

Data Mining in Customer Profitability Analysis

Drazena Gaspar^{1,*}, Ivica Coric², Mirela Mabic¹

¹Faculty of Economics, University of Mostar, Bosnia and Herzegovina

²HERA Company, K.P.Kresimira IV bb, Bosnia and Herzegovina

Copyright © 2015 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract The aim of this paper is a presentation of data mining model that could be used for the measurement of current and forecasting of the future customer profitability. The purpose of this model is to forecast activities of individual customers in the future, and to value that company could expect in doing business with them. Modern customer profitability analysis shows that product cost is just one part of the relation enterprise-customer. A general framework for defining customer profitability, besides pure financial items, has to include a lot of non-linear and non-financial elements. Data mining methods do not use conventional learning methods that suffer from imperfections such as inability to explicitly transfer the knowledge from experts to machines or nonexistence of experts' will for knowledge transfer. Data mining can identify and adopt patterns and rules that exist in historical data stored in databases and/or data warehouses. It can work equally well with nonlinear and nonfinancial elements of environment which have influence on profitability results. Neural networks approved their capability for approximate description of any continuous function. Together with robust methods of genetic algorithms used in the learning process of networks, they make a good choice in the process of selecting methods for forecasting customer profitability. The proposed model for the forecasting of the customer profitability uses two data mining methods: neural networks and genetic algorithm. The paper presents results of empirical research related to forecasting of customer determination to specific segment made in a company which produces and distributes products like dry fruits, nuts, seeds and cereals for the market of South-East Europe.

Keywords Customer Profitability, Data Mining, Genetic Algorithm, Neural Network

1. Introduction

Analysis of business operations shows that product cost is just one part of the relation enterprise-customer. This relation is burdened with other different costs generated by

business activities like inbound logistics, administration of customer orders, sale, marketing and post-sales services. Customers, consistent with their own characteristics, differ a lot in the allocation of these costs. The first theoretical analysis related to this issue has started with work of [1], [2] and [3] in 1990s. The main contribution of all these works is that each monetary unit of revenue does not participate in profit in the same way. So, the focus has been changed on a particular customer and measurement of the revenue's monetary unit "value" emerged from doing business with this customer. This value is represented by the difference between revenue and costs that arise from doing business with the customer. While revenue presentation is a relatively simple, presentation of costs is a complex problem. Calculation of customer profitability starts with a deduction of product costs and proceeds with the recognition and deduction of other types of costs which strain on this business transaction.

Modern customer profitability analysis (CPA) refers to trends in business profitability research, measured at the customer level. Van Triest [4] distinguishes costs related to physical product delivery, targeted customer costs and the costs that could not be allocated to the individual customer. Shapiro [1] divides costs to the costs before sale, production costs, distribution costs and costs for post-sale service.

Calculation of individual/group customer profitability gives a basis for a detailed analysis of customer profitability distribution inside a company. Some empirical researches [3], [5], [6], [7] have results that make this field very interesting from the point of analysis and management of customer distribution. Management of unprofitable, but also of profitable customers, is a natural continuation of these researches. Correct and exact calculation of customer costs and identification of profitability dispersion, make the basis for defining of strategy for management of unprofitable customers and establishing the corresponding infrastructure for its implementation. Repeated evaluation of customer relationship, customer education, renewed negotiation about sales conditions, customers' migration, and as a final step, termination of business relations with the customer, are only some of the activities of the highly intensive process of customer management [8].

Enforcement of strategic customer management based on retrospective analysis has become insufficient for survival on the global market. One different approach to profitability, view on prospective customer profitability, brings a completely new knowledge important for decision making process. Prospective customer analysis forecasts elements of business relationship with the customer during its future lifetime (as the customer of the company), finding basis for forecasting in the retrospective analysis. Data mining methods, as neural networks and genetic algorithms, are especially suitable for this type of analysis.

