INFO- 2103
Primljeno / Received: 2013-04-13

UDK: 314:061:316.77:007
Izvorni znanstveni rad/Original Scientific Paper

SOCIO-DEMOGRAPHIC CHARACTERISTICS AS FACTOR OF CONSUMERS' BEHAVIOR

SOCIO-DEMOGRAFSKA OBILJEŽJA KAO FAKTOR PONAŠANJA POTROŠAČA

Sanja Bijakšić, Brano Markić, Slavo Kukić
Faculty of Economics, University of Mostar, Mostar, Bosnia and Herzegovina
Ekonomski fakultet, Sveučilište u Mostaru, Mostar, Bosna i Hercegovina

Abstract
Knowledge management is a general framework, concept that comprises the knowledge discovery process, learning techniques and collection of data mining algorithms. The paper researches the possibilities and advantages of supervised learning and knowledge base within the whole knowledge management in the process of market research. Knowledge management is a general concept of organizing, extracting, deployment und using the knowledge within an organization. The final goal is addition of the new knowledge in all functional areas of organization and use of knowledge to achieve strategic advantages and realize the defined and accepted goals of the organization. The paper is focused on marketing and discovering customers' behavior determined by their demographic characteristics. It is shown in the paper that the supervised learning and knowledge base as components of knowledge management may be successfully implemented in marketing in description and forecasting of customers' behavior.

Sažetak
Upravljanje znanjem je opći okvir, koncept koji obuhvaća proces otkrivanje znanja, tehnike učenja i skup algoritama data mininga. Rad istražuje mogućnosti i prednosti nadziranog učenja i baze znanja unutar cjeline upravljanja znanjem u procesu istraživanja tržišta. Upravljanje znanjem je koncept organiziranja, ekstrakcije, distribucije i uporabe znanja unutar organizacije. Krajnji cilj je dodavanje novih znanja u svim funkcijskim područjima i uporabe tog znanja radi postizanja strategijskih prednosti i realizacije definiranih i prihvaćenih ciljeva organizacije. Rad se fokusira na marketing i otkrivanje ponašanja kupaca određeno njihovim demografskim obilježjima.U radu je pokazano da nadzirano učenje i baza znanja kao komponente upravljanja znanjem se mogu uspješno implementirati u marketingu za opisivanje i predviđanje ponašanja kupaca.

1. INTRODUCTION

Researchers and scientist constantly investigate new methods how to acquire and manage knowledge resources effectively in order to reach the competitive advantages. The creation, distribution and implementation of the marketing knowledge powerfully contributes to products and services improvement, shortening market channels, consumer satisfaction, promotion, and pricing. Many analyses show that firms with superior marketing knowledge and capabilities have the business performances upon its competitors /**1**/. Marketing knowledge management is a system consisting of four processes extraction, knowledge representation, knowledge dissemination and knowledge implementation. Knowledge extraction is a complex process of discovering hidden relationships that exist among data and cannot be uncovered by implementing trivial (simple) sequence of steps, by simple algorithm. Therefore we stress out that the relationships among data are hidden and can be discovered by application of adequate algorithms.
Knowledge representation is the second step of knowledge flow process. Knowledge can be

present in different representation forms such as diagrams, plain text, semantic networks, first order predicates, production rules, frames, and object-attributes-values. In this paper the extracting knowledge will be in the form of production rules presented as clauses of Prolog programming language.

Knowledge dissemination is the third step of the marketing knowledge management where knowledge has to be provided to employees in marketing business function and shared with other sectors and functional areas in organization. Finally, the implementation is the verification of the total marketing knowledge management activities. The utilization of extracting knowledge to reach the competitive advantages represented in financial results is the purpose of the whole marketing knowledge management process in economic sense. Knowledge is the result of a learning process. The goal of the learning process is to generate new knowledge about the relationships and connections that exist between the data and variables in the system which is being analyzed. These knowledge and skills allow us to direct the system to set goals and predict their behavior in the future. Today a number of algorithms that allow supervised learning are developed: inductive rules, associative rules, neural networks, regression, Bayes classification algorithm, k-nearest neighbour classifiers, support vector machine (SVM), etc. /**2**/

2. SUPERVISED AND UNSUPERVISED LEARNING

Supervised learning assumes that cause and consequence relationships between the attributes (variables) exist. Appropriate combinations of values of input variables (these values identify the object) classify the object to a class (output). Data set is usually in the form of relational tables and it is divided into a training set and test set.

