Predicting the Success of Bank Telemarketing Using Artificial Neural Network

Mokrane Selma

***Abstract*—**The shift towards decision making (DM) based on artificial intelligence (AI) techniques will change the way in which consumer markets and our societies function. Through AI, predictive analytics is being used by businesses to identify these patterns and major trends with the objective to improve the DM and influence future business outcomes. This paper proposes an Artificial Neural Network (ANN) approach to predict the success of telemarketing calls for selling bank long-term deposits. To validate the proposed model, we uses the bank marketing data of 41188 phone calls. The ANN attains 98.93% of accuracy which outperforms other conventional classifiers and confirms that it is credible and valuable approach for telemarketing campaign managers.

***Keywords*—**Bank telemarketing, prediction, decision making, artificial intelligence, artificial neural network.

I. Introduction

TELEMARKETING is a personalized and interactive method for consumers to learn about products, services or opportunities. When the consumer has limited time or is immobile, telemarketing delivers products and services, and makes buying opportunities possible with local and distant vendors [1].

Many consumers would be greatly inconvenienced without telephone access to such products and services as; travel reservation, catalog purchases, subscription renewals, charity donations, banking services, insurance purchases, food, medicine purchases and personal services.

With a broader market available to vendors, more choices available to consumers and sales costs frequently reduced through telemarketing, consumers may benefit from lower, long-term product and services prices.

Through AI, many companies and businesses have benefited from the advantages of prediction analysis. This technique makes it possible to anticipate an event, predict incidents, and even behaviors based on data set, after having established a correlation between data variables, it tries to predict what will happen in the future. Anticipating actions in companies and being able to control them helps to prevent problems and conflicts from occurring in the future, and increasing productivity is an objective for many businesses. Today, the terms AI, big data or predictive analysis are very often used. They represent a new sector that is still unclear to the public.

Businesses today are oriented toward the predictive analytics to solve internal problems and to improve their competitiveness. Indeed, AI, data mining or machine learning can solve complex problems within businesses. Based on data, these tools are able to anticipate customer feedback and know their purchases in order to guide more accurately the marketing campaign. These actions help build customer loyalty and optimize marketing actions. Ultimately, predictive analytics is a tool that is nowadays perfectly integrated into the life of businesses, to become progressively indispensable.

M. Selma is with the Department of Economics, Abdelhamid Ibn Badis University of Mostaganem - 27000, Algeria (phone: +213(0)778757655; e-mail: selma.mokran@gmail.com,).

An intelligent system for prediction of the success of bank telemarketing using a data-driven approach is suggested by [2]. Research on the success of bank telemarketing has deserved increasing interest in recent years. Using data mining for bank, direct marketing was developed in [3]. In [4] authors used deep convolutional neural network for predicting the success of bank telemarketing.

Neural networks are widely used tools for classification [5], estimation [6], [7], and prediction [8], [9]. Artificial neural technique has been applied in many areas. The back propagation is one of the most learning algorithms used for the training process of ANN. ANNs are becoming a powerful tool for prediction. This method is used to explain and predict one or more observable and effectively measured phenomena. The analysis is done on all variables of the database using a supervised learning classification (the class is known) based on automatic learning. This technique aims to extrapolate new information from hidden information (this is the case of scoring), in this case there is a target variable to be predicted.

In this study, the data used for Bank Telemarketing prediction is a dataset publicly available for research. The dataset used for prediction was collected from a Portuguese retail bank. The approach used in this study is the ANNs of the type Multi-Layer Perceptron (MLP).

This paper is organized as follows. The description of the dataset and ANN with its architecture and learning algorithm are introduced in Section II. In Section III, results and discussions are given in order to demonstrate the effectiveness of the proposed technique. Section IV gives the conclusions.

II. Material and Methods

A. Dataset

The data used in this study are real data downloaded from UCI Machine Learning Repository [10]. It was gathered from a Portuguese retail bank [11] over five years, from May 2008 to June 2013. The dataset was collected by marketing campaigns of a Portuguese banking institution, based on phone calls. In order to access if the product (bank term deposit) would be subscribed "yes" or "no", all clients were contacted. Dataset contains 41188 phone contacts composed of 21 attributes with a label attribute as shown in Table I.

TABLE I
ATTRIBUTE INFORMATION

Num.	Attribute Name	Description	Type
1	Age	It is age of client.	Numeric
2	Job	It is type of client's job.	Categorical
3	Marital	It is client's marital status.	Categorical
4	Education	What is the highest education of client?	Categorical
5	Default	Does client has credit?	Categorical
6	Housing	Does client has housing loan?	Categorical
7	Loan	Does client has personal loan?	Categorical
8	Contact	What is a contact communication type of client?	Categorical
9	Month	What is the last month of the year contracting to the client?	Categorical
10	Day of Week	What is the last day of the week contracting to the client?	Categorical
11	Duration	How long does it contact to the client?	Numeric
12	Campaign	Number of contacts performed during this campaign and for this client	Numeric
13	Pdays	Number of days that passed by after the client was last contacted from a previous campaign	Numeric
14	Previous	Number of contacts performed before this campaign and for this client	Numeric
15	Poutcome	Outcome of the previous marketing campaign	Categorical
16	Emp.var.rate	Employment variation rate	Numeric
17	Cos.price.idx	Consumer price index	Numeric
18	Cons.conf.idx	Consumer confidence index	Numeric
19	Euribor3m	Euribor 3 month rate	Numeric
20	Nr.employed	Number of employees	Numeric
21	Label	Has the client subscribed a term deposit?	Categorical

B. ANNs

Neural networks [12] are tools widely used for classification, estimation, prediction and segmentation. They are derived from biological models and are composed of elementary units: neurons. They are organized according to an architecture and they are well adapted to the problems including continuous variables possibly too noisy. They get good performance.