Neural networks do not use conventional learning methods that suffer from imperfections such as inability to explicitly transfer of knowledge from experts to machines or nonexistence of experts' will for knowledge transfer. On the contrary, the knowledge stored in historical cases makes the basis for neural network learning. It means that instead of request that expert articulates knowledge and then enables the machine to absorb such articulated knowledge, neural networks are capable to identify and absorb hidden knowledge and patterns of behavior that are stored in historical data of retrospective customer analysis. They could work equally well with nonlinear and nonfinancial elements of environment which have influence on profitability results. Neural networks approved their capability for approximate description of any continuous function. That characteristic makes them a good choice in the process of selecting methods for forecasting customer profitability. But, beside obvious advantages, the usage of neural network has some potential limitations. Per example, back-propagation topology of neural network adjusts the network's weights and biases by calculating the gradient of the error. Usually, back-propagation neural networks are applied with random initial weight setting because of symmetry breaking [9]. However, training the neural networks with random initial weights may cause two main drawbacks: trapping into local minima and converging slowly [9]. In view of these limitations of back-propagation neural networks, global search techniques (e.g. genetic algorithm) could be used to overcome these shortcomings [10], [11]. A number of works compare the evaluation between back-propagation neural network and genetic algorithm for training neural networks [12]. Genetic algorithms (GAs) developed to mimic some of the processes observed in natural evolution are a class of global search algorithms techniques. They have been shown in practice to be very effective at function optimization, and searching large or complex (multimodal, discontinuous, etc.) spaces to

find nearly global optimum efficiently [9].

In genetic algorithm, fitness function will evaluate the adaptation of each individual; it is a key point to decide whether the outcome is good or not. Selection operator will choose adaptive parents depending on their fitness values. By this step, the population tends towards better individuals [9].

This paper presents results of empirical research, based on neural networks and genetic algorithm, related to forecasting of customer determination to specific segments.

2. Objectives

Business subjects present the profitability of their business using aggregate values. Measures like market share, relative market share, market penetration, brand penetration, customer satisfaction, total revenue, total costs, marketing costs and so on, show the way how company works with their customers as a whole. Continuous management tendency for improving the business results lead to necessity of analyzing individual customers. From that perspective, it is not enough anymore that company is agile and efficient in its business. The company should have to be the part of the value chain and to have suppliers and customers that will not decrease its profit margin. Results of customer profitability analysis presented in this paper, show that every customer is not equally valuable, i.e. profitable, for company. That is the reason why it is not enough for a company to rely on detail analysis of existing customer data, but it should be focused on proactive and predictive customer management. Prerequisites for that are big volumes of quality data (data warehouse) and understanding of data mining methods.

Customer profitability is represented by the difference between revenue and costs that arise from doing business with the customer in a defined period of time. The most of empirical phenomena, including sale, are following Pareto's rule 20-80: 20% of products or customers realize 80% of sales. But, detailed analysis of customer profitability phenomenon denied Pareto's rule. Kaplan [13] in his explanations mentioned "whale curve" which made the relation between cumulative profit and customers (Figure 1.). If there are on x-coordinate customers ordered by the profit they make, curve gives an interesting result: 20% of the most profitable customers generate 150-300% of total profit, next 70% are around break even and last 10% generate the loss of 50-200% of total profit, leading cumulative profit at the level of realized profit (curve ended in the point of 100% profit).

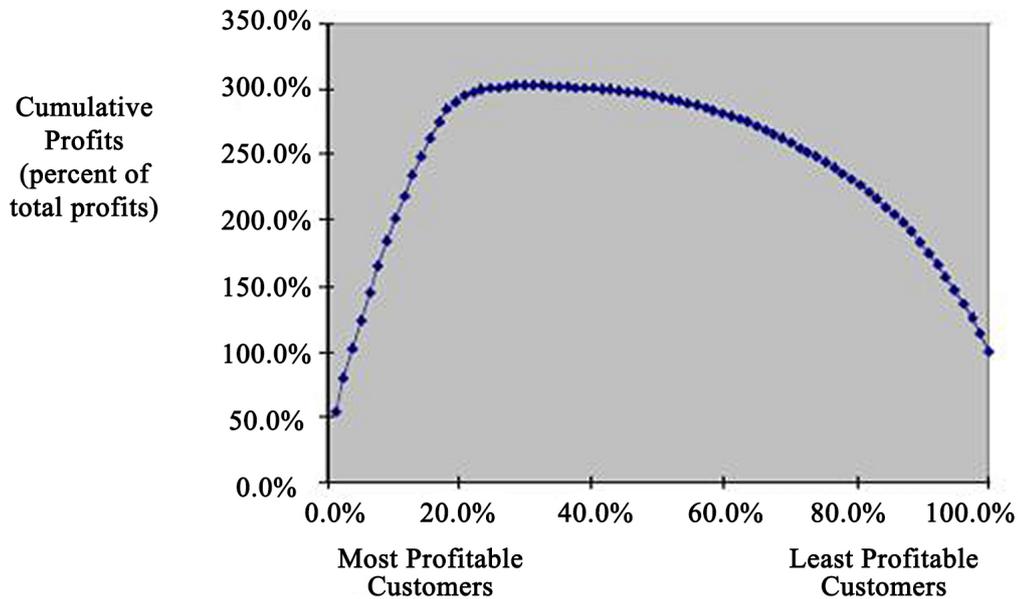


Figure 1. Cumulative customer profitability - „whale curve“ [13]

