

Customer retention is all about retaining the existing customers, which forms the major concern for CRM. Successful customer retention begins with the entry of customer to the organization and proceeds as long as the relationship lasts with the organization [8]. The ability to retain customers depends not only on products and sales but also on the services offered to customers from time to time. The elements of customer retention include personalized marketing campaigns, loyalty programs, and complaints management. Data mining can be very effectively applied to CRM applications to get vital information about the customer

base. Various data mining models such as decision trees may be used to model CRM applications [9].

13.4.4 Customer Development

Customer development strategies involve additional efforts to increase the value or profitability of existing customers. Elements of customer development include the prediction of the total net income of an organization, lifetime value of the customers, and promotion activities of the customer. Various customer segments play a major role in designing these strategies. The objective of the customer development strategy relates primarily to selling additional products and services to customers in the current database and selling higher value or higher profit margin products and services to the same customer segments.

13.5 CRM DATA MINING MODELS

Each data mining technique can perform one or more of the various types of data modeling, which includes association, classification, clustering, forecasting, regression, sequence discovery, and visualization. Table 13.1 shows the data mining modeling techniques and their applications.

13.5.1 Association

Association modeling is an important technique in predictive analytics and is used to detect associations between discrete products or attributes. With association models, it is possible to predict which items are most likely to appear together, as well as the strength of the relationships between them. The aim of the model is to establish relationships between items that exist together in a given record. For example, consider a supermarket that gathers data on customers' purchasing habits. The association rule learning can be used to determine which products are frequently bought together in the supermarket, and this information can be used for marketing purposes. From the sales data of the supermarket, it was indicated that if a customer buys computers and monitors together, he or she is likely to buy UPS also. Such kind of information can be used for decision-making about marketing activities,which includes promotional pricing, product placements, or market basket analysis. Common tools employed for association modeling are statistics and A priori algorithms.

13.5.2 Classification

Classification is a supervised data mining function used to predict group membership for data instances. The function assigns items in a collection to predefined

classes or categories based on certain criteria. The model helps in predicting customer behaviors through classification. For example, using classification model, we can identify loan applicants as low, medium, or high credit risks. Neural networks, decision trees, and if-then-else rules are the common tools used for classification. It has many applications in customer segmentation, marketing, and credit analysis.

13.5.3 Clustering

Clustering can be defined as the task of segmenting a heterogeneous population into a number of homogenous clusters [10]. It is an automated process that groups related records together on the basis of having similar values for attributes. Clusters are unknown at the time the algorithm starts i.e., there are no predefined clusters. The objective of this model is to discover segments or clusters and then examine the values and attributes that define the clusters or segments. Clustering is used for anomaly detection, exploring data, etc. Common tools used for clustering are neural networks and discrimination analysis.

13.5.4 Forecasting

Forecasting estimates the future value based on patterns in records. It can also be defined as the estimation of some variable of interest at some specified future date. A typical example of a forecasting model is demand forecasting, which involves informal methods, such as educated guesses, and quantitative methods such as use of historical sales data. Demand forecasting may be used in making pricing decisions as well. Neural networks and survival analysis are the common tools used for forecasting.

13.5.5 Regression

Regression is a kind of statistical estimation technique used to map each data object to a real value. Regression modeling has many applications in trend analysis, business planning, time-series prediction, marketing, financial forecasting, and biomedical and environmental modeling. There are regression algorithms estimating the value of the target in terms of function of the predictors for each of the cases in the build data. Common tools for regression include linear regression and logistic regression.

13.5.6 Sequence Discovery

Sequence discovery is a data mining function that identifies associations or patterns between data examples over time [11]. In sequence discovery, the

approach devised is to model the states of the process generating the sequence or to extract and report trends and deviation over time. Statistics and set theory are the common tools for sequence discovery.

13.5.7 Visualization

Visualization refers to the presentation of data in a way that enables users to view complex patterns [12]. The main objective of data visualization is to communicate information clearly and efficiently by means of graphics. The visualization model is used along with other data mining models to provide a better understanding of the discovered patterns or relationships. Examples of the visualization model are 3D graphs, Hygraphs, and See Net.

13.6 DATA MINING APPLICATIONS

Data mining activities fall into three general categories, from a process orientation [13] as shown in Fig. 13.4.

1. *Discovery*—the process of looking in a database to find hidden patterns without a predetermined idea about what the patterns may be.
2. *Predictive modeling*—the process of using patterns discovered from the database to predict the future.
3. *Forensic analysis*—the process of applying the extracted patterns to find anomalous or unusual data elements.

