NEURAL NETWORK APPLICATION FOR PREDICTION

ELDA XHUMARI

Department of Informatics, Faculty of Natural Sciences, University of Tirana
Email: elda.xhumari@fshn.edu.al

Abstract. Neural networks are one of the areas of research in the field of artificial intelligence, based on attempts to reproduce the human neural system. The ability of the neural system to learn and correct mistakes, which should allow simulating, albeit roughly enough, the work of the human brain [1]. The most important feature of the network, which attests to its wide capabilities and huge potential, is the parallel processing of information, which speeds up the information processing. In addition, with a large number of interneuronal connections, the network acquires resistance to errors that occur on certain lines. Nowadays, neural networks are used to solve a big number of problems, one of which is the prediction problem.

Keywords: Neural networks, prediction, training, NN architecture, NN types, activation function

I. INTRODUCTION

Recent studies have shown the classification and prediction power of the Artificial Neural Networks. It has been demonstrated that a neural network can approximate any continuous functions. Neural networks have been successfully used for forecasting financial data series because they have the advantage that can approximate any nonlinear functions without any apriori information about the properties of the data series.

Another advantage of the usage of neural networks for prediction is that they are able to learn from examples only and that after their learning is finished, they are able to catch hidden and strongly non-linear dependencies, even when there is a significant noise in the training set.

The disadvantage is that NNs can learn the dependency valid in a certain period only. The error of prediction cannot be generally estimated.

The purpose of the study is to develop methods for a forecasting model to estimate the change of prices in real transactions market of Real Estate based on neural network technologies, which can significantly improve the efficiency of the Real Estate Agencies.

At the initial stage, will be analyzed the known methods of predicting data, using artificial neural networks.

The second task consists on developing my own model and a neural network Real Estate value algorithm that shows high performance in solving large-scale problems.

The third task is the development of the program and the study of the Real Estate value algorithm. At the decision of the specified problems will help modern methods of the theory of artificial neural networks, gradient optimization methods, possibilities of Matlab package [2].

II. METHODOLOGY OF THE EXPERIMENTS

In this section we introduce how to solve these problems, several experiments were carried out. The purpose of the experiments was to predict the value of Real Estate in the city of Tirana. Data on Real Estate objects (apartments) received during the period from February 01 2017, to May 15, 2017 were used. Each of the experiments can be divided into several stages.

The first step was the formation of a training sample. At this stage is determined how the data will be represented (input and output neurons). We have time series, i.e., a variable x changing in time x_t (t=1,2,...) and we would like to predict the value of x in time $t+h$.

The prediction of time series using neural network consists of teaching the net the history of the variable in a selected limited time and applying the taught information to the future. Data from past are provided to the inputs of neural network and we expect data from future from the outputs of the network.

As we can see, the teaching with teacher is involved. For more exact prediction, additional information can be added for teaching and prediction, for example in the form of interventional variables (intervention indicators) - see the figure 2. However, more information does not always mean better prediction; sometimes it can make the process of teaching and predicting worse. It is always necessary to select really relevant information, if it is available.

Various types of neural networks can be used for prediction, such as backpropagation [3], RB functions [4], feedforward network [5] and others. As a network architecture for our experiments is used multilayer perceptron backpropagation network, which helps on solving a lot of problems in different fields.

Training set can be made easily directly from the time series. Certain number of measured values is used as inputs and the value to be predicted (i.e., the value in the future, in some chosen distance after these input measured values) is used as required output. Input

part of the time series is called window, the output part is the predicted value. By shifting the window over time series the items of training set are made. It is advised to left part of time series for testing, i.e., to not use this part during learning, but to use it to test how successfully the network learned to predict our data.

The training set obtained in this way can be then adjusted for the needs of a particular neural network. For example, it may be necessary to adjust the values to a certain interval, such as (0,1).