So, “whale curve” once again point out that not all customers are of equal value. According to the results of “whale curve” the customers should be divided into three different layers [14]:

1. Top segment customers: The most valuable and loyal customers. The company should do everything that can keep that customer and have to devote to them more attention and award them for their loyalty.
2. The second segment customers: Customers with average-small profit. Those customers have potential for growth and transition into top segment customers.
3. The third segment customers: Customers whose servicing cause loss for company.

Total result of company could be improved by actions customized to different customer segments. Kaplan [9], according to their nature, classifies those actions into three groups:

- Process improvement
- Price decisions
- Modify relationships with customers.

The third group opens the room for the possibility of improvement of customer loyalty and profitability measurement in order to represent and the forecast relationship between individual customers and company. The quantity and characteristics of the parameters of customer profitability point to the complexity of problem, measurement and forecast. Generally, the profitability is presented as financial item. But, a lot of nonfinancial factors like customer satisfaction, loyalty or reputation, have an important influence on the financial performance of the company, especially on long term. So, per example, customer satisfaction is preceding customer loyalty, while the result of customer loyalty is his profitability. The

important place in research of customer profitability belongs defining and measuring of customers’ costs.

3. Materials and Methods

The data source for research related to customer segmentation was a data warehouse of the company which produces and distributes products like dry fruits, nuts, seeds and cereals for the market of South-East Europe and for the period 2008 - 2013.

All customers are, according to data from Table 1, classified into corresponding segments (defined in the previous paragraph). The combination of data mining methods, neural networks and genetic algorithm, was used during research in two different ways.

The first way of use (Application 1) should confirm the capacity of classification of used data mining methods related to aligning of customers in the adequate segment.

The second way of use (Application 2) should present predictive capacity of combination of neural network and genetic algorithm. The starting point was the same data set, but limited by time. Namely, the idea was to use the same indicators from former example, but limited to first quarter.

The main hypothesis was: If network shows the ability to predict correct determination to customer segment that will be at the end of the year, based on value of input variables for the first quarter, it means that network has predictive capacity.

Proven hypothesis means that customer management could get mechanism for better defining a course of action related to correcting customers’ behavior and directing of customers into the preferred segment.

Table 1. Data used in customer segmentation

No.	Data	Description	In/Out variable
1	Year	Year to which data refers	N/A
2	Customer	Customer to which data refers	N/A
3	Customer size	Customer size (1-small, 2-medium, 3-big)	Input
4	Customer origin	Customer origin (1-domestic, 2-EU, 3-nonEU)	Input
5	Cost	Total costs of business with customer in observed year	N/A
6	Order cost	Total costs of customer's order processing and fulfillment costs in observed year	Input
7	Shipping cost	Total cost of shipping orders costs in observed year	Input
8	Purchase cost	Total costs of purchase and warehousing costs for delivered products to customer in observed year	Input
9		Total costs of raising purchase orders to suppliers for delivered products to customer in observed year	Input
10		Total costs of receiving shipments from suppliers for delivered products to customer in observed year	Input
11	Revenue	Total revenue realized by the customer in observed year	Input
12	No of deliveries	Total number of deliveries to customer in observed year	Input
13	No of products	Total number of different products in customer trade in observed year	Input
14	No of delivery places	Total number of different places where products were delivered to customers in observed year	Input
15	Net margin	Net margin realized in customer trade in observed year	N/A
16	No of returns	Total number of returns of good in customer trade in observed year	Input
17	Value of returns	Total value of returns of good in customer trade in observed year	Input
18	Value of discount	Total value of the discount given to customers in customer trade in observed year	Input
19	Customer payment	The average time of delay in customer payments in observed year (1-less than average, 2-more than average)	Input
20	Customer segment	Determination to segment based on realized customer business results in observed year	Output

In the both ways of usage, the input set of data (customer, customer size, customer origin, cost, order cost, shipping cost, purchase cost, costs of raising purchase orders, costs of receiving shipments, revenue, number of deliveries, number of products, number of delivery places, net margin, values of returns, number of returns, discount value and customer payment) was divided to:

- Training data (60%),
- Cross Validation data (15%) and
- Testing data (25%).