Four applications are discussed in the following sections, and each of them uses data mining activities, discovery, and predictive modeling. The discovery process is used to identify customer segments. The application categories discussed in the following sections describe some sort of predictive modeling. Each business is concerned with predicting the behavior of its customers through the knowledge gained from data mining. Table 13.1 shows some modeling techniques in data mining and their applications.

13.6.1 Retail

Retail can perform sales targeting from individuals or businesses to end users, and retailers can keep detailed records of every shopping transaction enabling them to better understand their customer segments. Retailers deal with target customers based on customer demographics, lifestyle, and purchase behavior. Various applications are as follows [6]:

Performing basket analysis: Basket analysis reveals which items customers tend to purchase together. It helps direct marketers, inventory management, improving store layout strategies, stocking, and promotions.

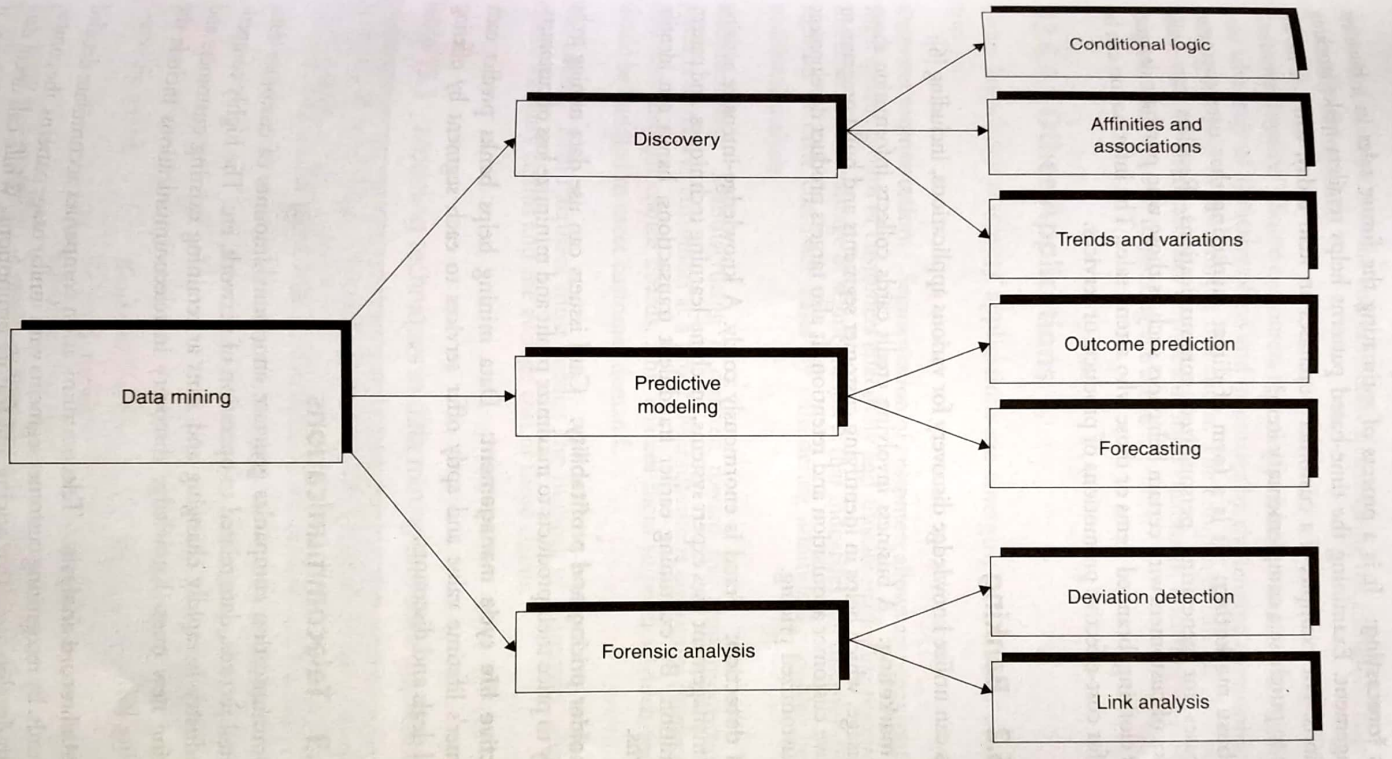

Fig. 13.4 Breakdown of data mining from a process orientation

Sales forecasting: It is a process of estimating the future sales in a business management. Examining the time-based patterns helps retailers make stocking decisions. For example, if a customer purchases an item today, when is he/she likely to purchase a complementary item?

