Artificial Neural Networks in CRM

Rexhep Rada[1] and Bashkim Ruseti[2]

[1]Department of Mathematics and Informatics, Faculty of Natural Sciences,
"Aleksandër Xhuvani" University, Elbasan, Albania
rexhep.rada@gmail.com
[2]Department of Mathematics, Statistics and Applied Informatics,
Faculty of Economics, University of Tirana, Tirana, Albania
bruseti@gmail.com

Abstract. In recent years, Data Mining becomes a very important technique supporting in business decisions. From the modern techniques, artificial neural networks have improved a lot the companies in their success. Artificial neural networks allow creating and using complex functions in a natural mode, to make useful prediction. In this paper we will evaluate the use of ANN in a Customer Relationship Management (CRM) database. Our goal is to define the target group, who spends more in mobile communication in Albania.

Keywords: Artificial Neural Networks, Customer Relationship Management (CRM), Data Mining

1 Introduction

The process of determining the target group interested more in a certain product is on the top of the agenda of every business company. Technical implementation of these methods in Western countries is widespread, but in Albania doesn't exist. Companies use as prediction tools, the descriptive statistics, without entering in more sophisticated methods. [2]

In this article we will try to give an overview of the application of ANN in mobile communication. It consists in analyzing an old database of customers, connecting the characteristics of the clients with the amount of money that they spend in phone calls. The data are taken from a survey in 930 users. [3]

2 Methodology

The data are dividing in two parts: Data Training Set and Data Validation Set. This case study is a based in supervised learning, more specifically in back-propagation algorithm.

We try to regulate the weights in ANN for improving prediction from training set. [4]

S. Markovski, M. Gusev (Editors): ICT Innovations 2012, Web Proceedings, ISSN 1857-7288

The training data consist in a large database with more than 620 records, with respective input and output. In this case the variables are: age, marital status, social status, personal income, family income, number of persons in family, and the outcome is the monthly expenses for mobile communication. (Table 1)

Table 1. Database sample input

Name	Surname	Age	marital status	Social status	Personal income	Family income	Nr. of persons in family	Nr. of brothers / sisters	Monthly expenses
Ina	Sina	19	single	student	0	50,000	4	2	1,500
Miri	Hoxha	28	single	employed	100,000	200,000	3	3	5,000
Refit	Sulmina	20	single	employed	10,000	50,000	4	3	4,000
Blerta	Teta	35	married	employed	10,000	40,000	5	7	2,500
Valbona	Filja	39	married	employed	17,000	100,000	8	7	2,000

2.1 Variables

We have to make every variable numerical. But beyond this, the numbers have to been normalized. We do this because we don't want that big numbers of not weight variables affect more in the final prediction. [1]

We can use from dummy variable, which equals 0 and 1 in two value variable domain to proportional value in the]0;1[. Let see a version of numerical table (table 2).

Table 2. Database sample input with numerical variables.

Name	Surname	Age	marital status	Social status	Personal income	Family income	Nr. of persons in family	Monthly expenses	Nr. of brothers/ sisters
Ina	Sina	19	1	0.5	0	50,000	4	1,500	2
Miri	Hoxha	28	1	1	100,000	200,000	3	5,000	3
Refit	Sulmina	20	1	1	10,000	50,000	4	4,000	3

Blerta	Teta	35	0	1	10,000	40,000	5	2,500	7
Valbona	Filja	39	0	1	17,000	100,000	8	2,000	7

After the numeric step, we can perform other processing methods like min-max standardization, to normalize the values with the aim that the highest number doesn't affect more in prediction. [5]

2.2 Results

For each exemplar of 310 records of validation dataset (remaining from the database with 930 records) we make the prediction with several ANN models. After that we compare the outcome of the prediction with the outcome in dataset. We see the percentage of the right predictions. Table 3 displays the results.

Table 3. Full results

Architecture (with the number of the hidden layers)	**Dataset training**			**Dataset validation**		
	MSE	**1-MSE**	**Good classification rate**	**MSE**	**1-MSE**	**Good classification rate**
Net_00	9.6	90.4	89.9	20.6	79.4	74.5
Net_01	3.7	96.3	96.4	4.54	95.46	82.5
Net_02	1	99	95.35	2.64	97.36	86.8
Net_03	0.5	99.5	97	0.8	99.2	90.1

We notice from tables, that the use of hidden layers improves substantially the model.

In training sample, the best version (Net 03) has the lowest mean square error (0.5%) and allows us the classification rate of 97%. In validation set we notice that the best network is the same, the classification rate jumped to 90.1% with MSE=0.8%.

3 Conclusion and future work

On finding the right target group of a product depends the profit of every company. This article gave an overview of Artificial Neural Networks and its applications. This technique have to be used even in Albania to improve the nowadays analyzes. We described a case study in mobile communication. Is used a database from a survey of

930 records. Comparing the prediction after training with the real results in database we got an amazing result. In validation set we notice that the best network gave a classification rate of 90.1% with MSE=0.8%.

The future direction of this study will be on improvement of the activation function. These helpful results will encourage the Albanian company for a fast introduction of Data mining techniques in their company.

References

1. Classification with Artificial Neural Networks And Support Vector Machines: Paolo Valigi, Vidas Gulbinas, Rainer Westphal, Khaled Mohamed Almhdi And Rainer Reuter.
2. CRM through DM: A Case Study Angelo M. Cister, Nelson F. F. Ebecken.
3. Bankruptcy Prediction for Credit Risk Using Neural Networks: Amir F. Atiya.
4. Neural Networks By Christos Stergiou And Dimitrios Siganos.
5. Artificial Neural Networks (Ann)Goodchild, Longley.