The training data (observations) are used to learn classifier. The logic of supervised learning is described by the following three steps:

1. Construct a model based on training dataset
2. Describe and distinguish classes for future prediction
3. Predict some unknown class labels.

Dataset table consists of header (attributes) and rows (records or tuples). One of the attributes is output attribute and it has as values different classes into rows (objects).

Unsupervised learning does not have defined class values in advance. The number of classes is unknown in advance and it can be 4, 10, 35 or any other number. Data in dataset are grouping and forming new categories, classes or clusters. In unsupervised learning, all the rows (objects) are influenced by latent variables. In many situations the target cannot be defined and it is necessary to consider learning without supervision. Cluster analysis and genetic algorithms are two examples of unsupervised learning. At the end of learning process there is new knowledge that must be represented on an adequate way and usable in the future. The knowledge is represented by production rules: IF (condition) THEN (action). After implementation of induction rules algorithms and creation the decision tree in the paper is built the knowledge base as the set of clauses of Visual Prolog programming language. This software tool enables the development of expert systems, and represents the knowledge by clauses. The Visual Prolog syntax ensures at relatively simple and easy way the knowledge representation about the objects properties as well as the relationships among objects and their properties.

2.1. Supervised learning in marketing research of customer preferences

The behavior of an individual and group of customers is more frequently the subject of a multidisciplinary theoretical interest. In fact, it is almost impossible to imagine that the research teams devoted to this type of research are consisted only of a single professional orientation – economists or sociologists, or social psychologists and psychologists. The knowledge about customer preferences and behavior can be extracted, represented, disseminated and used only as the result of teams that include a variety of professional orientations - and the most realistic of those who, in the true sense, are polyvalent.

In doing so, the focus of theoretical inquiry, is not only the behavior of consumers in the process of purchasing products and services. Research interests, on the contrary, extend to other forms of manifestation of consumer behavior - how to conduct the process of using products and services, and the impact of the overall process of consumer behavior - which, again, includes the purchase and use or deprivation of purchased products and services - to consumers and society.

If, however, we want to explain these claims in a simpler way, the study of consumer behavior focuses on the search for answers to the questions of how individuals or organizations choose to invest their available resources - time, money and effort - in the individual components of the consumer as a whole structure. The study of consumer behavior, after all, cannot avoid some other activities - collecting data on variables that influence consumers' decisions and overall behavior. In addition, n the background of this research is, finally, the need to facilitate the understanding of how consumers buy and use the purchased goods and services. All pointed out earlier, of course, this implies a detailed knowledge of the structure of the consumer, the factors that shape the structure, all of which, again, is a function of creating a successful marketing strategy. In fact, the creation and design of each marketing program requires insight into the preferences, expectations, and beliefs, as well as insights into other processes taking place in the minds of consumers. Does he know, for example, how to segment and access to different groups of consumers in terms of their response to new products, in order to prepare better strategy in terms of access, time and expectations of producers? Quality information about consumer behavior, then, can help in designing the best solutions and with the aim of protecting consumers. Finally, researching consumer behavior may help consumers to discover and understand some of the elements of their own behavior - and this, in turn, can increase the level of satisfaction of needs. The knowledge concerning the nature of his/her personality is needed. Because of the characteristics of the personality, in the end, it depends on what and when consumers buy,. In connection with the personalities are, then, some other variables that significantly shape the behavior of individual consumer. In connection with it is, in a certain way, motivation of each individual consumer. In connection with the personality is, then, the perception, the way we see the world around, because it is subjective - and that means two will never experience the same product, service or event in the same way. The learning process as a variable that also determines the behavior of individual consumers is connection with the personality, to some extent. True, the behavior of individual consumers is not only determined by the variables that are in relation to his/her personality. The structure of individual consumer behavior, for example, is to a large extent determined by the family where he/she belongs because it has a significant impact on the shaping of attitudes and skills. With awareness of the addressed theoretical premises we entered to the study of consumer preferences and market demand for bottled water. Therefore, the research is examining exactly these connections - those ties between some socio-demographic characteristics of consumer and consumption of bottled and tap water. The initial working hypothesis (H1) is that the demographic characteristics of consumers directly affect their tendency to consume bottled water. By supervised learning techniques and algorithms it is possible to extract the knowledge from data in survey and utilize this knowledge for future research and predictions.

