Third International Conference
Modelling and Development of Intelligent Systems
October 10 - 12, 2013
"Lucian Blaga" University
Sibiu - Romania

Example of developing a loyalty program using CRM, SQL-queries and Rapid Miner tool

Iryna Zolotaryova, Iryna Garbuz, Mykhailo Dorokhov

Abstract

The object of paper is the implementing of loyalty programs using methods for intelligent processing of information and analysis the effectiveness of such a program. The subject is researching the loyalty program effectiveness based on Data Mining for the company "Auto-Maximum". The aim is to study the effectiveness of loyalty program based on Data mining. Results of performance are analyzed and described concept of direct marketing and loyalty programs features, research loyalty program development tools and algorithms for data mining. The results can be used in research institutions different companies that are aimed at marketing.

1 Introduction: defining main points of a loyalty program

The process of developing a loyalty program we start with describing the company for that we will develop the loyalty program. The firm "Auto-Maximum" has over 15 years of successful experience in the Ukrainian market of cosmetic service center in Kharkiv. Today the station is a recognized leader in this field, providing services to the population and not just implementing technology cosmetic service of cars, but also a whole range of new products, which includes everything you need for a successful business organization.

At the station the following services are offered for the cosmetic maintenance of vehicles: repair chips and cracks, installation (replacement) auto glass; toning of car windows; book lights; removing dents without paint; depth cleaning salon; professional polished body and glasses; bumper repair; accurate color matching; touch up minor defects; painting of individual components and car parts; installation of xenon light.

Before starting to develop the loyalty program itself, it is necessary to define the basic components of the program and the criteria.

Objectives. As with any project, the development of loyalty program begins with setting goals. A common mistake here is the lack of clear definitions. Of course, the key is to increase customer loyalty. To do this, you must define the parameters by which it will be possible to evaluate its success and effectiveness. If we consider a comprehensive loyalty, let, for example, indicators of transactional loyalty will be to increase the number of clients, done depleting repeat purchase, 40%, and the index increase customer satisfaction by 30%. Of course, that prior to the development of loyalty program we define for themselves their original positions in order to understand, from what we have to make a start and what we want to achieve.

Target Audience. Program is aimed at anyone? Who do we want to keep? Whose loyalty we want to raise? In accordance with the determined conditions, and the program. For example, we want our clients are committed to those who commit at least 5 purchases per year (if it comes about, say, a clothing store), or uses our products for at least 1 time per week (if we're talking about food). In this case, once again I would like to stress the importance of market research and identify key basic parameters prior to the development of loyalty program.

Type of program. The best known and most widely used instrument - discount programs. Their essence is to provide customer benefits in the form of a refund of the paid value of the goods at the moment of purchase. There is a purely material gain. The second, also quite common - among lotteries that have made certain purchases in a given period of time. And even if the prize is not quite need a man - all the same emotions that accompanied the receipt of the prize, will leave a positive impression of your company. Another variety, have been gaining in popularity - accumulative discount programs. They benefit directly dependent on the participant: the more often and for a large amount buy, the more benefit you get. The fourth type - the bonus program. Their essence is that when shopping, the customer receives a certain conditional points, having accumulated a certain number of which he has the right to exchange them for goods or services on your own. By the way, on a product (as opposed to the lottery) this person is wanted and needed.

The other important component of customer loyalty programs are gift certificates in the form of a plastic card. This option is much more practical and presentable than usual, the paper certificate. Gift card will reflect your corporate identity, advertising your organization, and after use can be presented as a discount card, or used in the prize draw.

Privileges. The most complex, interesting and creative stage - the definition of what to offer the customer, in addition to the main component - the bonuses. And there scope for creativity is so wide that you can lose sight of the most important - the needs of customers! It is they who must determine the entire range of additional benefits. We can distinguish three stages of its preparation: a preliminary study of the resulting list on a limited sample of clients; a large-scale survey; creative development of all possible privileges. The final list is generated taking into account factors such as the feasibility of the franchise, the competence of the company in its implementation, and cost.

Financial concept. The most sensitive issue is related to the assessment of future expenses for the program and they should be met. Thus, the costs associated with the cost of accrual of bonuses, discounts, production of advertising and souvenir production, club cards, acquisition or development of specialized software, compensation of employees responsible for the operation of the program. Covering the costs may come at the expense of the annual contribution of participants, foreclosures club card, etc. Also, there are bottlenecks, such as, for example, the account of the bonuses in the company's accounting or determination techniques for tracing the influence of loyalty programs for sales, profitability and revenue. All this requires a thorough study before implementation software.

Management and Communication. There are three areas of communication: between the company and customers, between the company and the external environment, as well as within the company. The system of interaction with the participants can be built, for example, on the basis of telephone communication with a call-center manager and loyalty programs, newsletters, holiday greetings, prompt responses to emails and customer complaints on the message in the forum. Communication with the external environment may include media publications, participation in conferences on loyalty marketing. Intra-corporate communications related to the interaction of all departments of the company, assess the effectiveness of its activities, etc.