Database marketing: It is a form of direct marketing that uses customer databases for generating personalized communications. Retailers can build profiles of customers with certain behaviors such as those who purchase designer labels clothing, branded items or those who attend sales. The information can be used for cost-effective promotions of products or services.

13.6.2 Banking

Banks can utilize knowledge discovery for various applications, including [6]:

Card marketing: A business involving credit cards collects information about card usage, which helps in identifying customer segments and build programs to improve customer acquisition and retention. It also targets product development and customized pricing.

Fraud detection: Fraud is enormously costly. A knowledge-intensive activity, fraud management uses expert systems, machine-learning techniques, and pattern recognition. By examining earlier fraudulent transactions, banks can identify patterns.

Cardholder pricing and profitability: Card issuers can use data mining technology to price their products to maximize profit and minimize loss of customers.

Predictive life cycle management: Data mining helps banks predict each customer's lifetime value and aptly offer services to each segment by offering special deals and discounts.

13.6.3 Telecommunications

Telecommunication companies generate enormous amounts of customer data, call detail records, data related to operation of network, etc. The highly competitive industry is rapidly changing and aims at retaining existing customers and attracting new ones. Knowledge discovery in telecommunications include the following [6]:

Call detail record analysis: Telecommunication companies accumulate detailed call records. By recognizing customer segments with similar usage patterns, the companies can develop attractive pricing and feature promotions. While call detail data can be used to identify suspicious calling patterns, a customer's credit score is

often incorporated into the analysis before determining the likelihood that fraud is actually taking place.

Customer loyalty: Some customers repeatedly change providers, or "churn," to take advantage of attractive offers and incentives by challenging companies. Data mining can help in minimizing this churn. The companies can use data mining to identify the characteristics of probable loyal customers. Focusing on customer lifetime value helps in improving the scene.

13.6.4 Other Applications

Other knowledge discovery applications are emerging in a variety of industries [6]:

Customer segmentation: Segmentation of customers allows an organization to target specific groups of customers effectively and allocate marketing resources to best effect. The advantage of segmentation is that industries can discover discrete segments in their customer bases by considering additional variables beyond traditional analysis.

Manufacturing: Manufacturing is directed toward the mass production of products for sale to customers at a profit. Manufacturers customize products for customers through choice boards. This enables them to predict which features should be bundled to meet customer demand.

Warranties: Manufacturers must predict the number of customers who will submit warranty claims and the average cost of those claims.

Table 13.1 Modeling techniques in data mining and their applications

Category of modeling techniques	Modeling techniques	Applications
Classification models	Neural networks, decision trees, logistic regression	Voluntary churn prediction; cross-, up-, and deep-selling
Clustering models	*K*-means, Two Step, Kohonen network/ self-organizing map.	Segmentation
Association and sequence models	A priori, generalized rule, induction, sequence	Market basket analysis, web path analysis

13.7 DATA MINING TOOLS

Various data mining tools available in the market can be grouped into the following three categories: (i) application-specific tools, (ii) integrated DSS/OLAP/DM tools, and (iii) general purpose tools [3].

13.7.1 Application-Specific Tools

The application-specific tools in Table 13.2 are rapidly gaining momentum. They offer business solutions rather than a technology searching for solution. Tools under this category include KD1 (focuses on retail), ESTARD Data Miner (focuses on marketing and insurance industry), Unica Detect, and Unica Predictive Insight (focuses on fraud detection, marketing).

Table 13.2 Application-specific tools

Unica Detect, Unica Predictive Insight http://www.unica.com/
ESTARD Data Miner http://www.estard.com/
KD1 (Knowledge Discovery One) www.kd1.com

13.7.2 Integrated DSS/LAP/DM Tools

The integrated DSS/OLAP/DM tools in Table 13.3 address real and compelling business requirement of having a single multifunctional decision-support tool that can provide management reporting, online analytical processing, and data mining capabilities within a common framework. Examples in this category are Cognos Scenario, Business Objects, etc.

Table 13.3 Integrated DSS/OLAP/DM tools

Cognos Scenario www.cognos.com/businteII/products/index.html
Business Objects www.sap.com/solutions/sapbusinessobjects/index.epx

13.7.3 General Purpose Tools

The general purpose tools in Table 13.4 are non application specific and contain a large segment of the market. Examples are SAS Enterprise Miner, IBM Intelligent

Miner, IBMSPSS Modeler, Unica Pattern Recognition Workbench, Ghost Miner, XL Miner, Weka, Rapid Miner, Oracle Darwin, etc.