The main hypothesis (H1) is tested by several socio-demographic characteristics of population - by gender, age, educational level and average monthly household income. Accordingly, the research includes four auxiliary hypotheses:

Hypothesis (H1a): Age affiliation is inversely proportional to the tendency of consumption of bottled water.

Hypothesis (H1b): Tendency to use bottled water is directly proportional to the education level.

Hypothesis (H1c): Increased average monthly household income implies higher level of bottled water consumption.
Hypothesis (H1d): The frequency of the consumption of bottled water is not directly related to gender affiliation.
The knowledge extracted from the data in survey will be represented in the form of production rules and will be used in the future. Moreover, the acquired knowledge about research technology and steps in the research are applicable to the study of consumer preferences and behavior for other products and services. This fact is of special value for the process of knowledge management in marketing.

3. APPLIED METHODOLOGY

The study, which aims to check the previous, additional and initial hypotheses, was on the territory of Bosnia and Herzegovina. The results however show the same or similar value as in other countries, particularly in ones which are characterized by overall economic and social development similar to the conditions present in the BIH.
The study, which tested the hypotheses, was conducted on a stratified sample of 150 respondents - where the stratification was done according to the four elements - gender, age, education, and the average monthly income of the household - and was done in the area of five countons in the Federation.

3.1. Induction rules

Induction rules are data mining methodology applicable in many real situations to classification problems. The learning goal is to create the classification model which will, based on the attributes input values, predict the class to which the object belongs. It is supervised learning method which builds decision trees from the samples set where the inputs and outputs are known in advance. For example, data set for transformation to decision trees using induction rules is given in the following table: /**3**/

Attribute 1	Attribute 2	Attribute 3	Class
A	70	True	Class 1
A	90	True	Class 2
A	85	False	Class 2
A	95	False	Class 2
A	70	False	Class 1
B	90	True	Class 1
B	78	False	Class 1
B	65	True	Class 1
B	75	False	Class 1
C	80	True	Class 2
C	70	True	Class 2
C	80	False	Class 1
C	80	False	Class 1
C	96	False	Class 1

Table 1. Data set for classification

The decision tree consists of nodes and connections between them. The nodes represent the attributes and connections (branches of the tree) the attributes values.

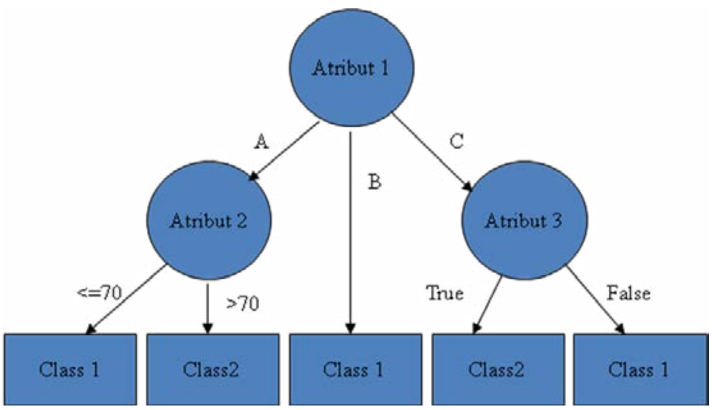


Figure 1. Decision tree for given dataset

The leaves represent the classes. This algorithm starts with all elements of the learning set. Following this we will choose some attribute that separates the collection. Each value of this attribute is a branch, and it forms a proper subset which contains the values of the attribute. Attribute values will split a table, using the node selection operations. The algorithm is applied recursively for each child node until all elements of the set does not belong to the same class. Each path to the leaf in the tree makes a rule of classification. The key decision is the selection of attributes for a node. Selection of attributes for ID3 and C4.5 algorithm is the minimum entropy. This choice is based on the mathematical theory of information.

