International Journal of Informatics and Communication Technology (IJ-ICT)
Vol. 3, No. 2, June 2014, pp. 88~96
ISSN: 2252-8776

Decision Support System for a Customer Relationship Management Case Study

Özge Kart*, Alp Kut*, Vladimir Radevski**
* Departement of Computer Engineering, Dokuz Eylul University
** Faculty of Contemporary Sciences and Technologies, SEE University

Article Info

Article history:

Received Feb 28, 2014
Revised May 19, 2014
Accepted May 25, 2014

Keyword:

CRM
Customer relationship
Management
Decision support

ABSTRACT

Data mining is a computational approach aiming to discover hidden and valuable information in large datasets. It has gained importance recently in the wide area of computational among which many in the domain of Business Informatics. This paper focuses on applications of data mining in Customer Relationship Management (CRM). The core of our application is a classifier based on the naive Bayesian classification. The accuracy rate of the model is determined by doing cross validation. The results demonstrated the applicability and effectiveness of the proposed model. Naive Bayesian classifier reported high accuracy. So the classification rules can be used to support decision making in CRM field. The aim of this study is to apply the data mining model to the banking sector as example case study. This work also contains an example data set related with customers to predict if the client will subscribe a term deposit. The results of the implementation are available on a mobile platform.

Corresponding Author:

First Author,
Departement of Computer Engineering,
Dokuz Eylul University,
Tinaztepe Campus, Buca 35390 Izmir, Turkey.
Email: ozge@cs.deu.edu.tr

1. INTRODUCTION

In this new era, companies have begun to give more attention to customers' personal preferences. A high consideration of the customers' *personal preferences* is considered to be a very important aspect of the nowadays functioning of the companies. CRM focuses to the managerial aspects of organizational communication to the customers and prospects. Raising the customer satisfaction is one of the main subjects of interest in this area. One to one marketing strategies began to come to the fore. These are only few contemporary actual examples putting the whole CRM concept in the focus of many researchers [1][2].

"CRM is an integrated information system that is used to plan, schedule and control the pre-sales and post-sales activities in an organization. CRM embraces all aspects of dealing with prospects and customers, including the call centre, sales-force, marketing, technical support and field service. The primary goal of CRM is to improve long-term growth and profitability through a better understanding of customer behavior. CRM aims to provide more effective feedback and improved integration to better gauge the return on investment (ROI) in these areas." [3]

Data mining has a great contribution to the extraction of knowledge and information which have been hidden in a large volume of data[4]. The concept of customer satisfaction and loyalty (CS&L) has attracted much attention in recent years. A key motivation for the fast growing emphasis on CS&L can be attributed to the fact that higher customer satisfaction and loyalty can lead to stronger competitive position resulting in larger market share and profitability [5].

Zengyou He et al. [6] implemented class outlier factors as loyalty scores for finding customers who are about to lose the loyalty segment. In their study, they considered the class outlier detection problem 'given a set of observations with class labels, find those that arouse suspicions, taking into account the class labels'. A semantic outlier is a data point, which looks regular according to data points in another class while seems irregular according to data points in the same class. They developed the idea of class outlier and proposed new solution as an extension of well known outlier detection algorithms to this case. Shin-Yuan Hung et al. [7] compared several data mining techniques that can give a 'propensity-to-churn' point at regular intervals to every mobile operator customers. They used customer demographics, billing information, contract/service status, call detail records, and service change log as customer data. As a result, both decision tree and neural network techniques gave successful churn prediction models. S.M.S. Hosseini et al. [8] proposed a new procedure, which is an expansion of RFM (Recency Frequency Monetary) model by adding one parameter, inserting WRFM-based method to K-means algorithm implemented in DM with K-optimum with respect to Davies– Bouldin Index, and then classifying customer product loyalty in B2B concept. The results provided a higher ability to the company to specify its customer loyalty in marketing strategy. R.S. Chen et al. [9] developed classification of chosen customers into clusters using RFM model to determine high-profit, gold customers. They used data mining techniques and discovered the actual consumption pattern of customers and behavioral changes in trends, which will allow management to detect potential changes of customer preference, and to prevent customer loses. Chao-Ton Su et al. [10] proposed an E-CKM model taking advantage of estimating methods based on data mining, for the development of innovative products that meet potential requirements of customers. They used web-based surveys and data mining techniques to obtain customer knowledge from different market segments.