As a software tool was used NeuroSolutions for Excel version 6.20 of company NeuroDimension. Tool Express Builder was used for comparison of results of different types of neural network: MLP-Multi Layer Perceptron, LR-Linear Regression, PNN-Probabilistic Neural Network, RBF-Radial Basis Function, MLPPCA-Multi Layer Perceptron with Principle Component Analysis, SVM-Classification Support Vector Machine, GFF- Generalized Feedforward, TDNN- Time-Delay Network, TLRN-Time-Lag Recurrent Network and RN- Recurrent Network. The number of epochs was 100.

Test results of the first set of data, which were used for testing classification ability of the model, were declared architecture of MLP with PCA network as the best (Figure 2). A Multi-Layer Perceptron (MLP) is a kind of feed-forward ANN model (i.e. forward direction links), consisting of three layers; the input, hidden and output layers [15]. This kind of networks is called the autoassociative network [16]. The performance of a MLP depends on its generalization capability, meaning the data representation. This could suggest a need for eliminating correlation in the data before they are being presented to an MLP. This can be achieved by applying the Principal Component Analysis (PCA) technique onto input data sets prior to the MLP training process as well as interpretation stage [15].

Performances of winning network are given in tables 2 and 3. Moving of MSE through epochs during the learning process of the network is shown on Figure 4. The network had a minimal error in the case of the training set in 98 epochs, while in the case of validation set network reached the best result in 67 epochs.

Performance Metrics				Training			Cross Validation			Testing		
Model Name	MSE	r	Correct	MSE	r	Correct	MSE	r	Correct	MSE	r	Correct
MLP-1-O-M (Multilayer Perceptron)	0,023982	0,756951	66,88%	0,015328	0,842895	65,96%	0,032923	0,772785	66,84%			
LR-O-B-M (Linear Regression)	0,026808	0,722547	69,86%	0,018361	0,806211	66,81%	0,027239	0,804491	68,62%			
LR-O-B-L (Linear Regression)	0,024747	0,747486	71,03%	0,024323	0,742531	66,81%	0,026462	0,815606	72,19%			
MLP-1-B-L (Multilayer Perceptron)	0,018189	0,822504	82,22%	0,013192	0,867681	81,28%	0,022665	0,851812	83,67%			
PNN-O-N-N (Probabilistic Neural Network)	0,016416	0,842373	72,95%	0,01126	0,885829	66,81%	0,025646	0,82824	72,70%			
RBF-1-B-L (Radial Basis Function)	0,016713	0,857745	84,88%	0,022939	0,819791	89,79%	0,031156	0,8224	88,27%			
GFF-1-B-L (Generalized Feedforward)	0,017431	0,838749	84,03%	0,021048	0,804705	87,23%	0,018315	0,8779	88,01%			
MLPPCA-1-B-L (MLP with PCA)	0,003527	0,969349	92,55%	0,009272	0,918751	91,91%	0,017779	0,894897	92,35%			
SVM-O-N-N (Classification SVM)	0,004737	0,984878	100,00%	0,02211	0,832042	77,87%	0,02963	0,824129	82,40%			
TDNN-1-B-L (Time-Delay Network)	0,023072	0,785122	89,99%	0,026799	0,706861	85,96%	0,036721	0,760319	82,65%			
TLRN-1-B-L (Time-Lag Recurrent Network)	0,073394	0,041839	48,18%	0,078102	0,003046	46,81%	0,101076	-0,03536	50,00%			
RN-1-B-L (Recurrent Network)	0,091671	0,365175	75,00%	0,057666	0,679641	76,17%	0,090133	0,536377	70,66%			
MLP-2-B-L (Multilayer Perceptron)	0,020215	0,802158	85,36%	0,013541	0,864035	83,83%	0,025875	0,830001	87,24%			
MLP-1-B-M (Multilayer Perceptron)	0,027671	0,716024	69,44%	0,021486	0,770605	67,66%	0,030248	0,785597	68,62%			
MLP-2-O-M (Multilayer Perceptron)	0,024053	0,756683	62,82%	0,014735	0,849085	59,15%	0,02822	0,799927	65,05%			
MLP-2-B-M (Multilayer Perceptron)	0,033034	0,675251	68,38%	0,027417	0,725153	66,81%	0,038393	0,749492	68,62%			
MLPPCA-1-O-M (MLP with PCA)	0,019589	0,809379	70,94%	0,013748	0,859719	65,96%	0,018983	0,868381	67,35%			
MLPPCA-1-B-M (MLP with PCA)	0,023941	0,760223	72,33%	0,018972	0,799002	66,38%	0,026783	0,808034	71,68%			
GFF-1-O-M (Generalized Feedforward)	0,025154	0,747114	68,38%	0,018844	0,813563	62,98%	0,032421	0,779883	68,88%			
GFF-1-B-M (Generalized Feedforward)	0,028373	0,708752	70,94%	0,015741	0,836308	66,81%	0,028975	0,793989	70,92%			
RBF-1-O-M (Radial Basis Function)	0,029805	0,692611	60,36%	0,023587	0,751491	54,47%	0,032512	0,772728	62,76%			
RBF-1-B-M (Radial Basis Function)	0,054272	0,417369	57,69%	0,049832	0,538692	52,34%	0,074638	0,509454	58,16%			
TDNN-1-O-M (Time-Delay Network)	0,03546	0,612147	54,91%	0,021937	0,777892	59,57%	0,050446	0,58931	53,06%			
TDNN-1-B-M (Time-Delay Network)	0,067276	0,13341	61,22%	0,062053	0,213808	59,57%	0,093301	0,307771	57,91%			
RN-1-O-M (Recurrent Network)	0,025253	0,74357	71,69%	0,01703	0,825895	65,11%	0,030076	0,789067	71,43%			
RN-1-B-M (Recurrent Network)	0,029617	0,688176	65,81%	0,016494	0,828966	65,53%	0,030169	0,780644	67,09%			
TLRN-1-O-M (Time-Lag Recurrent Network)	0,035269	0,620863	62,61%	0,020754	0,781244	60,43%	0,038221	0,717076	66,07%			
TLRN-1-B-M (Time-Lag Recurrent Network)	0,026416	0,746007	63,78%	0,023229	0,756326	61,70%	0,038226	0,72081	60,97%			