The key concept is the information gain (gain). Gain (S, A) is the expected reduction in entropy because set S is sorted by attribute A. The decision tree (Figure 1) for the given data set can also be written in the form of the pseudo code:

```
If Attribute 1= A   Then
If Attribute 2<=70  Then
 Classification=Class1;
 Else
 Classification= Class2;
     Elseif Attribute 1= B Then
 Classification=Class 1;
          Elseif Attribute 1= C Then
               If Attribute 3=True Then
 Classification =Class 2;
                    Else
 Classification =Class 1.
```

Pseudo code can be written in the form of rules:

```
Rule 1: If Attribute 1= A   AND Attribute 2 <=70 Then Class 1;
Rule 2: If Attribute 1= A   AND Attribute 2   >70 Then Class 2;
Rule 3: If Attribute 1= B Then Class 1;
Rule 4: If Attribute 1=C AND A3=True Then Class 2;
Rule 5: If Attribute 1=C AND A3=False Then Class 1;
```

or the rules written as a knowledge base using the first order predicate in Prolog:[1]

```
class 1:-verify(A1=A,A2<=70);
class 1:- verify (A1=B,_);
class 1:- verify(A1=C,A3=False);
class 2:-provjeri(A1=A,A2>70);
class 2:-provjeri(A1=C,A3=True);
classify (A1,A2,A3,Class1):- A1=a,A2<70,A3="irrelevant".
classify (A1,A2,A3,"Class1"):- A1="b",A2=4,A3="irrelevant".
classify (A1,A2,A3,"Class1"):- A1="c",A2=4,A3="false".
classify (A1,A2,_A3,"Class2"):-A1="a",A2>70,A3="irrelevant".
classify (A1,A2,A3,"Class2"):- A1="c",A2=4,A3="true".
```

Using the production rules in the form of Prolog clauses replace the whole data set from Table 1 with the knowledge base. This knowledge base consists of ten clauses.

4. THE RESEARCH RESULTS

The methodology in this paper is based on induction rules and knowledge base built by using the R programming language. At the same time R denotes three things: data analysis software, calculator and programming language.[2]

4.1. Data set

Data set is the result of pre-processing activity followed by receiving completed surveys from respondents. Survey consists of the questions about the customers' socio-demographic characteristics at the market of drinking water: gender, age, education and household revenues.

[1] A production rule represents conditional knowledge. It is an implication in the form "clause head <= clause body". A fact is a Prolog clause without a clause body.

[2] R costs nothing and is completely free. To install R on your computer visit the site http://cran.r-project.org/mirrors.html and choose the nearest mirror

Respondent	Gender	Age	Education	Revenues	Purchase
1	male	31	bachelor	2500	does not buy
2	female	27	master	950	buys
3	male	48	high school	850	does not buy
...	...	...	...	...	..
149	female	29	bachelor	2800	buys
150	female	19	high school	950	does not buy

Table 2. Data set with socio-demographic customers characteristics

Data analysis in the R language is relatively simple and easy. So to display data in the form of data cube is sufficient following sequence of statements stored in script file named dCubeERP.R :

```
library(RODBC)
waDBuy<-read.table("C:\\Radovi2013\\waDBuy.csv",header=T,sep=";")
w<-odbcConnectAccess("C:\\Radovi2013\\wD")
 sareSQL<-sqlQuery(w, "SELECT * FROM waterAlldata")
erSQL=sqlQuery(w, "SELECT Gender,Age,Education, Revenues, Purchase FROM waterAllData")
library(scatterplot3d)
X<-erSQL$Education
Y<-erSQL$Revenues
Z=c(erSQL$Purchase)
scatterplot3d(X,Y,Z,pch=16,type="h",main="data cube",xlab="Education level", ylab="Revenues", zlab="Purchase")
```

Now we execute the script using the statement
>source(„C:/Opatija/dCubeERP.R")
The result is data cube where X axis represents education level[3], Y axis represents level of household revenues and Z axis represents purchase (1- means buys; 2- means does not buy).