Technology. Requires careful attention to the creation of a single control center loyalty program, coordinating key areas of its operation, the center for the processing of incoming calls; IT-system maintenance program; logistics system and algorithm implementation procedures of the program.

Database. Loyalty program - a great tool to collect and store data about customers. Before starting the program is to determine what data and how much should be entered into the database, how and with what frequency analyzed what this will require resources, both technical and human. Unfortunately, many companies implementing loyalty programs and have extensive databases, they are used inefficiently. The reasons for this - lack of knowledge of how to efficiently use the information collected, how to develop individual offers for each customer segment, the technical complexity of the implementation of analytical processes, the unreliability of data, etc.

Closing the program. The problem to which few people paid attention to when you start the program because of optimistic beginning. But do not forget that every project has its own life cycle. There may come a time when the program will cease to be effective. Decide in advance with the critical exponents for the attainment of which is necessary to minimize the program. Will it be transformed into something new normed? If not, how staff will be disbanded and its service? How does the database will be used? These and other questions worth pondering in advance.

The final concept of the program include such elements: objectives (the increase in the number of repeat orders and increase the return up to 15%), target audience (group of clients with large number of orders for a large sum), type of program (discount), management and communication (between company and customers), technology (Data mining. CRM, Rapid Miner).

2 Developing of the software for collecting information

To develop our loyalty program we must have the information about: information about range of services provided by this company; general information about clients; information about frequence of services and sum for them; information about cooperation clients and the firm etc.

So, to answer to these questions we create a special program. The main aim of this program is to provide an opportunity to collect information about clients and service. This program is a kind of Customer Relation Management. The manager of the firm record information about every service, order etc during the day. As result, we accumulated data about clients, services and provided services. Moreover, this program provides an opportunity to get information for making decision. But, first of all, let us consider the model of the data base which our application bases on.

The normal ER-model (Fig. 1) consists of 4 entities: “Customer”, “Service”, “Sale header” ,“Sale detail”. Entity “Customer” includes information about clients such as Name, Surname, Phone, Fax etc. Entity “Service” contains list of services provided by the company. And entities “Sale header” and “Sale detail” contain information about orders: date, client, what services were ordered and the sum of order. There are connections between entities such as "one to many".

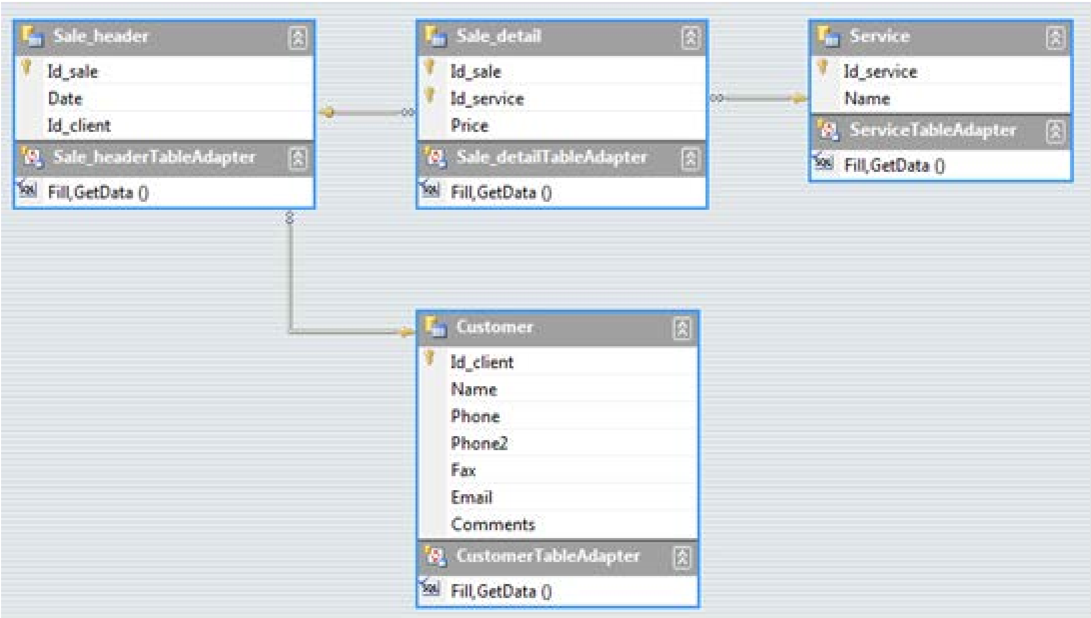

Figure 1. Logic model of data base

The logic model matches completely with physical model of data base (Fig. 2).