1. MLP

MLP is one of the most used neural networks for approximation problems, classification and prediction. It is usually composed of three layers of neurons totally connected.

A MLP is a set of interconnected neurons. This constitutes an extension of the perceptron model, with one or more hidden layers between the input and the output. Each neuron in a layer is connected to all the neurons of the previous layer and the next layer. The activation functions used in such networks are mainly threshold or sigmoid functions. It can solve non-linearly separable problems and more complicated logical problems. In our case we will study the three-layer networks.

- The input layer: it is the first layer of the network where no calculation is made. The input cell only receives in input the values of attributes: $x_1, x_2, \dots, x_p$
- The output layer: it is the last layer of the network that contains one or more decision cells.
- The hidden layer: it is located between the input layer and that of the output; a MLP may contain one or more hidden layers. The only peculiarity is that starting from the input layer toward that of output and by passing through the hidden layers, all neurons of a layer are the entries of each neuron of the next layer (only the layer that comes just after).

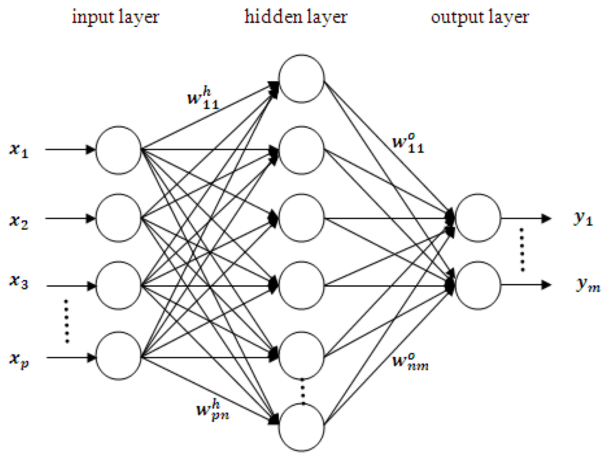

Fig. 1 A MLP with one hidden layer

Recall that the output of the elementary cell (linear perceptron) considered for the development of a MLP is a differentiable function of real variables, to thereby a perfect example of differentiable function of real variables is the sigmoid function defined by:

$$\sigma(y) = \frac{e^y}{e^y+1} = \frac{1}{1+e^{-y}} \tag{1}$$

It can be noted that the derivative of the function σ is simple to calculate:

$$\acute{\sigma}(y) = \frac{e^y}{(1+e^y)^z} = \sigma(y)(1-\sigma(y) \tag{2}$$

It is essential that this calculation is simple because the derivative of this function will be used in the update rule of weight by the back-propagation algorithm [11].

We can now define a MLP considered in the following of this paper with the elementary cells of perceptron type topreal entries $\vec{x} = (x_1, \dots, x_p)$ is defined by the real synaptic weights $\vec{w} = (w_1, \dots, w_p)$, and the output y is calculated by:

$$y(\vec{x}) = \frac{1}{1+e^{-y}} \; where \; y = \vec{x}\vec{w} = \sum_{i=1}^{p} w_i x_i \tag{3}$$

2. Learning by Back Propagation Algorithm

The principle of the algorithm is to minimize an error function E. It is then to calculate the contribution of each synaptic weight to this error. Indeed, each of the weight influences the corresponding neuron, but the change to this neuron will affect all neurons of the following layers. The error therefore measures the gap between the desired and calculated outputs on the full sample. The problem is to determine a vector $\overrightarrow{w}$ which minimizes $E(\overrightarrow{w})$. The error is for example:

$$E = E_{(\vec{x},\vec{y})}(\overrightarrow{w}) = \frac{1}{2}\sum_{j=1}^{m}(y_j^d - y_j)^2 \qquad (4)$$

The back-propagation algorithm is then written:

Input: a sample Ω^{ANN} *of* R^p*R^m

A MLP with an input layer C_0, q-1 *hidden layers* $C_1, \ldots, C_{q\text{-}1}$, *and an output layer* C_q,

Random initialization of weights in $[0,1]$ *for each neuron* j *of the network.*

Set the value of ε.