The core part of CRM activities is to understand customer requirements and retain profitable customers. Data mining techniques such as classification, clustering etc, have important role to play in CRM applications. With data mining applications, databases, records in large companies can be converted into meaningful information. In fact through these processes important knowledge and information are extracted from the large volume of data, where they have been hidden previously [11][5].

The main purpose of this study is an implementation of Data Mining algorithm for extracting hidden information from the corporate databases and datasets that companies can use in decision making process. This valuable information is accessed via a WCF service and presented on a mobile platform [12]. As a case study, customer dataset provided from a bank was used to predict if the client will subscribe a term deposit Bayesian Classification is implemented on this dataset. A decision support system is generated to help the institution to predict the behavior of a new customer. This prediction is presented on a mobile platform. The decision support system is accessed via WCF service.

2. PROPOSED METHOD/ALGORITHM

Swift [13] described CRM dimensions as: Customer Identification, Customer Attraction, Customer Retention and Customer Development. CRM starts with customer identification. This first step is about discovering the entities that are tend to become customers or who bring the most profit to the company. In addition, it contains analyzing customers who are about to lost to the competition and how they can be won again. After identifying the segments of potential customers, companies may consume effort and resources for attracting the potential customer segments.

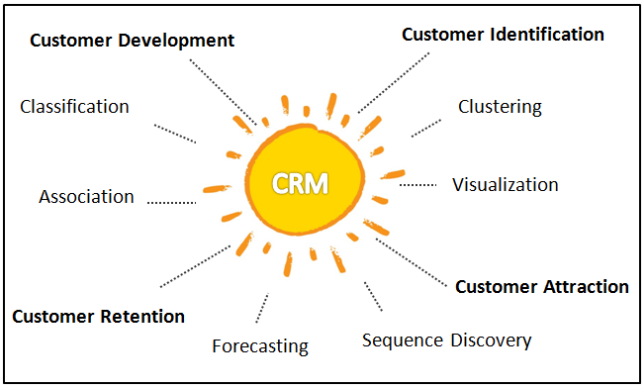
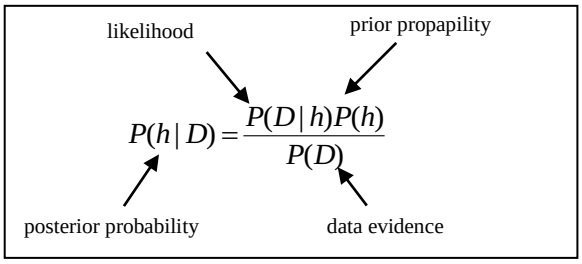


Figure 1. CRM Dimensions and Data Mining Models

In CRM, Customer retention is the main interest. Customer satisfaction, which means the comparison of customers' needs and how much they are satisfied, is the major factor for retaining customers. Components of customer retention contain loyalty programs, one-to-one marketing and complaints management. Customer development contains transaction density, values of transactions and customer profitability. Components of customer development contain up/cross selling and market basket analysis [15]. Ahmed [14] described the types of data mining model as Association, Classification, Clustering, Forecasting, Regression, Sequence Discovery and Visualization (Figure 1). In this study, Bayesian classification which is one of the well known classification algorithms is implemented.

Bayesian classification can predict class membership probabilities. Naïve Bayesian classifiers are based on Bayes' theorem. They are shown to be very powerful tool in data mining and decision support systems consequently. In this approach learning is formulated as a form of probabilistic inference, using the observations to update a prior distribution over hypotheses in Bayes classification [4].