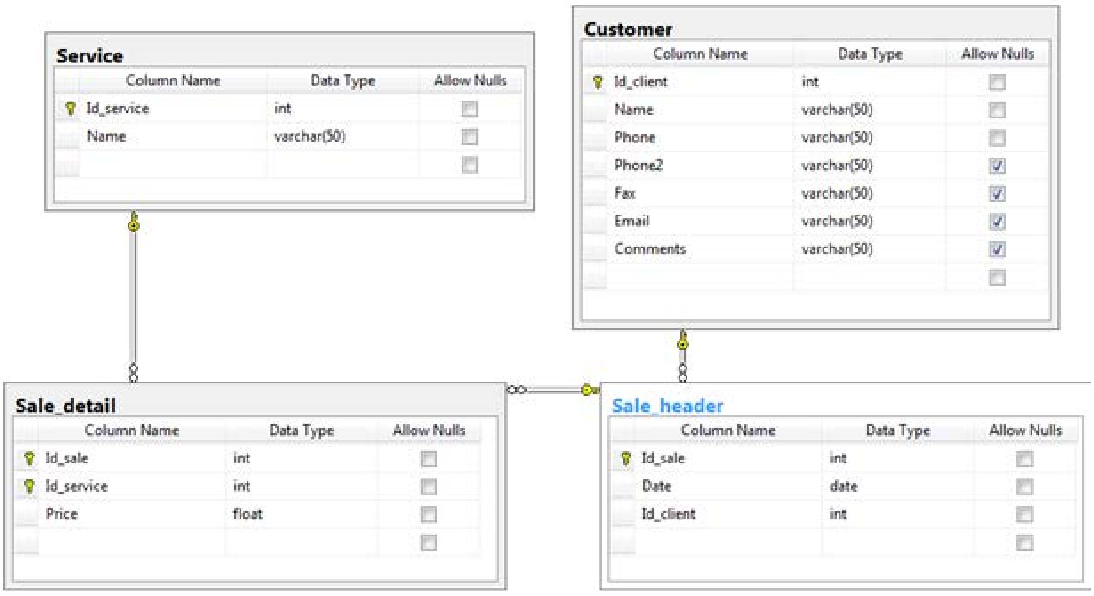

Figure 2. Physical model of data base

Also we construct graphic user interface which consists of units (Clients, Services, Orders and Reports) with main window and additional dialog windows. Also, the program provides an analytical functionality. With the help of this program, a manager can get useful information for making decision such as monthly orders, quantity of services for a month, 3 months and year.

3 Analyzing of tools for Data Mining

At first we analyzed (Fig. 3) statistics about data mining/analytic tools which are most popular in the past 12 months for a real project. Direct comparison votes may not be representative, because of different verification strategies in 2011 and 2012, but clearly that the leading open source tools were Rapid Miner, R, and KNIME. Among commercial tools, the top tools were SAS, MATLAB, and IBM SPSS Modeler (former Clementine).

Rapid Miner Tool is one of the leading open source data mining software suites. With more than 400 data mining modules or operators, it is one of the most comprehensive and most flexible data mining tools available. With over 10,000 downloads from SourceForge.net each month and more than 300,000 downloads in total, it is also one of the most widespread-used data mining tools. According to polls by the popular data mining web portal KDnuggets.com among several hundred data mining experts, Rapid Miner was the most widely used open source data mining tool and among the top three data mining tools overall in 2011 and 2012.

Rapid Miner supports all steps of the data mining process from data loading, pre-processing, visualization, interactive data mining process design and inspection, automated modeling, automated parameter and process optimization, automated feature construction and feature selection, evaluation, and deployment. Rapid Miner can be used as stand-alone program on the desktop with its graphical user interface (GUI), on a server via its command line version, or as data mining engine for your own products and Java library for developers.

Rapid Miner is provided by Rapid-I, an open source data mining and business intelligence software and consulting company based in Dortmund, Germany. Rapid Miner has users in more than 30 countries and Rapid-I serves customers on four continents. Rapid-I won the Open Source Business Award 2008.

Which data mining/analytic tools you used in the past 12 months for a real project (not just evaluation) [912 voters]	
RapidMiner (345)	37.8%
R (272)	29.8%
Excel (222)	24.3%
KNIME (175)	19.2%
Your own code (168)	18.4%
Pentaho/Weka (131)	14.3%
SAS (110)	12.0%
MATLAB (84)	9.2%
IBM SPSS Statistics (72)	7.9%
Other free tools (67)	7.3%
IBM SPSS Modeler (former Clementine) (67)	7.3%
Microsoft SQL Server (63)	6.9%
Statsoft Statistica (57)	6.2%
Other commercial tools (56)	6.1%
SAS Enterprise Miner (50)	5.5%
Zementis (34)	3.7%
Orange (25)	2.7%
Oracle DM (19)	2.1%
KXEN (19)	2.1%
Salford CART Mars other (15)	1.6%
VisuaLinks (12)	1.3%
Viscovery (10)	1.1%
Angoss (8)	0.9%
TIBCO Insightful Miner (7)	0.8%
Miner3D (7)	0.8%
REvolution Computing (4)	0.4%
Megaputer Polyanalyst/TextAnalyst (3)	0.3%
Portrait Software (2)	0.2%
Data Applied (2)	0.2%
Centrifuge (2)	0.2%