Figure 2. Summary of all Networks for Application 1

Performance Metrics				Training			Cross Validation			Testing		
Model Name	MSE	r	Correct	MSE	r	Correct	MSE	r	Correct	MSE	r	Correct
MLP-1-O-M (Multilayer Perceptron)	0,0225	0,7952	53,72%	0,0273	0,7809	52,70%	0,041	0,6541	37,13%			
LR-O-B-M (Linear Regression)	0,0216	0,8007	70,20%	0,0271	0,7715	71,17%	0,0411	0,6593	72,09%			
LR-O-B-L (Linear Regression)	0,0209	0,8082	71,44%	0,0262	0,7804	73,42%	0,0386	0,6852	71,54%			
MLP-1-B-L (Multilayer Perceptron)	0,0168	0,8508	86,79%	0,0194	0,8428	87,84%	0,029	0,7696	77,24%			
PNN-O-N-N (Probabilistic Neural Network)	0,0128	0,888	72,23%	0,0253	0,7913	73,42%	0,0353	0,7129	72,09%			
RBF-1-B-L (Radial Basis Function)	0,0227	0,7897	74,15%	0,0335	0,7197	76,13%	0,045	0,623	71,54%			
GFF-1-B-L (Generalized Feedforward)	0,1395	0,2357	73,81%	0,1339	0,2845	69,82%	0,175	0,1418	62,60%			
MLPPCA-1-B-L (MLP with PCA)	0,0152	0,8655	87,58%	0,0148	0,8844	87,39%	0,0281	0,7778	75,34%			
SVM-O-N-N (Classification SVM)	0,0048	0,9842	98,76%	0,0343	0,7591	79,28%	0,0415	0,7018	73,71%			
TDNN-1-B-L (Time-Delay Network)	0,0091	0,9247	94,70%	0,0262	0,7898	75,23%	0,0363	0,7201	65,58%			
TLRN-1-B-L (Time-Lag Recurrent Network)	0,0774	-0,0121	48,98%	0,0803	-0,0691	48,65%	0,0886	-0,0519	41,46%			
RN-1-B-L (Recurrent Network)	0,0704	0,0867	46,15%	0,0651	0,2914	49,10%	0,073	0,1992	32,79%			
MLP-2-B-L (Multilayer Perceptron)	0,0166	0,8526	91,52%	0,0225	0,8192	86,49%	0,0282	0,7754	80,49%			
MLP-1-B-M (Multilayer Perceptron)	0,0204	0,8133	70,36%	0,0254	0,789	70,72%	0,0374	0,6868	76,96%			
MLP-2-O-M (Multilayer Perceptron)	0,0212	0,8054	69,46%	0,0282	0,7618	69,82%	0,04	0,6594	71,00%			
MLP-2-B-M (Multilayer Perceptron)	0,0238	0,7878	70,14%	0,0284	0,7662	71,62%	0,0392	0,6644	75,07%			
MLPPCA-1-O-M (MLP with PCA)	0,0204	0,816	66,63%	0,0257	0,7843	59,91%	0,0386	0,6719	56,91%			
MLPPCA-1-B-M (MLP with PCA)	0,0226	0,792	70,48%	0,0289	0,7552	71,62%	0,0448	0,6215	71,54%			
GFF-1-O-M (Generalized Feedforward)	0,0193	0,8252	69,68%	0,0261	0,782	68,92%	0,0376	0,6888	66,67%			
GFF-1-B-M (Generalized Feedforward)	0,0221	0,7957	69,68%	0,0284	0,76	71,62%	0,0359	0,7008	72,36%			
RBF-1-O-M (Radial Basis Function)	0,029	0,7472	56,90%	0,0344	0,7192	56,31%	0,0409	0,6599	47,70%			
RBF-1-B-M (Radial Basis Function)	0,0586	0,6994	56,67%	0,0649	0,648	57,21%	0,0685	0,5942	63,41%			
TDNN-1-O-M (Time-Delay Network)	0,028	0,7588	51,02%	0,0317	0,7427	52,70%	0,0488	0,5838	37,13%			
TDNN-1-B-M (Time-Delay Network)	0,0604	0,0059	48,87%	0,0669	0,0052	45,50%	0,0701	0,0631	61,52%			
RN-1-O-M (Recurrent Network)	0,0233	0,7864	67,76%	0,0295	0,7549	68,02%	0,0408	0,6599	65,85%			
RN-1-B-M (Recurrent Network)	0,0237	0,7797	68,89%	0,0314	0,7372	68,47%	0,0368	0,6906	71,00%			
TLRN-1-O-M (Time-Lag Recurrent Network)	0,0213	0,8046	66,97%	0,0285	0,7615	66,67%	0,0423	0,6394	64,23%			
TLRN-1-B-M (Time-Lag Recurrent Network)	0,0245	0,8012	66,29%	0,0361	0,6892	65,32%	0,0426	0,6303	69,65%			