Given a hypothesis h and data D which concerned with the hypothesis:



(1)

P(h): independent probability of h (prior probability)
P(D): independent probability of D (data evidence)
P(D|h): conditional probability of D given h (likelihood)
P(h|D): conditional probability of h given D (posterior probability)

The goal of Bayes Theorem is to specify the most probable hypothesis from the given data D. Prior probability of h, P(h): is the probability of being h is a correct hypothesis. Prior probability of D, P(D): is the probability of training data D will be observed .Conditional Probability of observation D, P(D|h): is the probability of observing data D given some world in which hypothesis h holds [11].

Naive Bayes aims to simplify the estimation problem by assuming that the different input features are conditionally independent. That is, they are assumed to be independent when conditioned on the class. Mathematically, for inputs x ∈ R^d, it is expressed as:

$$P(X|C)=\prod_{i=1}^{d} p(x_i|C) \tag{2}$$

For this reason, it is only needed to get P(Xi | C) for every possible couple of a category and a feature-value.

Bayesian categorization normally deals with categorical data. But in our case, continuous attributes exist in dataset. To classify continuous attributes, Gaussian distribution is used. If attribute X has continuous values instead of categorical values, for calculating P (Xi | Y), Gaussian distribution is used to calculate the probability of X (P(X|Y)).

For each combination of a continuous value Xi and a class value for Y, yk, has a mean, μik , and standard deviation(variance) σik based on values Xi in class yk. For estimation P(Xi | Y=yk) of this example, Gaussian distribution of Xi is defined by μik(mean) and σik (variance) depends on Y:

$$P(X_i|Y=y_k)=\frac{1}{\sigma_{ik}\sqrt{2\pi}}\exp\left(\frac{-(X_i-\mu_{ik})^2}{2\sigma_{ik}^{\ 2}}\right) \tag{3}$$

3. RESEARCH METHOD

The dataset used in this case study, is about a marketing campaign of a bank. It consists of 30902 records. Attributes of the dataset are related to customer information such as age, job, marital status etc and related with the last contact of the current campaign such as contact (communication type), duration (last contact duration, in seconds) and campaign (number of contacts made during the campaign for the client).

Bayesian classification method is implemented to analyze this dataset. The classifier will predict the customers belongs to which class that should have highest posterior probability. The customer information accumulated by a Portuguese banking institution is used to identify customers and provide decision support.

A data model is generated based upon the history of the customers in the bank. In this application, the dataset is obtained from the UCI machine learning repository (http://archive.ics.uci.edu/ml/). Aim of the classification is to predict if the client will subscribe a term deposit or not [16].

Note: Unknown values are omitted from dataset.

Banks have numerous individual retail customers. They uses CRM because of its analytical abilities. CRM helps the banks to increase the cross sell performance and manage the churn rates (customer defection rates). Data Mining models may be used to define the customers which are eager to confirm cross sell offers, which are about to be lost and what can be done to win them again.

Microsoft SQL Server 2008 is used as database management system. The Data Mining operation is applied to data. This operation is converted to a WCF service.

Architectural design of the application is as in the Figure 2. WCF service [17] connects to database and reads data from database or writes into the database. From the application, WCF service is called for computing the Bayesian classification. Phonegap[18] which is a "write once, run everywhere" platform connects to WCF services. It enables to run the application on all operating systems like IOS, Android, Windows mobile etc...