Figure 3. Survey Results

Besides Rapid Miner, Rapid-I offers data mining training courses, consulting, professional support, adaptations and extensions of Rapid Miner, individual software development, integration, and other data mining services.
Some of important features of Rapid Miner are listed as follows: Rapid Miner (previously known as YALE) is based on modular operator concept which facilitates rapid prototyping of data mining processes by way of nesting operator chains and using complex operator trees; a large numbers of operators (more than four hundred) defined in Rapid Miner along with its plugins cover nearly all key aspects of Data mining handling data transparently and without the need to know the different data formats or different data views; Rapid Miner provides flexible operators for data input and data output in different file formats such as excel files, files, SPSS files, data sets from well-known databases such as Oracle, my SQL, Microsoft SQL Server, Sybase, and dBase.
Owing to the modular operator concept, the data mining processes are optimized because, by substituting or replacing one particular operator at a time and leaving rest of the data mining process design untouched, its performance can be evaluated.
Rapid Miner follows a multi-layered data view concept which enables it to store different views

on the same data table and therefore facilitates cascading multiple views in layers through a central data table. Rapid Miner data core is typically similar to a standard database management system. Rapid Miner has a flexible interactive design which lets user to additional meta data on the available data sets to enable automated search and optimized preprocessing which are both needed for an effective data mining processes. Rapid Miner also acts as a powerful scripting language engine along with a graphical user interface.

Since using Rapid Miner, data mining processes are designed as operator trees defined in XML, where operators are not defined in a graph layout so as to be positioned and connected by a user. Therefore data flow normally follows "depth first search," resulting in optimization of data mining processes.

4 Clustering of the customer database and choosing the target group

Clustering is a data mining method that analyzes a given data set and organizes it based on similar attributes [1-4]. Clustering can be performed with pretty much any type of organized or semi-organized data set, including text, documents, number sets, census or demographic data, etc. The core concept is the cluster, which is a grouping of similar objects. Clusters can be any size – theoretically, a cluster can have zero objects within it, or the entire data set may be so similar that every object falls into the same cluster [5-8].

As one of the main benefits from the application of the loyalty program is an opportunity to focus on a specific group of customers who make the most out of to a company, so an important point of their promotion effectiveness is the process of a segmentation of client base and selection of the most attractive consumers. And then you can build a relationship with clients in certain segments that have common characteristics. This allows you to create specific marketing programs.

To cluster clients of the company we use annual report presented by our CRM. This report contains information about how many services for what amount of money the clients made during the half of year. So, in this analytical report we summarize quantity of services and their cost for each client separately. This information helps us choose the most interesting group of clients for us. The most interesting group for us is the group that ordered more services and for the largest sum than the average amount.

Clustering is a great first step to use when looking at a large data set. In order to perform clustering, some setup is required. First, the data set must be prepared and cleaned (Replace Missing Values). Second, the numerical data must be separated into a subset (Work on Subset). Third, the clustering algorithm must be defined and applied (Clustering). Lastly, the output must be examined in order to check for quality and usefulness.

There are many different types of clustering algorithms. Some of the most advanced methods in 2012 revolve about support-vector models (SVM), the CLOPES and COBWEB algorithms, or clustering by expectancy. Unfortunately, these clustering methods require an intense amount of computing power. The K-Means algorithm is the simplest clustering method and also probably the most efficient given limited technology. It may not be cutting edge, but the results are still valid and useful for any data miner looking for the broadest of insights.

First of all, we import our report to Rapid Miner. Then with the help of Rapid Miner's operators we build the model of clustering process (Fig. 4). Our model contains 3 operators: Retrieve as input, nominal to numerical as a convertor and clustering operator as the main component. We use K-means algorithm and determine that all data will be divided into 5 clusters (k=5). We choose k-means algorithm for several reasons: as we have not large database this algorithm works really fast; simplicity of algorithm's logic; opportunity to choose quantity of clusters etc.

Retrieve reads an object from the data repository. Nominal to Numerical operator changes the type of selected non-numeric attributes to a numeric type. It also maps all values of these attributes to numeric values. Clustering operator performs clustering using the k-means algorithm. Clustering is concerned with grouping objects together that are similar to each other and dissimilar to the objects belonging to other clusters. Clustering is a technique for extracting information from unlabeled data. K-means clustering is an exclusive clustering algorithm i.e. each object is assigned to precisely one of a set of clusters.

Figure 4. Clustering Process Model

The clustering process results are grouped into 5 clusters (Fig. 5). Centroid result table presents the average characteristics of each cluster. We will choose cluster 4, reasons are followed: this cluster has high level of average amount of orders; this cluster has high level of average amount of money spent for orders; this cluster has 28 items, so it is not very large. But according to 80/20 Pareto theory, 20 percent of clients generate 80 percent of profit; this cluster has the most close location of its items.

Text View   Folder View   Graph View   Centroid Table   Centroid Plo

Attribute	cluster_0	cluster_1	cluster_2	cluster_3	cluster_4
Количество	1.847	10.857	4.882	3.188	7.214
Сумма	645.678	9928.571	3565.263	1873.419	6505.357

```
Cluster 0: 236 items
Cluster 1: 7 items
Cluster 2: 76 items
Cluster 3: 160 items
Cluster 4: 28 items
Total number of items: 507
```

Figure 5. Cluster Model Results and Centroid Table

Market basket analysis aims to detect relationships or associations between specific items in a large "catalog" of objects. A simple example would be the occurrence of diapers and baby formula in the same sales transaction. However the real value from association rules analysis is finding connections between seemingly non-intuitive items: it would be indeed surprising (and valuable) to the retailer if it was found that there was a strong association between beer and diapers! But how do you know if this association rule has statistical value and not just some random coincidence? One way to test this is by calculating the Lift Ratio.