Available data are often divided into three set: learning set, validating set and testing set. These sets can overlap and do not have to be continuous. The learning set is a sequence that is shown to the neural network during the learning phase. The network is adapted to it to achieve required outputs (in other words, weights in the network are changed based on this set). The difference to required output is measured using the validating set and this difference is used to validate whether the learning of the network can be finished. The last set, testing set, is then used to test whether the network is able to work also on the data that were not used in the previous process.

To summarize, the learning set is used for creating a model, validation set is used for verifying the model, and the testing set is used for testing of the usability of the model.

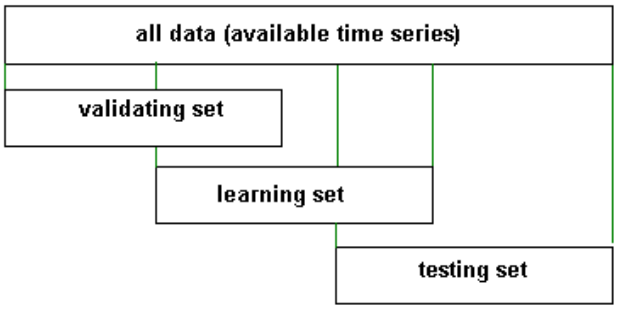

Figure 1. Validating, learning and testing set of data

Data preprocessing is important as well. Especially for neural networks that can have outputs only in a certain interval it is important to realize that it is not possible to predict values outside of this interval. Data normalization is then required for the network to be able to get meaningful outputs. The original data is converted to the form in which they can be fed as inputs for the neural network. For example, there are quantitative data (floor, number of rooms, living space) and non-quantitative data (region, type). Non-quantitative data are indexed, and then all data are normalized: $x = \{x_1, x_2, ... x_n\}$ set of each factor, that we have to normalize, for $n = \overline{1, i}$, i - number of elements of the set, x (at our sample i =400), then:

$$x_{norm} = \frac{x_i}{x_{\max}},$$

where $x_{\max}$ - maximal value x, x_{norm} - normalized value. As a result all the data values are from the interval [0,1].

III. EXPERIMENT RESULTS

Experiments are made using different types of architecture of neural network. On the first experiment the data used for training was used for testing too and on the second experiment, testing was performed on data not used for training. The results are shown in the table 1.

Table 1. Experiments results

Architecture neurons	MSE of Experiment 1				MSE of Experiment 2			
	Number of Epochs				Number of Epochs			
Hidden layers neurons	5	100	1000	10000	5	100	1000	10000
5	0.1253	0.1185	0.1356	0.1384	0.1893	0.1360	0.1506	0.1665
10	0.1697	0.1368	0.1380	0.1383	0.2356	0.1400	0.1472	0.1445
20	0.2106	0.1373	0.1392	0.1430	0.3692	0.1445	0.1415	0.1481

CONCLUSION

Neural networks are suitable for predicting time series mainly because of learning only from examples, without any need to add additional information that can bring more confusion than prediction effect. Neural networks are able to generalize and are resistant to noise. On the other hand, it is generally not possible to determine exactly what a neural network learned and it is also hard to estimate possible prediction error.

However, neural networks were often successfully used for predicting time series. They are ideal especially when we do not have any other description of the observed series.

REFERENCES

1 David H. Hubel, Torsten N. Wiesel (2005). Brain and visual perception: the story of a 25-year collaboration. Oxford University Press US. p. 106.

2 www.mathworks.com - Introduction to Dynamic Neural Networks - MATLAB & Simulink. *(Retrieved 2017-06-15)*

3 E.R. Kandel, J.H. Schwartz, and T.M. Jessell. Principles of neural science Appleton & Lange, 2000.

4 Chen, S., Billings, S. A., Cowan, C. F. N. and Grant, P.M., "Nonlinear System Identification using Radial Basic Functions,"*Intl. J. Systems Sci.*,**21**, 2513 (1990).

5 Fine, T. L., "Feed Forward Neural Network Methodology," Springer, N.Y. (1999).

★★★