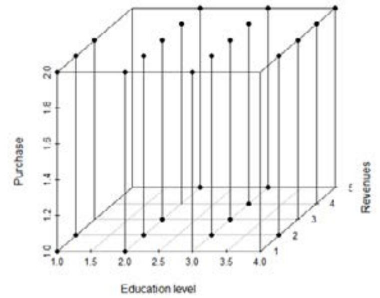


Figure 2. Data cube

The four auxiliary hypotheses will be tested by a few statements in R again stored into script file. The first hypothesis (H1b) is tested by a few statements of R language:

```
> h1AB<-sqlQuery(w, "SELECT Age,count(*) AS TB FROM waterAllData Where Purchase='buys' GROUP BY Age")
> h1AT<-sqlQuery(w, "SELECT Age,count(*) AS T FROM waterAllData GROUP BY Age")
> dataAP=rbind(c(1,2,3,4),c(h1AB$TB/h1AT$T))
> plot(dataAP[1,],dataAP[2,],type ="l",lwd=2,col="blue", xlab="Age", ylab="Frequency of purchase")
```

[3] All attributes (education level, household revenues and purchase) are represented numerically: education level: 1-elementary school, 2-high school, 3-bachelor, 4-master or Ph degree, level of household revenues: 1- revenues until 500BAM, 2- from 501 BAM to 1000 BAM, 3 - 1001 BAM to 1500 BAM, 4- 1501 BAM to 2000 BAM and 5 – greater than 2001 BAM, purchase: 1- does not by, 2- buys.

The result is in the form of diagram where it is obvious that the age is inversely proportional to the tendency of consumption of bottled water. Such visualization has confirmed the first auxiliary hypothesis.

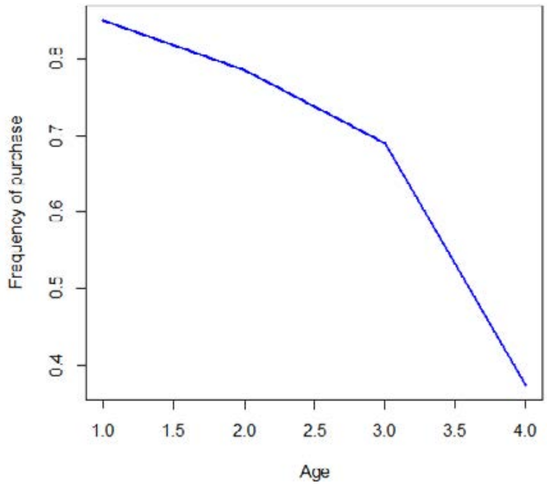


Figure 3. Age versus the frequency of purchase

If the age increases, the consumption of bottled water decreases.

Similar to the testing of the first hypothesis it is possible to test the second, the third, and the fourth hypothesis.

```
>h1EB<-sqlQuery(w, "SELECT Education,count(*) AS TB FROM waterAllData Where Purchase='buys' GROUP BY Education")
>h1ET<-sqlQuery(w, "SELECT Education,count(*) AS T FROM waterAllData GROUP BY Education")
>dataEP=rbind(c(1,2,3,4),c(h1EB$TB/h1ET$T))
>plot(dataEP[1,],dataEP[2,],type ="l",lwd=2,col="red", xlab="Education level", ylab="requency of purchase")
```

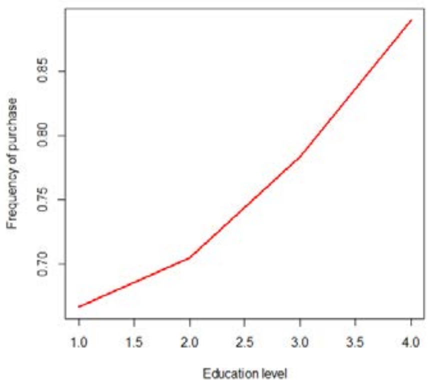
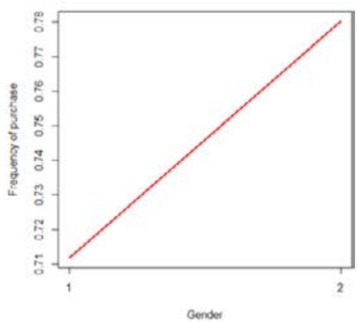


Figure 4. Educational level versus the frequency of purchase

The results are shown in the form of diagram where it is clear that the tendency to consume bottled water is directly proportional to the education level.