Table 13.4 General purpose tools

SAS Enterprise Miner www.sas.com/technologies/analytics/datamining/miner
IBM Intelligent Miner www.01.ibm.com/software/data/iminer
IBM SPSS Modeler www.spss.com/software/modeling/modeler-pro
Unica Pattern Recognition Workbench http://www.unica.com
Ghost Miner http://cncmining.com/
XLMiner www.resample.com/xlminer/download.shtml
Weka www.cs.waikato.ac.nz/ml/weka
Rapid Miner http://rapid-i.com/
Oracle Darwin www.oracle.com/technology/documentation/darwin.html

13.8 CASE STUDIES

A case study is given in Rygielski *et al.* [6] regarding neural networks from the website of Neo Vista. NeoVista Solutions, Inc. provides comprehensive, enterprise-level data mining solutions and professional services, and their Decision Series suite of knowledge discovery tools solve data mining challenges in a variety of markets, including retail, insurance, telecommunications, and healthcare. NeoVista Solutions' Decision Series suite includes pattern discovery tools based on neural networks, clustering, genetic algorithms, and association rules.

The problem is defined as follows: A large retailer, with over $1 billion in sales, found its profits were suffering due to less-than-optimal seasonal product demand forecasting. The retailer was under-stocked on its most popular items and over-stocked on its slow-moving products at critical selling periods. NeoVista designed and implemented a solution that combined elements of the clustering and neural network technology, which entitle the retailer to automatically review

its point-of-sale history and equate store groupings to sales patterns. This solution enabled the management to explore the lowest level of detail and forecast stocking requirements for individual stock keeping units (SKUs) on a store-by-store basis. Also, a strategy combining neighborhood demographics with historic sales patterns helps the management to continuously fine-tune their replenishment system. This study enabled the management to forecast seasonal trends at the store-item level. In addition, the Decision Series tools showed that clustering similar items into actionable groups streamlined the process of ordering. As a result, the company can now predict demand for SKUs and operate a just-in-time inventory program far more effectively.

Another case study with a wireless telecommunications company with a powerful server and a data mining software package investigated data mining opportunities with the help of Data Miners. The study revealed that a major factor for churn was overcalls: New customers were making too many calls during their first month. The customers learned about the excess usage only after the arrival of the first bill, sometime during the middle of the second month. By that time, more large bills would be generated, which would make the customers even more dissatisfied. However, the data mining group identified appropriate data feeds. With a relatively simple programming, it was possible to identify these customers within days of their first overcall. This information helped the customer service center to contact customers at risk and move them to appropriate billing plans even before the first bill. This simple system was a big win for data mining, simply because having a data mining group—with the skills, hardware, software, and access—was the qualifying feature for putting together this triggering system.

SUMMARY

CRM is the strategy for building, managing, and strengthening long-lasting customer relationships. The basic idea of CRM is that it helps businesses use technology and human resources to gain insight into the behavior and value of customers. First a description of the customer life cycle is given, which helps in understanding customers as it progresses through different stages. The next section shows various business processes that are organized around the customer life cycle. From the architecture point of view, the entire CRM framework can be classified into three key components: operational CRM, analytical CRM, and collaborative CRM.

The CRM framework is used for analyzing customer profitability and improving marketing effectiveness. Many organizations have collected and stored abundant data about their customers, suppliers, and business partners. However,

the inability to discover valuable information hidden in the data prevents these companies from transforming this data into knowledge. Each data mining technique can perform one or more of the various types of data modeling, which includes association, classification, clustering, forecasting, regression, sequence discovery, and visualization. A graphical representation of CRM framework with CRM dimensions and CRM data mining models is given, which gives a better understanding about the organization of the entire framework.

A brief description about data mining applications in the fields of retail, banking, telecommunications, and other applications is also covered in the chapter. The chapter also covers a categorization of various data mining tools available in the market.

A case study using neural networks from the web site of Neo Vista is also presented to show how data mining tools solve data mining challenges in a variety of markets, including retail, insurance, telecommunications, and healthcare. The study shows how clustering and forecasting was used by the management to enhance business opportunities and improve customer relationship.

EXERCISES

A. Multiple Choice Questions

1. Customer relationship management is the strategy for building, managing, and strengthening long-lasting _____.
   a. Customer management
   b. Customer relationships
   c. Customer information
   d. None of these

2. The _____ helps in understanding customers as it progresses through different stages.
   a. Customer relationship management
   b. Customer life cycle
   c. Business intelligence
   d. None of these

3. _____, which is also referred to as customer acquisition, is the initial phase of CRM framework.
   a. Customer information
   b. Customer analysis
   c. Customer identification
   d. None of these

4. _____ are people in the target market who are not yet customers.
   a. Prospects
   b. Former customers
   c. Responders
   d. None of these