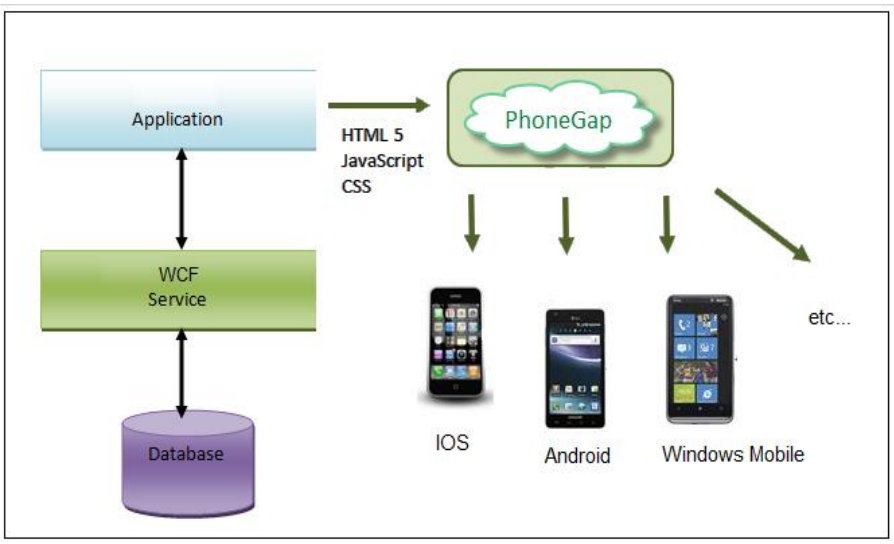


Figure 2. Architectural design

The application is developed on Android platform (Android 4.1.2) using PhoneGap (version 2.1.0) technology to connect to the service and return the result. PhoneGap is an open-source development tool for mobile cross-platform App publication that uses device-agnostic wrappers like HTML, Javascript, and CSS, that can be rapidly deployed on Android, Blackberry, and iPhone platforms.

Bayesian calculation is done by using the dataset and the probabilities of each variables are written into database when the method BayesianCalculation() is called from WCF service.

At first, model is constructed with training set and tested with test set. Figure 3 shows the model construction diagram. In training set, output classes of samples are known. There are categorical and continuous attributes in dataset. Figure 3 includes some of them.

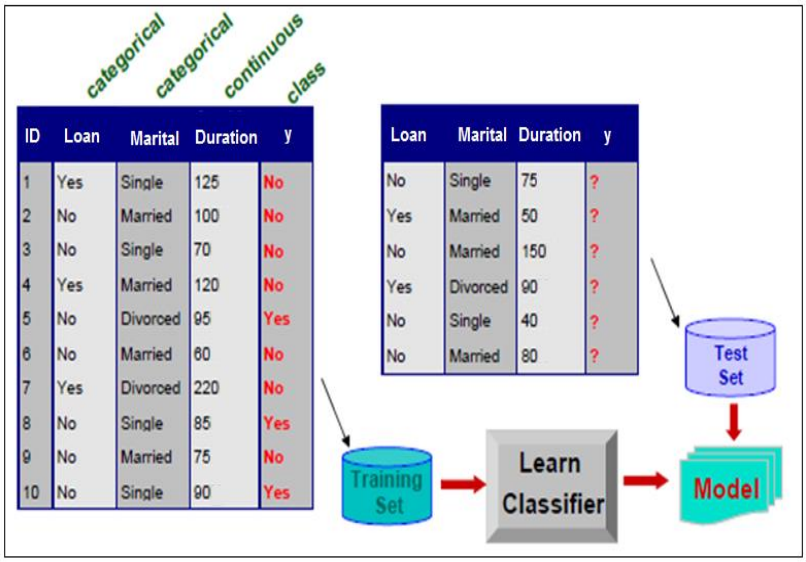


Figure 3. Model construction

Table 1. Probabilities of categorical attributes

		No	Yes
job	management	0,233450798	0,258865248
	blue-collar	0,189799553	0,116134752
	technician	0,182865371	0,163342199
	services	0,08798454	0,067375887
	admin.	0,113068849	0,122340426
	unemployed	0,029707097	0,040336879
	entrepreneur	0,034064643	0,022163121
	housemaid	0,02932818	0,020833333
	retired	0,045697397	0,103501773
	self-employed	0,037626464	0,035682624
	student	0,016407108	0,049423759
marital	single	0,279830245	0,365469858
	married	0,607328256	0,520611702
	divorced	0,112841499	0,11391844
education	secondary	0,524800121	0,476507092
	primary	0,142397029	0,111702128
	tertiary	0,332802849	0,41179078
default	no	0,981963548	0,992464539
	yes	0,018036452	0,007535461
housing	no	0,475995605	0,664893617
	yes	0,524004395	0,335106383
loan	no	0,8213785	0,91001773
	yes	0,1786215	0,08998227
contact	cellular	0,911181842	0,922429078
	telephone	0,088818158	0,077570922