Hypothesis (H1c): Increased average monthly household income implies a higher tendency to consume bottled water. The test of this hypothesis enables the next set of R statements:

```
h1RB<-sqlQuery(w, "SELECT Revenues,count(*) AS TB FROM waterAllData Where Purchase='buys' GROUP BY Revenues")
h1RT<-sqlQuery(w, "SELECT Revenues,count(*) AS T FROM waterAllData GROUP BY Revenues ")
dataRP=rbind(c(1,2,3,4,5),c(h1RB$TB/h1RT$T))
plot(dataRP[1,],dataRP[2,],type ="l",lwd=2,col="blue", xlab="Revenues", ylab="Frequency of purchase")
```



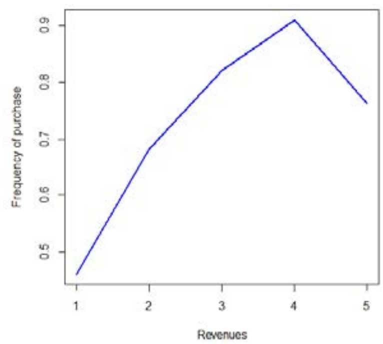


Figure 5a. Hypothesis (H1c): Increased average monthly household income im-

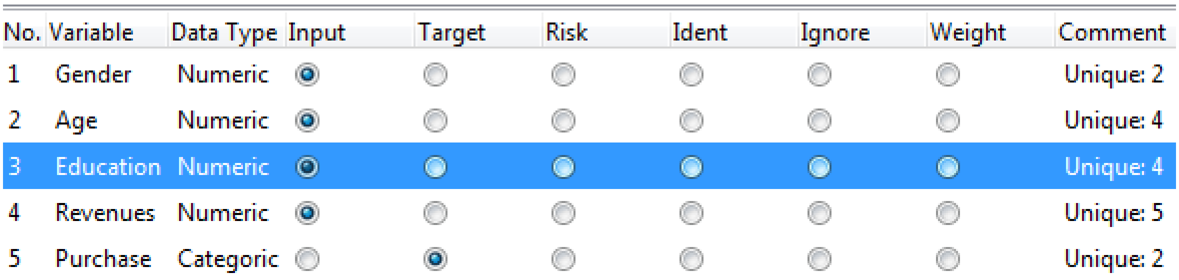
Figure 5b. Hypothesis (H1d): The frequency of the consumption of bottled water is not directly related to gender

The test of H1d hypothesis enables the next set R statements:

```
h1GB<-sqlQuery(w, "SELECT Gender,count(*) AS TB FROM waterAllData Where Purchase='buys' GROUP BY Gender")
> h1GT<-sqlQuery(w, "SELECT Gender,count(*) AS T FROM waterAllData GROUP BY Gender ")
> dataGP=rbind(c(1,2),c(h1GB$TB/h1GT$T))
> plot(dataGP[1,],dataGP[2,],type ="l",lwd=2,col="red", xlab=" Gender ", ylab="Frequency of purchase", xaxt = "n")
> axis(1, at=1:2, labels= dataGP[1,])
```

The final goal of this specific research is not to test four auxiliary hypotheses. The goal of the research is to create knowledge base using induction rules and directly test the initial working hypothesis (H1) that the demographic characteristics of consumers directly affect their appropriation and the tendency to consume bottled water. Again we use R language and the tool Rattle.[4] After entering data from the survey as data set, Rattle will automatically recognize the five variables, their data types, and input and target variables. Input variables are Gender, Age, Education, Revenues while target variable is Purchase.

[4] Rattle is an abbreviation from English words the **R** **A**nalytical **T**ool **T**o **L**earn **E**asily. Full instructions are available from http://rattle.togaware.com. After installing the required libraries be sure to restart the R console to ensure R can find the new libraries. Assuming R is installed we can then install the RGtk2 and rattle packages with: > install.packages("RGtk2") > install.packages("rattle"). Once installed we simply start Rattle by loadingthe rattle package and then evaluating the rattle() function:> library(rattle).