Figure 3. Summary of all Networks for Application 2

4. Results

Test results of the second data set, which were used for testing predictive ability of the model, were declared MLP – 2- B-L network (Figure 3) as the optimal one.

Table 2. Metrics of Best-Performing networks for Application 1

	Training	Cross Val.	Testing
# of Rows	939	235	392
MSE	0,0035272	0,0092725	0,0177791
Correlation (r)	0,9693491	0,9187507	0,8948969
# Correct	869	216	362
# Incorrect	70	19	30
% Correct	92,55%	91,91%	92,35%

Table 3. Final and Minimum MSE for Application 1

Best Networks	Training	Cross Validation
Epoch #	98	67
Minimum MSE	0,0622935	0,0741663
Final MSE	0,0633549	0,0796188

The results of MLP network are presented in tables 4 and 5. Moving of MSE through epochs during the learning process of this network is shown on Figure 5. In the case of the training, set network had a minimal error in 100 epochs, while in the case of validation set network reached the best result in 78 epochs.

Table 4. Metrics of Best-Performing networks for Application 2

	Training	Cross Val.	Testing
# of Rows	886	222	369
MSE	0,0165998	0,0225251	0,0282442
Correlation (r)	0,8525781	0,8191502	0,7754392
# Correct	809	192	297
# Incorrect	75	30	72
% Correct	91,31%	86,49%	80,49%

Table 5. Final and Minimum MSE for Application 2

Best Networks	Training	Cross Validation
Epoch #	100	78
Minimum MSE	0,0738607	0,1179428
Final MSE	0,0738607	0,1245319