Table 2. Mean and variance of continuous attributes

	No		Yes	
	Mean	Variance	Mean	Variance
age	40,77863666	10,37905095	41,73891844	13,64783701
balance	1402,718086	3086,941414	1870,782358	3593,489828
duration	218,3497783	205,4261557	507,03125	370,4182292
campaign	2,861960517	3,092719771	2,105718085	1,831433975

Some of the inputs such as age, balance, duration and campaign have continuous values in the dataset. So they are handled by Gaussian distribution. Probabilities of all categorical attributes when output class y is 'No' and output class y is 'Yes' are calculated and shown in Table 1 for the dataset used in the application. Mean and variance of each continuous attributes when output class y is 'No' and output class y is 'Yes' are calculated and shown in Table 2.

As an example, following sample customer data is an evidence for Bayesian Classifier.

Table 3. Example customer data

age	job	marital	education	default	balance	housing	loan	contact	duration	campaign	y
34	self-employed	married	tertiary	no	1045	yes	yes	telephone	65	2	?

Probabilities of each classes for given sample customer data (shown in Table 3) and Gaussian distributions for continuous attributes and each general class probabilities (P(yes) and P(no)) are all multiplied for finding the probable class of given customer. All values and result of multiplication are shown in Table 4. According to the table, Probability of No is higher than Probability of Yes.

Table 4. The terms for calculating percentage of results

	No	Yes
self-employed	0,037626464	0,035682624
married	0,607328256	0,520611702
tertiary	0,332802849	0,41179078
no	0,981963548	0,992464539
yes	0,524004395	0,335106383
yes	0,1786215	0,08998227
telephone	0,088818158	0,077570922
G(age= 34)	0,031054774	0,024890009
G(balance= 1045)	0,000128371	0,000108125
G(duration= 65)	0,001469764	0,000528436
G(campaign= 2)	0,124080091	0,217467943
no/y - yes/y	0,853994758	0,146005242
RESULT	**3,85451E-14**	**8,01878E-16**

The percentage of results may be calculated as following: (X*100)/(X+Y)
When it is applied to the sample data: (3,85451E-14 * 100) / (3,85451E-14 + 8,01878E-16)

According to result of the calculation it can be expressed that the output class of the sample customer is no with 97% probability.

When the application is run, following screen shown in Figure 4 appears and user enters the parameters.

Figure 4. Application's Initial Screen

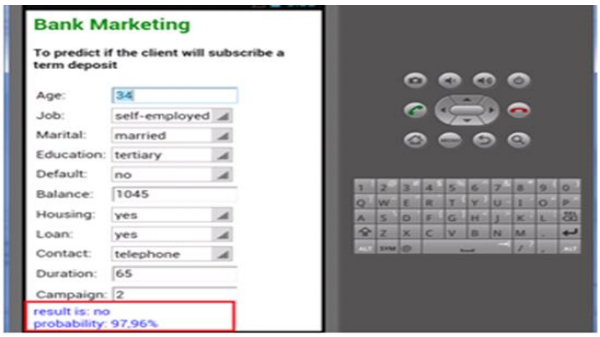


Figure 5. Result of the classification and probability of the result

After clicking the Show Result button, program evaluates the values entered by user and shows the classification result. "Classify" method is called from WCF when the button is clicked. And the service returns the result. Figure 5 shows the result:

4. RESULTS AND ANALYSIS

Accuracy rate of the model used in this application is computed by Cross Validation. Cross validation allows using the whole dataset in computing. The dataset is divided into two parts randomly. First part is used for model construction. The model is tested with second part of dataset and the accuracy rate is

computed. After that, a model is constructed with second part and tested with first part. Accuracy rate is computed. Finally, the model is constructed using whole dataset. Average of the computed accuracy rates is the accuracy rate of the constructed model. Figure 6 shows the Cross Validation:

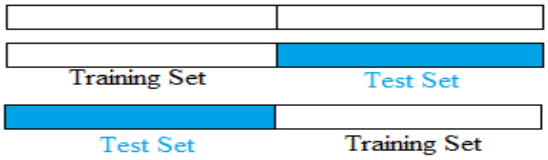


Figure 6. Cross Validation

Table 5. First cross validation results

		Classified As	
		Yes	No
Actual Class Value	Yes	555	351
	No	1005	13540

In this study, the first half of dataset is used to train and second half is used to test. After that, the second half of dataset is used to train and first half is used to test the algorithm. Cross validation showed that the accuracy rate of the model is 84%.

First test results are showed in Table 5. The first part of dataset is used for testing. There are 906 samples which belong to actual class value 'yes'. After the test, 555 of them are classified as 'yes' which is the right result and 351 of them are classified as 'no'. There are also 14545 samples which belong to actual class value 'no. After the test, 13540 of them are classified as 'no' which is the right result and 1005 of them are classified as 'yes'.

As a result, accuracy rate of first training is;
Truly classified samples: 555+13540=14095
All samples in dataset: 15451
Accuracy rate: (14095 * 100) / 15451 ~= 91%

Second test results are showed in Table 6. The second part of dataset is used for testing. There are 3606 samples which belong to actual class value 'yes'. After the test, 552 of them are classified as 'yes' which is the right result and 3054 of them are classified as 'no'. There are also 11846 samples which belong to actual class value 'no. After the test, 11514 of them are classified as 'no' which is the right result and 332 of them are classified as 'yes'.

As a result, accuracy rate of second training is;
Truly classified samples: 552+11514=12066
All samples in dataset: 15452
Accuracy rate: (12066* 100) / 15452 ~= 78%

Table 6. Second cross validation results

		Classified As	
		Yes	No
Actual Class Value	Yes	552	3054
	No	332	11514

Average accuracy rates of first and second tests give the accuracy of the implementation which is 84, 5 %.

5. CONCLUSION

Today information is vital for companies and Customer Relationship Management and Data Mining techniques aim to determine the customers with high profitability. The information obtained by applying sophisticated Data mining techniques to CRM problems have a strategic importance for companies. An important competitive advantage is evident as implication of using such systems at the decision making level.

In this paper we present an analysis of a massive volume of customer data which is afterwards classified based on the customer behaviors. Naive Bayesian classification is used as classifier to predict if the client will subscribe a term deposit. The classifier reported highly acceptable accuracy like 84.5% for all the tested data by doing cross validation. The accuracy is determined as the percentage of the correctly classified instances from the test set. In other words, classification centers around exploring through data objects (training set) to find a set of rules which determine the class of each object according to its attributes. Since the accuracy of the model is acceptable, the model can be used to classify data tuples whose class labels are not known. The classification rules can be used to support decision making for achieving a good CRM for businesses.

In this paper, the data obtained from a bank is analyzed. Bayesian classification is applied to the data. Bayesian classification is implemented as a WCF service. From an Android device, this service is called and result of classification of a new customer data is showed.

Dataset is stored in Microsoft SQL Server 2008. Bayesian classification method is implemented on a WCF (Windows Communication Foundation) service project using Microsoft Visual Studio 2010. The result of the implementation is presented on Android platform using Phonegap technology.

The developed service can handle different datasets with different number of attributes. Hereby this model can also be used for other type of companies to predict their customer behaviors with specific historical data. In this application, the bank can predict if a customer will subscribe a term deposit or not. It can manage direct marketing campaigns using this prediction. The results can be accessed anytime and anywhere through a mobile device.

With application of CRM in mobile platform, changes and updates can be done seamlessly from anywhere and anytime and with no delay.