No.	Variable	Data Type	Input	Target	Risk	Ident	Ignore	Weight	Comment
1	Gender	Numeric	●	○	○	○	○	○	Unique: 2
2	Age	Numeric	●	○	○	○	○	○	Unique: 4
3	Education	Numeric	●	○	○	○	○	○	Unique: 4
4	Revenues	Numeric	●	○	○	○	○	○	Unique: 5
5	Purchase	Categoric	○	●	○	○	○	○	Unique: 2

Figure 6. Variables in induction rules algorithms

Among the set of models we choose one of the most common and popular data mining models, which is decision tree. The algorithm uses a recursive partitioning approach and is implemented in the 'rpart' package. The result of the application of this recursive algorithm to the data from the questionnaire presented in the form of data sets is given in the form of decision tree and production rules that form the knowledge base about the behavior of customers in the market of bottled drinking water. This behavior is explained by socio-demographic characteristics of gender, age, education and household income.

Summary of the Decision Tree model for Classification (built using 'rpart'):

```
n= 105
node), split, n, loss, yval, (yprob)
    * denotes terminal node
 1) root 105 25 buys (0.2380952 0.7619048)
   2) Revenues< 1250  41 15 buys (0.3658537 0.6341463)
     4) Age>=50  12  5 "does not buy" (0.5833333 0.4166667)
       8) Revenues < 750   4  1 "does not buy" (0.7500000 0.2500000) *
       9) Revenues >=750  8  4 "does not buy" (0.5000000 0.5000000)
        18) Gender = male  3  1 "does not buy" (0.6666667 0.3333333) *
        19) Gender = female 5  2 buys (0.4000000 0.6000000) *
     5) Age< 50  29  8 buys (0.2758621 0.7241379)
      10) Gender=male 1 8  4 "does not buy" (0.5000000 0.5000000)
        20) Education= elementary or high school 5  2 "does not buy" (0.6000000 0.4000000) *
        21) Education= bachelor or master  3  1 buys (0.3333333 0.6666667) *
      11) Gender=female 21  4 buys (0.1904762 0.8095238) *
   3) Revenues>=1250  64 10 buys (0.1562500 0.8437500) *
Classification tree:
rpart(formula = Purchase ~ ., data = crs$dataset[crs$train, c(crs$input,
    crs$target)], method = "class", parms = list(split = "information"),
    control = rpart.control(minsplit = 2, minbucket = 3, usesurrogate = 0,
        maxsurrogate = 0))
Variables actually used in tree construction:
[1] Age      Education Gender    Revenues

Root node error: 25/105 = 0.2381
n= 105
    CP nsplit rel error xerror    xstd
1 0.04      0      1.00   1.00 0.17457
2 0.02      2      0.92   1.28 0.18867
3 0.01      6      0.84   1.32 0.19028
```