We discussed a couple of key measures which are needed to correctly apply association rules for market basket analysis: Support and Confidence. A third quantity of importance is lift. Lift Ratio is defined as the ratio between the confidence of a rule and a benchmark confidence. Benchmark confidence is simply the confidence of the two items if they were independent. Lift ratio higher than 1.0 indicates that there is some value to the rule.

There are three steps for generating usable recommender systems using a data mining tool such a Rapid Miner: data preparation, rule generation and rule application. Data may consist of products sold, transaction details and customer ids if we are discussing retail. Data has to be transformed into a format where each row is a separate transaction id (or session id) and each column is a product sold (page viewed). The cells of this table will be 0's or 1's depending upon whether that particular product was sold during a given transaction. To transform data we used the SQL query.

The well-known algorithm proceeds in two steps within the support-confidence framework (minimum support and confidence thresholds have to be fixed by the user) in order to extract association rules:

• find frequent item sets (the sets of items which occur more frequently than the minimum support threshold) with the frequent item set property (any subset of a frequent item set is frequent; if an item set is not frequent, none of its supersets can be frequent)for efficiency reasons. Thus starting from k = 1, generates item sets of size k + 1 from frequent item sets of size k,

• generate rules from frequent item sets and filter them with the minimum confidence threshold.
Our process model consists of four elements: input Retrieve, Numerical to Binominal for conversion the FP-Growth element for calculating frequencies of items in the data and Create Association Rules element for creating rules. Then, we describe all components of our model.

Retrieve (Rapid Miner Core) can be used to access the repositories. It should replace all file access, since it provides full metadata processing, which eases the usage of Rapid Miner a lot. In contrast to accessing a raw file, it provides the complete metadata of the data, so all metadata transformations are possible.

Numerical to Binominal changes the type of the selected numeric attributes to a binominal type. It also maps all values of these attributes to corresponding binominal values. The Numerical to Binominal operator changes the type of numeric attributes to a binominal type (also called binary). This operator not only changes the type of selected attributes but it also maps all values of these attributes to corresponding binominal values. Binominal attributes can have only two possible values i.e. 'true' or 'false'. If the value of an attribute is between the specified minimal and maximal value, it becomes 'false', otherwise 'true'. Minimal and maximal values can be specified by the min and max parameters respectively. If the value is missing, the new value will be missing. The default boundaries are both set to 0.0, thus only 0.0 is mapped to 'false' and all other values are mapped to 'true' by default.

FP-Grow the efficiently calculates all frequent items from the given Example Set using the FP-tree data structure. It is compulsory that all attributes of the input Example Set should be binominal. In simple words, frequent item sets are groups of items that often appear together in the data. It is important to know the basics of market-basket analysis for understanding frequent item sets.

The market-basket model of data is used to describe a common form of a many-to-many relationship between two kinds of objects. On the one hand, we have items, and on the other we have baskets, also called 'transactions'. The set of items is usually represented as set of attributes. Mostly these attributes are binominal. The transactions are usually each represented as examples of the Example Set. When an attribute value is 'true' in an example; it implies that the corresponding item is present in that transaction. Each transaction consists of a set of items (an item set). Usually it is assumed that the number of items in a transaction is small, much smaller than the total number of items i.e. in most of the examples most of the attribute values are 'false'. The number of transactions is usually assumed to be very large i.e. the number of examples in the Example Set is assumed to be large. The frequent-item sets problem is that of finding sets of items that appear together in at least a threshold ratio of transactions. This threshold is defined by the 'minimum support' criteria. The support of an item set is the number of times that item set appears in the Example Set divided by the total number of examples. The 'Transactions' data set at "Samples/data/Transactions" in the repository of Rapid Miner is an example of how transactions data usually look like.

The discovery of frequent item sets is often viewed as the discovery of 'association rules', although the latter is a more complex characterization of data, whose discovery depends fundamentally on the discovery of frequent item sets. Association rules are derived from the frequent item sets. The FP-Growth operator finds the frequent item sets and operators like the Create Association Rules operator uses these frequent item sets for calculating the rules.

This operator calculates all frequent item sets from an Example Set by building a FP-tree data structure on the transaction data base. This is a very compressed copy of the data which in many cases fits into main memory even for large data bases. All frequent item sets are derived from this FP-tree. Many other frequent item set mining algorithms also exist e.g. the Apriori algorithm. A major advantage of FP-Growth compared to Apriori is that it uses only 2 data scans and is therefore often applicable even on large data sets.