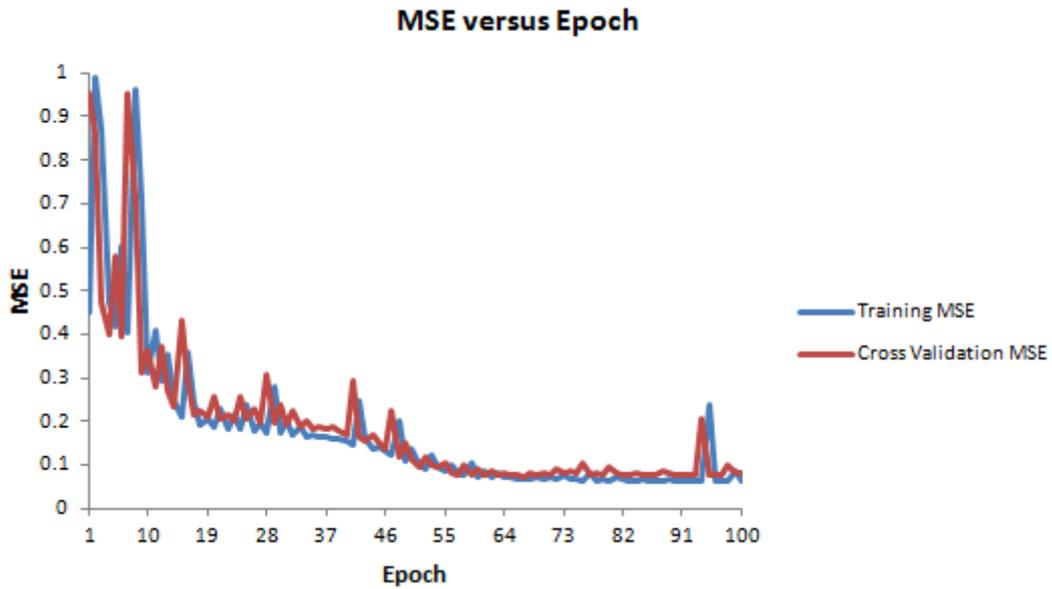


Figure 4. MSE versus Epoch for Application 1

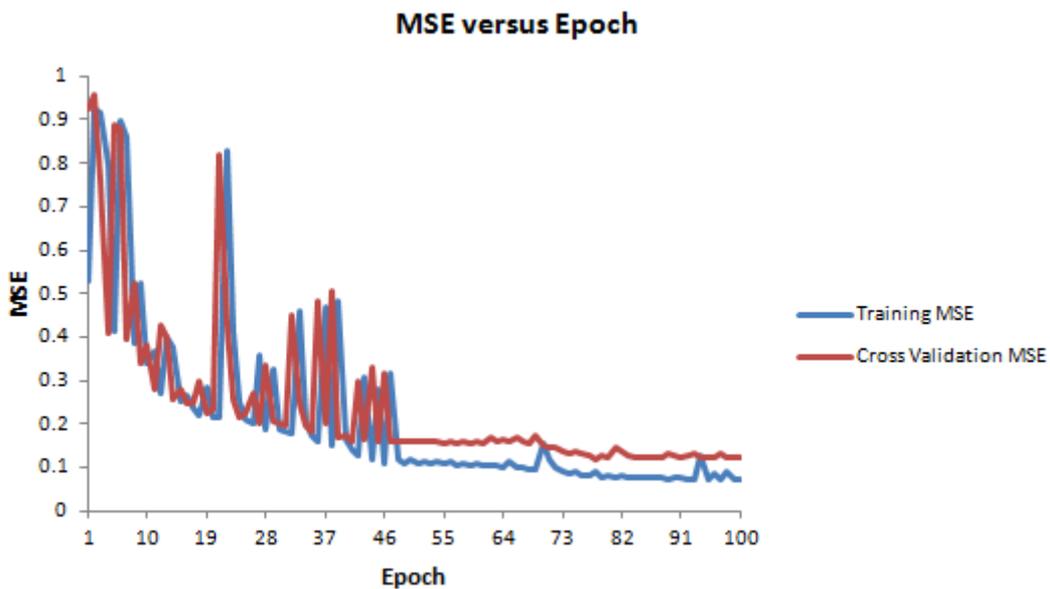


Figure 5. MSE versus Epoch for Application 2

After initial network training for application 2 and achieved results which were presented previously, genetic algorithm was used for network optimization. Training parameters were set on as follows: the number of epochs on 100, maximum number of generations on 100 and the value of a population of 50. During the process of optimization, genetic algorithm optimized selection of input variables, step size, value of momentum and the number of neurons in the hidden layer of the network. The goal of such optimization is finding parameters which will result in a minimal error of the network. After a cycle of optimization by use of genetic algorithm was finished, the error of the network was reduced as it is presented in Table 6.

Table 6. Final and Minimum Minimum MSE for Application 2 after generic optimization.