Repeat

Take an example $(\vec{x}, \overrightarrow{y^d})$ *in* Ω^{ANN}

Calculate the output $\vec{y}$ *of MLP for input* $\vec{x}$

---Calculation of δ_j *back-propagation---*

For *all output cells* ***j***, *calculate*

$$\delta_j = y_j(1 - y_j)(y_j^d - y_j)$$

End For

For *each layer from* q-1 *to* 1

For *each cell* j *of the current layer*

$\forall i \in Succ(j)$(Succ(j) *is the set of cells which take as input the output of the cell* j) *calculate*

$$\delta_j = y_j(1 - y_j)\sum_i(\delta_j * w_{ij})$$

End For

End For

--- Update weights ---

For *all weights*

$$w_{ji} \longleftarrow w_{ji} + \varepsilon\delta_j x_{ji}$$

End For

End Repeat

Output: *a MLP defined by the initial chosen structure and weights* w_{ji}

III. Results and Discussions

In this section; to show the contribution of the bank telemarketing prediction by MLP, a model is developed; able to predict the success of the bank telemarketing using real data.

In this section, to show the contribution of the prediction by ANN approach, ANN technology has been approved on bank telemarketing.

Fig. 2 shows accuracy rate of ANN method that is quite high; however, the proposed technique can provide a high rate than other methods. This method has proven to be effective in prediction the success of bank telemarketing with a high rate of precision.

IV. Conclusion

In this study, a non-traditional approach based on the ANN technology has been developed in order to predict the bank telemarketing. 98.93% accuracy rate has been achieved with the proposed method. The obtained results have shown that the effectiveness of the presented approach is able to predict seismic hazard more accurately than methods routinely used in Polish coal mines. Experimental results show that the MLP has good predictive effect and high accuracy.

```
Time taken to build model: 198.02 seconds

=== Evaluation on training set ===
=== Summary ===

Correctly Classified Instances        4075               98.9318 %
Incorrectly Classified Instances        44                1.0682 %
Kappa statistic                          0.9439
Mean absolute error                      0.0131
Root mean squared error                  0.1003
Relative absolute error                  6.7191 %
Root relative squared error             32.1345 %
Total Number of Instances             4119

=== Detailed Accuracy By Class ===

TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class
  0.925     0.003      0.977     0.925     0.95        ?        yes
  0.997     0.075      0.991     0.997     0.994       ?        no

=== Confusion Matrix ===

   a    b   <-- classified as
 417   34 |    a = yes
  10 3658 |    b = no
```

Fig. 2 MLP Prediction results

Major advantage of using ANN over other statistical and numerical techniques lies on their capability to capture non-linear relationship among concerned variables and optimization can be done very fast, without mathematical form of the relation between input and the output data is necessary. Their ability to learn by example makes them very flexible and powerful.

In our future work; it is recommended to improve the effectiveness of the proposed model by several means. It would be important to optimize the ANN model by combining it with other techniques such as genetic algorithms and metaheuristics.

References

[1] P. Kotler, and K. L. Keller. "Marketing Management".*14th Edition, New Jersey: Prentice Hall*, 2012.

[2] S. Moro, P.Cortez, and P. Rita. "A data-driven approach to predict the success of bank telemarketing". *Decision Support Systems*, 2014, 62, 22–31. http://dx.doi.org/10.1016/j.dss.2014.03.001.

[3] S. Moro, R. M. S. Laureano and P. Cortez. "Using Data Mining for Bank Direct Marketing: an Application of The CRISP-DM Methodology," *in Proceedings of the European Simulation and Modelling Conference - ESM'* 2011, Guimarães, 2011.

[4] K. H. Kim, et al., "Predicting the success of bank telemarketing using deep convolutional neural network," *in Proc. 7th International Conference of Soft Computing and Pattern Recognition (SoCPaR)*, Fukuoka, 2015.

[5] F. J. L. Iturriaga, and I. P. Sanz. "Bankruptcy visualization and prediction using neural networks: A study of US commercial banks". *Expert Systems with Applications*, 2015, 42 (6), p. 2857-2869.

[6] Kwon, H.-B., & Lee, J. "Two-stage production modeling of large U.S. banks: A DEA-neural network approach". *Expert Systems with Applications*, 2015, 42(19), 6758–6766. http://dx.doi.org/10.1016/j.eswa.2015.04.062.

[7] L. G. Brock III, and L. B. Davis. "Estimating available supermarket commodities for food bank collection in the absence of information". *Expert Systems with Applications*, 2015, 42(7), 3450-3461.

[8] C. Jeong, J. H. Min, and M. S. Kim. "A tuning method for the architecture of neural network models incorporating GAM and GA as

applied to bankruptcy prediction". *Expert Systems with Applications*, 2012, 39 (3), pp. 3650-3658.

[9] P. Du Jardin. "Predicting bankruptcy using neural networks and other classification methods: The influence of variable selection techniques on model accuracy". *Neurocomputing*, 2010, 73 (10), pp .2047-2060.

[10] A. Asuncion, and D. Newman, CA: University of California, School of Information and Computer Science, *UCI machine learning repository. Irvine*, 2012.

[11] S. Moro, P. Cortez and P. Rita. "A Data-Driven Approach to Predict the Success of Bank Telemarketing". *Decision Support Systems, Elsevier*, 2014, 62:22-31.

[12] S. Bajpai, K. Jain, and J. Neeti, "Artificial Neural Networks," the International Journal of Soft Computing and Engineering, vol. 1, no. 2011, pp. 2231-2307, 2011.