Please note that the given Example Set should contain only binominal attributes, i.e. nominal attributes with only two different values. If your Example Set does not satisfy this condition, you may use appropriate preprocessing operators to transform it into the required form. The discretization operators can be used for changing the value of numerical attributes to nominal

attributes. Then the Nominal to Binominal operator can be used for transforming nominal attributes into binominal attributes.
Please note that the frequent item sets are mined for the positive entries in your Example Set, i.e. for those nominal values which are defined as positive in your Example Set. If data does not specify the positive entries correctly, you may set them using the positive value parameter. This only works if all your attributes contain this value.
This operator has two basic working modes:

- finding at least the specified number of item sets with highest support without taking the 'min support' into account. This mode is available when the find min number of item sets parameter is set to true. Then this operator finds the number of item sets specified in the min number of item sets parameter. The min support parameter is ignored in this case.
- finding all item sets with a support larger than the specified minimum support. The minimum support is specified through the min support parameter. This mode is available when the find min number of item sets parameter is set to false.

5 Criterion and creating of Association Rules

Create Association Rules generates a set of association rules from the given set of frequent item sets. Association rules are if/then statements that help uncover relationships between seemingly unrelated data. An example of an association rule would be "If a customer buys eggs, he is 80% likely to also purchase milk." An association rule has two parts, an antecedent (if) and a consequent (then). An antecedent is an item (or item set) found in the data. A consequent is an item (or item set) that is found in combination with the antecedent.
Association rules are created by analyzing data for frequent if/then patterns and using the criteria support and confidence to identify the most important relationships. Support is an indication of how frequently the items appear in the database. Confidence indicates the number of times the if/then statements have been found to be true. The frequent if/then patterns are mined using the operators like the FP-Growth operator. The Create Association Rules operator takes these frequent item sets and generates association rules. Such information can be used as the basis for decisions about marketing activities such as, e.g., promotional pricing or product placements. In addition to the above example from market basket analysis association rules are employed today in many application areas including Web usage mining, intrusion detection and bioinformatics.
Let describe specifies the criterion which is used for the selection of rules.
Confidence. The confidence is defined conf(X implies Y)=sup (X ∪Y)/ sup (X) . Be careful when reading the expression: here sup (X∪Y) means "support for occurrences of transactions where X and Y both appear", not "support for occurrences of transactions where either X or Y appears". Confidence ranges from 0 to 1. Confidence is an estimate of Pr(Y | X), the probability of observing Y given X. The support sup (X) of an item set X is defined as the proportion of transactions in the data set which contain the item set.
Lift. The lift of a rule is defined as lift(X implies Y) = sup (X ∪ Y)/((sup (Y) x sup (X)) or the ratio of the observed support to that expected if X and Y were independent. Lift can also be defined as lift(X implies Y) =conf(X implies Y)/ sup (Y). Lift measures how far from independence are X and Y. It ranges within 0 to positive infinity. Values close to 1 imply that X and Y are independent and the rule is not interesting.
Conviction. Conviction is sensitive to rule direction i.e. conv(X implies Y) is not same as conv(Y implies X). Conviction is somewhat inspired in the logical definition of implication and attempts to measure the degree of implication of a rule. Conviction is defined as conv(X implies Y) =(1 - supp(Y))/(1 – conf(X implies Y)).
Gain. When this option is selected, the gain is calculated using the gain theta parameter.
Laplace. When this option is selected, the Laplace is calculated using the laplace k parameter.
As result of the process, corresponded rules with confidence 0.6 has been created. We obtained seven rules. As we defined the confidence equal to or more than 0.6, all our rules satisfy our limits. We can see that results contain two or even three services (units).

The suitable presentation of results is graph (Fig. 6). The relations between services include confidence and support parameters.