Optimization Summary	Best Fitness	Average Fitness
Generation #	6	7
Minimum MSE	0.0748949	0.095297847
Final MSE	0.0748949	0.095297847

5. Conclusions

The paper presents that modern customer profitability analysis refers to trends in business profitability research, measured at the customer level. Calculation of

individual/group customer profitability gives a basis for a detailed analysis of customer profitability distribution inside a company. Enforcement of strategic customer management based on retrospective analysis has become insufficient for survival on the global market. Modern approach to profitability, based on prospective customer profitability, brings a completely new knowledge important for decision making process. Prospective customer analysis forecasts elements of business relationship with the customer during its future lifetime, finding basis for forecasting in the retrospective analysis.

The paper presents the potential of neural networks as a tool in the process of analyzing and forecasting customer profitability. The advantages of neural networks lay in the characteristics of their structure and ability of learning and generalization. By using historical data, neural networks are capable to learn and make input-output data mapping. Also, they can adapt weights if new influences arrive from the environment, e.g. it means that they are capable to adjust themselves according to the changes in the environment. Neural network can be projected, not only to give answer to the setting of data parameters, but also to define the degree of confidence in their output result. One of the potential limitations of neural networks – setting initial weights – could be overcome by using genetic algorithm.

The results of presented research showed that classification based on neural networks and genetic algorithm - used for defining network weight factors - could be the efficiency used in customer segmentation. This combination of neural networks and genetic algorithm approved its applicability with classification precision of 91.91 - 92.55% and with prediction precision of 86.49 – 91.31% in the conducted research.

The further research should include new variables into model, taking into account of the number of parameters that defined customer profitability. Since neural networks have the ability to equally well process financial and nonfinancial factors like customer satisfaction or loyalty, it opens new possibilities for more complex applications of these methods, all with a main goal to ensure data necessary for decision making process related to customer management.

REFERENCES

- [1] Shapiro P, Rangan K, Moriarty R, Ross E. Manage Customers for Profits (Not Just Sales). Harvard Business Review September-October 1987.
- [2] Howell R A, Soucy S R. Customer profitability; as critical as product profitability. *Management Accounting*, October 1990: 43-47.
- [3] Foster G, Gupta M, Sjoblom L. Customer profitability analysis: Challenges and new directions. *Journal of Cost Management* (Spring 1996): 5-17.
- [4] Triest S. Customer size and customer profitability in non-contractual relationships. *Journal of Business & Industrial Marketing* 2005, 20/3: 148-155.
- [5] Cooper R, Kaplan R. Profit Priorities from Activity-Based Costing. *Harvard Business Review* 1991.
- [6] Storbacka E K. Segmentation based on customer profitability: retrospective analysis of retail bank customer bases. *Journal of Marketing Management* 1997, Vol. 13 No. 5:479-92.
- [7] Krakhmal V. Customer profitability analysis in service industries. BAA Annual Conference 11-13 Apr 2006, Portsmouth, UK.
- [8] Mittal V, Sarkees M, Murshed F. The right way to manage unprofitable customers. *Harvard Business Review* April 2008
- [9] Chang Y, Lin J, Shieh J, Abbod M.F. Optimization the Initial Weights of Artificial Neural Networks via Genetic Algorithm Applied to Hip Bone Fracture Prediction. *Advances in Fuzzy Systems* 2012
- [10] Al-Shareef A.J, Abbod M.F. Neural networks initial weights optimization. *Proceedings of the 12th International Conference on Modelling and Simulation (UKSim '10)*, March 2010, pp. 57–61.
- [11] Venkatesan D, Kannan K, Saravanan R. A genetic algorithm-based artificial neural network model for the optimization of machining processes. *Neural Computing and Applications* 2009, vol. 18, no. 2, pp. 135–140.
- [12] Sexton R.S., Gupta J.N.D. Comparative evaluation of genetic algorithm and backpropagation for training neural networks. *Information Sciences* 2000, vol. 129, no. 1–4, pp. 45–59.
- [13] Kaplan R, Narayanan V G. Measuring and managing customer profitability. *Journal of Cost Management* 2001, 15(5): 5–15.
- [14] Farris P W, Bendle N T, Pfeifer P E, Reibstein D J. *Marketing metrics*, Wharton School Publishing, USA, 2006
- [15] Mohamad-Saleh J, Hoyle B.S. Improved Neural Network Performance Using Principal Component Analysis on Matlab. *International Journal of the Computer, the Internet and Management* Vol.16. N.o.2 (May-August, 2008) pp 1-8.
- [16] Qiu J, Wang H, Lu J, Zhang B, Du K.L. Neural Network Implementations for PCA and Its Extensions. *ISRN Artificial Intelligence* 2012