REFERENCES

[1] Gao, Hua; , "Customer Relationship Management Based on Data Mining Technique," *International Conference on E -Business and E –Government*, 1–4, 2011
[2] Padmanabhan, B.; Tuzhilin, A.; , "On the Use of Optimization for Data Mining: Theoretical Interactions and eCRM Opportunities [J]," *Management Science*, 49,10, 271–343, 2003
[3] Petersen, G.; , "Customer Relationship Management In ROI: Building the CRM Business Case," Xlibris, 2003
[4] Han, J.; Kamber, M.; , "Data Mining: Concepts and Techniques [M]," CA: Morgan Kaufmann Publishers. San Francisco, 2001
[5] Kuykendall, L.; , "The data-mining toolbox [J]," *Credit Card Management*,12, 6, 30–40, 1999
[6] He Z.; Xu X.; Huang J.Z.; Deng S.; , "Mining class outliers: concepts, algorithms and applications in CRM," *Expert Systems with Applications*, 27, 681–697, 2004
[7] Hung S.Y.; Yen D.C.; Hsiu-Yu Wang; , "Applying data mining to telecom churn management," *Expert Systems with Applications* 31, 515–524, 2006
[8] Hosseini S.M.S.; Maleki A.; Gholamian M.R.; , "Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty," *Expert Systems with Applications*, 37, 5259–5264, 2010
[9] Chen R.S.; Wu R.C.; Chen J.Y.; , "Data Mining Application in Customer Relationship Management Of Credit Card Business," *International Computer Software and Applications Conference on*, 2005

[10] Sua C.T.; Chen Y.H.; Sha D.Y.; , "Linking innovative product development with customer knowledge: a data-mining approach," *Technovation*, 26, 784–795, 2006
[11] Carrier, C. G.; Povel, O.; , "Characterising data mining software," *Intelligent Data Analysis*, 7, 181–192, 2003
[12] Chung, H.M.; Gray, P.; , "Data mining [J]," *Journal of MIS*,16(1), 11–13, 1999
[13] Swift, R.S.; , "Accelerating customer relationships: Using CRM and relationship technologies," Upper Saddle River. N.J.: Prentice Hall PTR., 2001
[14] Ahmed, S. R.; , "Applications of data mining in retail business," *Information Technology: Coding and Computing*, 2, 455–459, 2004
[15] Ngai E.W.T.; Xiu L.; Chau D.C.K.; , "Application of data mining techniques in customer relationship management: A literature review and classification," *Expert Systems with Applications*, 36, 2592–2602, 2009
[16] Moro, S.; Laureano, R.; Cortez, P.; , "Using Data Mining for Bank Direct Marketing: An Application of the CRISP-DM Methodology," *European Simulation and Modelling Conference on*, 2011
[17] Löwy, J.; , "WCF Essentials In Programming WCF Services (1st. ed.)(1)," USA: O'Reilly, 2007
[18] Ghatol, R.; Patel, Y.; , "Beginning PhoneGap," Apress, 2012

BIOGRAPHIES OF AUTHORS

Özge Kart received her master degree in computer engineering from Dokuz Eylul University in 2013. Currently, she is a research assistant at the Department of Computer Engineering, Dokuz Eylul University in Turkey. Her research interests include data mining in large databases, social media mining, decision support systems, and industrial applications.

Alp Kut is a full professor of computer engineering at Dokuz Eylul University. He is head of the Department of Computer Engineering of Dokuz Eylul University since the fall of 2003. His works include data mining in databases, database management systems and distributed systems. He has many publications on a variety otopics, including, web-based systems and parallel systems.

Dr. Vladimir Radevski is Associate Professor at the Faculty of Contemporary Sciences and Technologies, South East European University in Tetovo, Republic of Macedonia, and director of its Skopje Study Center. He received his PhD from University of Paris 13, France and has been teaching at several international programs and universities. His research interests are in the domains of semantic web and artificial intelligence.