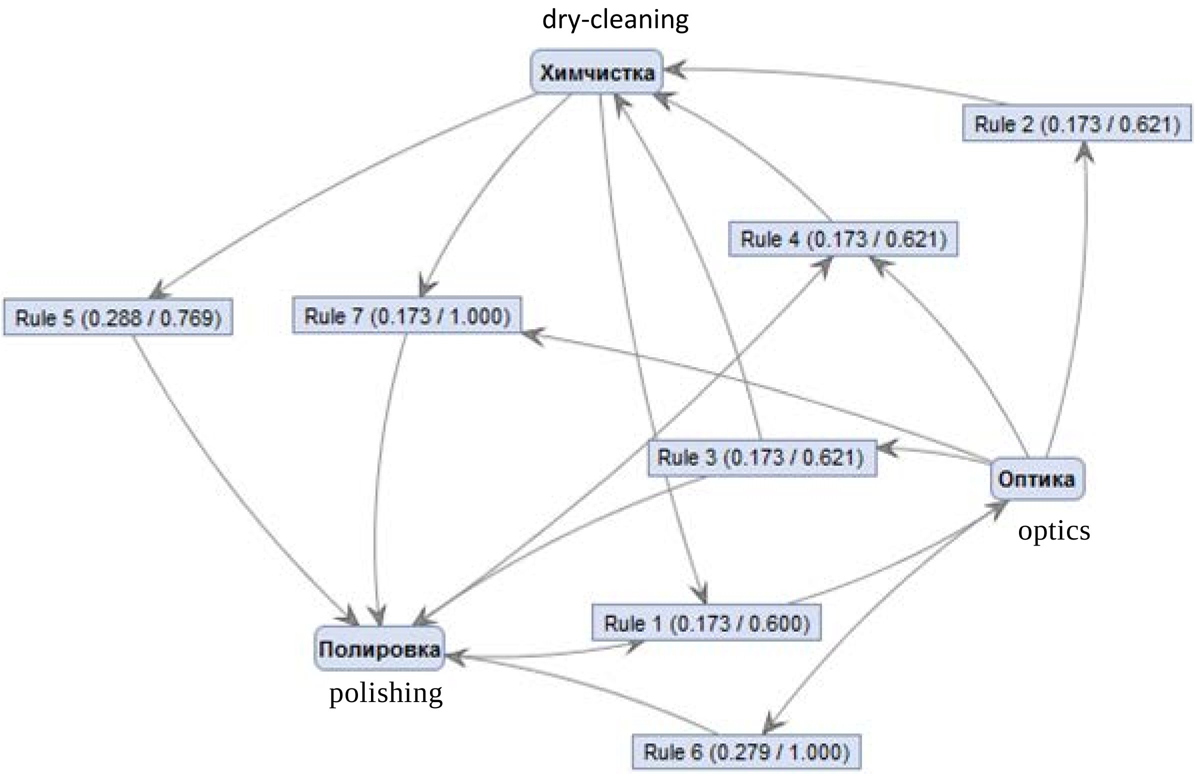

Figure 6. Results in the graph form

According to resulted association rules, we have strong relations between three services: “polishing”, “dry-cleaning” and “optics”. All rules cover these services. So, we should create the special offer for our clients that include these services. Prices for “polishing”, “dry-cleaning” and “optics” are presented at Fig. 7.

Service	Price
polishing	800 UAH
dry-cleaning	1200 UAH
optics	200 UAH
Sum for a package:	2200 UAH

Figure 7. Prices for services

Prime prices for “polishing”, “dry-cleaning”, “optics” are presented at Fig. 8. According these can be provided discount 15% for services package for clients without cost lost for the company.

Service	Price
Polishing	650 UAH
dry-cleaning	900 UAH
Optics	50 UAH
Sum for a package:	1600 UAH

Figure 8. Prime prices for services

6 Conclusions: evaluating of loyalty program effectiveness

The goal of any loyalty program - is to increase the efficiency of marketing and sales channels, increase customer loyalty, and as a consequence - increase profits. The goal of our program is the increase of quantity of repeat orders and, consequently, the increase of return. We can compare quantity of sets with selected services in one order during the quarter before and after implementing the loyalty program. So, we will analyze 4th quarter of 2012 year and 1st quarter of 2013 year. You can see result on diagrams (Fig. 8). For 4th quarter 2012 we have 15 sets of services with 2 repeat orders in it and for 1st quarter of 2013 - 24 (7). So, we can see increase in number of orders - 40% and in number of repeat orders – 350%.

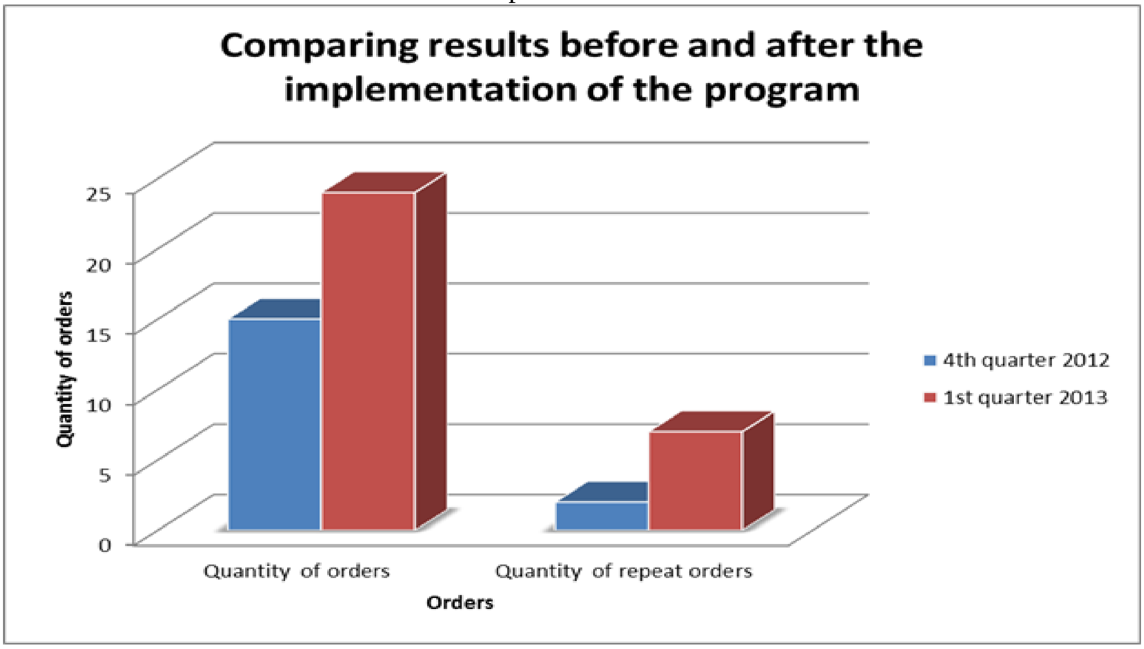

Figure 8. Diagram “Comparing results before and after the implementation of the program”

Also, we calculate the profit from this program (Fig. 9). In the 4th quarter 2012 company profits were 33.000 UAH, in 1st quarter 2013 they are 41140 UAH. The profit increase amounted 25%.