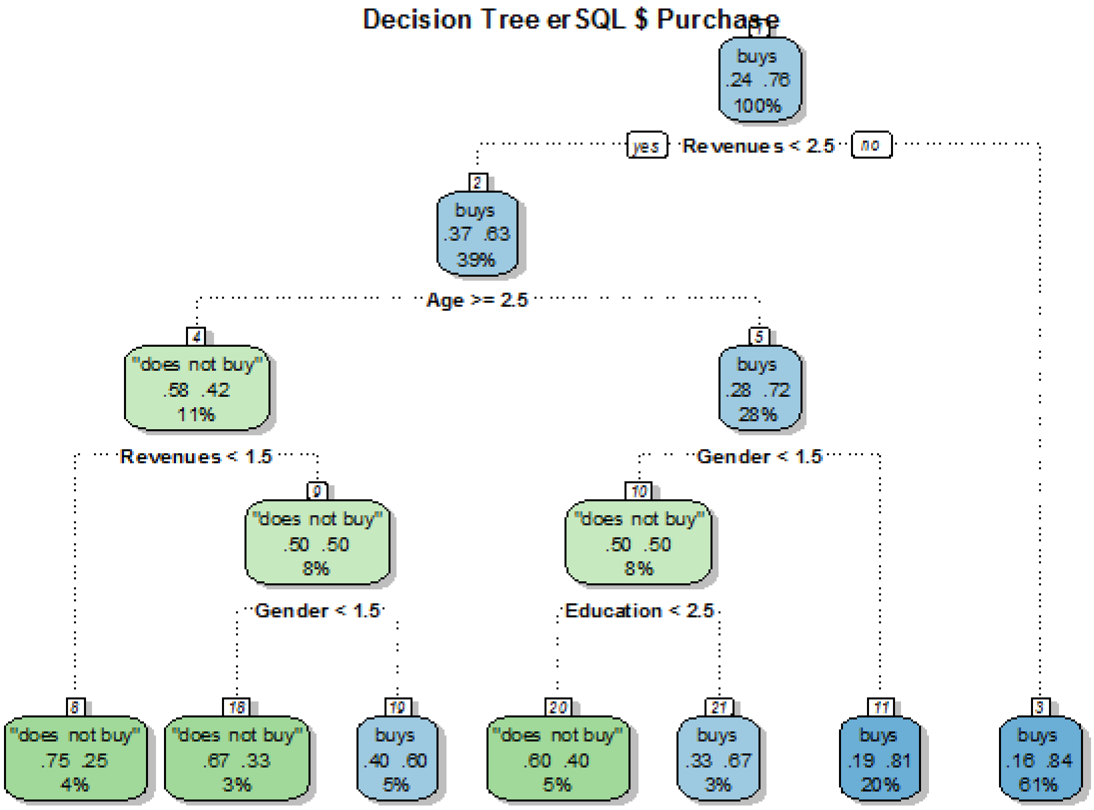


Figure 7. Decision tree

The decision tree reflects the knowledge hidden in data set getting from survey. The decision tree can be replaced by the set of decision rules.

Tree as rules:
Rule number: 3 [Purchase=buys cover=64 (61%) prob=0.84]
Revenues>=1250
Rule number: 11 [Purchase=buys cover=21 (20%) prob=0.81]
Revenues< 1250
Age< 30
Gender=female
Rule number: 21 [Purchase=buys cover=3 (3%) prob=0.67]
Revenues< 1250
Age< 30
Gender=male
Education=bachelor or master
Rule number: 19 [Purchase=buys cover=5 (5%) prob=0.60]
Revenues< 1250
Age>=30
Revenues>=750
Gender= female
Rule number: 20 [Purchase="does not buy" cover=5 (5%) prob=0.40]
Revenues< 1250
Age< 30
Gender= female
Education= elementary or high school
Rule number: 18 [Purchase="does not buy" cover=3 (3%) prob=0.33]
Revenues< 1250
Age>=30
Revenues>=750
Gender= female
Rule number: 8 [Purchase="does not buy" cover=4 (4%) prob=0.25]
Revenues< 1250
Age>=30
Revenues< 750

Now it is possible to transform this data into knowledge base in the form of production rules in programming language Prolog. This information or knowledge can be used to examine the purchase decision based on customer's gender, age, education and revenues. The advantage of this approach of creating knowledge base as the vital component of the whole knowledge management process in marketing is that it does not require detailed mathematical understanding of all operation in order to provide important information to marketing managers.

CONCLUSION

The paper clearly shows the complexity of marketing knowledge management and all necessary steps in this process: extraction, representation, dissemination and utilization of the knowledge. Knowledge extracting from data set is under influence of the research in marketing, data mining algorithms and software tools. The paper presents in details the knowledge extraction process in marketing where induction rules and R language have shown satisfactory development and implementation power. The knowledge is in the form of production rules and can also be used for future research of consumer preferences and behavior.

Notes

/1/ Z. Abbasi (2007) A review of models of knowledge management implementation in Organizations, Knowledge Management Conference.

/2/ Markic, B. (2011) Customer segmentation by integrating unsupervised and supervised learning, Proceeding from international conference Economic Theory and Practice: Meeting the New Challenges, Faculty of Economics University of Mostar, october 2011, Mostar.

/3/ Kantaradžić, M., (2003) *Data mining, Concepts, Models, Methods, and Algorithms*, Wiley-Interscience.

Literature

1. Husić-Mehmedović, M., Kukić, S., Čičić, M. (2012), Consumer Behaviour, School of Economics and Business University of Sarajevo, Sarajevo.
2. John Maindonald and John Braun, (2003) *Data Analysis and Graphics Using R - An Example-Based Approach,* Cambridge University Press.
3. Markic, B. (2011) Customer segmentation by integrating unsupervised and supervised learning, Proceeding from international conference Economic Theory and Practice: Meeting the New Challenges, Faculty of Economics University of Mostar, october 2011, Mostar.