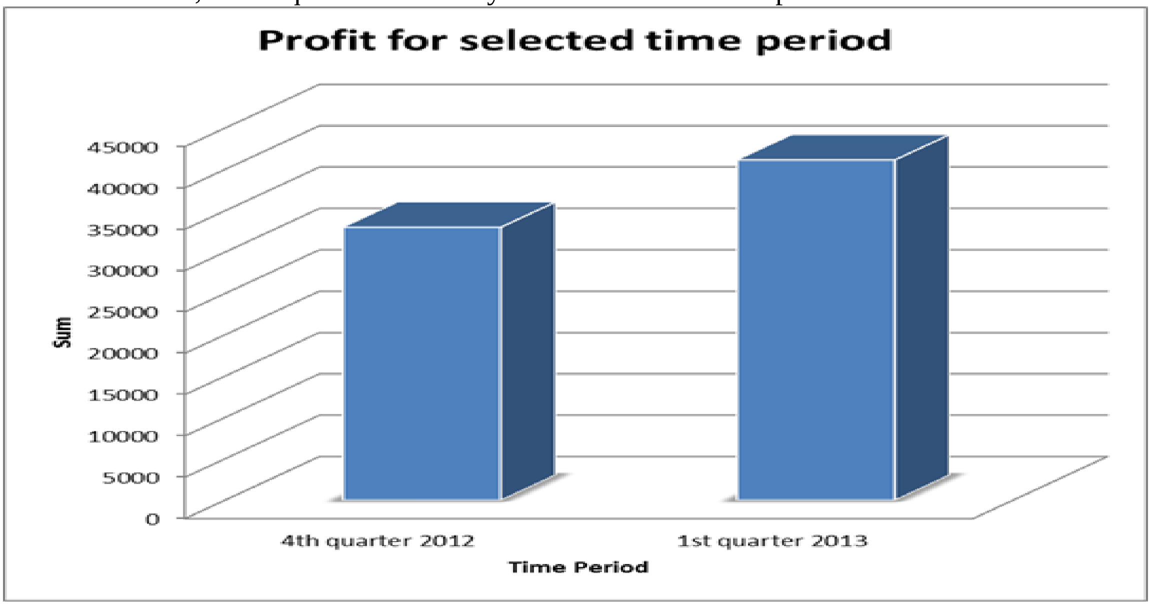

Figure 9. Diagram “Profit for selected time period”

Finally, our loyalty program resulted in not only the growth of client's orders included repeat orders but also the increase of company profit. And we can see these results only after 3 months. This is great result. We achieved our goals during 3 months of implementing of our program.

References

[1] A. Romeu. Cluster Detection in Laboratory Auction Data: A Model-Based Approach. *Panoeconomicus,* 58(4): 473–488, 2011.

[2] H. Shah-Hosseini. Binary Tree Time Adaptive Self-Organizing Map. *Neurocomputing*, 74 (11): 1823-1839, 2011.

[3] H. Chu Chai. Intelligent value –based customer segmentation method for campaign management: A case study of automobile retailer. *Expert Systems with Applications*, 34(4): 2754-2762, 2008.

[4] Y.S. Kim and W.N. Street. An intelligent system for customer targeting: A data mining approach. *Decision Support Systems*, 37, 215–228, 2004.

[5] R. Ng and J. Han. Efficient and Effective Clustering Method for Spatial Data Mining. *Proc. of 1994 Int'l Conf. on Very Large Data Bases (VLDB'94),* Santiago, 1994, 144-155.

[6] J.Sander and M. Ester M. Density-Based Clustering in Spatial Databases: The Algorithm GDBSCAN and its Applications. *Data Mining and Knowledge Discovery*, Vol. 2, No. 2, 169-194, 1998.

[7] P. Cheeseman and J. Stutz. Bayesian Classification (AutoClass): Theory and Results. *Advances in Knowledge Discovery and Data Mining,* AAAI/MIT Press, 1996, 153-180.

[8] T. Zhang and R. Ramakrishnan. BIRCH: An Efficient Data Clustering Method for Very Large Databases. *Proc. of ACM SIGMOD Int'l Conf. on Data Management,* Canada, 1996.

ZOLOTARYOVA Iryna
Kharkiv National University of Economics
Department of Informations Systems
Kharkiv
UKRAINE
izolotaryova@gmail.com

GARBUZ Iryna
Kharkiv National University of Economics
Department of Informations Systems
Kharkiv
UKRAINE
kafis@hneu.edu.ua

DOROKHOV Mykhailo
Kharkiv National University of Economics
Department of Informations Systems
Kharkiv
UKRAINE
michaeldorokhov@gmail.com