Using Ensembles of Neural Networks for Forecasting Telemetry Data

Yauheni Marushko
United Institute of Informatics Problems of National Academy of Sciences of Belarus, Belarus, Minsk, st. Surganova 6, marushkoee@gmail.com

Abstract: *In this paper we propose an approach to solving the problems of forecasting multivariate time series telemetry data that describe the state of small airborne objects. The main objective of the proposed method - it's automated design and development of neural network models for solving such problems, namely the choice of model parameters are close to optimal. The approach is based on the use of ensembles of neural networks. In this case learning algorithm uses some elements of evolutionary strategy. The article also describes the experiments and experimental data.*

Keywords**:** Forecasting, telemetry data, ensembles of neural network, evolutionary strategy.

1. INTRODUCTION

Telemetry covers the measurement of physical quantities characterizing the state of objects or processes, the transfer of the results of these measurements, recording and data processing.

The processing and analysis telemetry data in a continuous process is accompanied by noisy data. This usually creates by non-deterministic sources of noise. This makes it the preferred use of the technology of artificial neural networks.

The following benefits are the basis of the development of systems to analyze telemetry data using a neural network [1]:

- The ability for self-learning, the replacement of complex mathematical apparatus necessary volume of information, the initial function of its transformation and the structure of the neural network;
- The possibility of realizing nonlinear mappings with substantial non-linearities in the hidden neurons of multilayer neural networks;
- High degree of parallelism possible for a neural network software system in real time.

These advantages provide application of neural network algorithms in the diagnosis of complex dynamic objects. Neural networks appear here as the apparatus of formalization of complex algorithms that convert the information.

There are examples of the use of neural networks in on-board intelligent decision support systems for managing complex dynamic objects and diagnosis of its condition [2].

The problem of forecasting requires an individual approach to processing data when using traditional neural networks. Such an approach requires a lot of time and resources, as determined during the design of a large number of parameters of both the neural network model and the processing parameters of time series (prediction window, the time horizon, etc.). It is therefore necessary to develop a tool automatic settings and automatic choice of architecture. Such a mechanism can be implemented using neural network complexes [3, 4, 5, 6, 7]. There are several examples of such complexes, they are mainly designed to solve a particular practical problem, and have restrictions on use.

So an important task is to develop a neural network structure, which allows you to automatically search for the optimal solution for multivariate time series. Such structures can be used to determine the preferred structure of a neural network, or to determine the individual characteristics of the structure and process. Most often, these structures are implemented based on ensembles of neural networks, and therefore also open to question is the optimization of learning algorithms.

2. REPRESENTATION OF THE INPUT DATA

In the simplest case, the formation of the input vector of multivariate time series can be represented as follows: from the N time series takes the window size K, and concatenated into a single input vector, thus the input vector length is equal to K * N. At the exit, we have a vector that contains N elements, in the case of one-step prediction.

3. ENSEMBLE (COMMITTEE) OF NEURAL NETWORKS

Ensemble (committee) of neural networks (ENN) - is a set of several individual neural networks, which are independently solving one problem prediction [4, 5, 7]. Particular solutions of the individual neural networks are fed to the gating module, which gives the final decision (see Figure 1.).

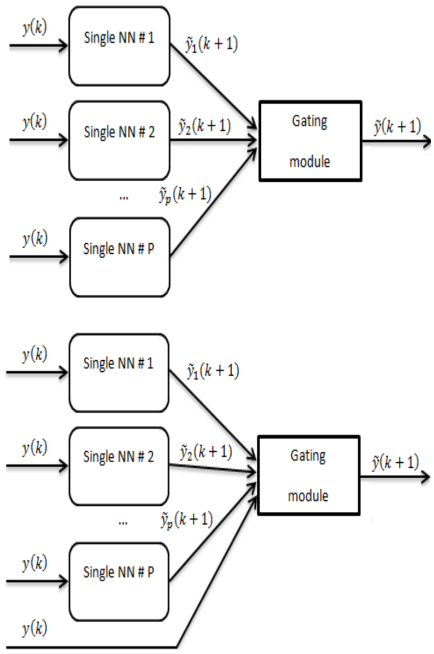

Fig. 1 - Two types of ANS: gating module examines the solution of particular neural network (up), gating module also receives external inputs (down).

As the gating module can be used by different types of neural networks, the averaging operation, calculating the median, and other ways to track the time series. When using a neural network, gating module learns to solve the problem using the output of experts.

There are two approaches to the use of ENN. The first approach involves the use of homogeneous ensemble of neural networks. This approach aims to obtain a stable output value, the elimination of the empirical selection of certain parameters.

On the other hand, ENN can be used for more complex model of the analyzed data to determine more sophisticated dependencies [3]. One variant of this prediction values of the time series with different "time window" [4].

Both approaches significantly extend the capabilities of neural networks and allow to exclude some of the empirical work.

4. TRAINING ENSEMBLES OF NEURAL NETWORKS

The basic operation here is a mutation [8], represented by formula (1), in which the connection weight is perturbed random value. Also, we calculate probability whether the operation is applied.

$$w_{n,I,j,k+1} = w_{n,I,j,k} + y \cdot Q \,, \qquad (1)$$

where $w_{n,I,j,k}$ is the weight between the i_{th} and j_{th} nodes of the of the n_{th} network in the population at the k_{th} epoch of training, Q is a number drawn at random from a Normal (0,1) distribution and y is some fixed (or weight dependant) multiplier. Typically a given network's ability to reproduce into the next epoch's (generation's) population is determined by its relative average Euclidean error, either in relation to the training set, or the validation set. Usually the population of networks is ranked at each epoch (generation), with the fittest Networks having a higher probability of offspring in the next generation.

The algorithm of formation of population consists of the following steps:

1. Initializing weights of ANN from ensemble.
2. Training elements of the ENN.
3. Survival of the fittest (as a criterion of fitness can be used one measure of error of ANN, or a combination thereof). Survival implies the elimination of a set of less adapted ANN, and duplication of the fittest.
4. Mutation of weighting links ANN ensemble.
5. Go to (2) or stop if stopping criteria met.

The perturbation in 2) described above in (1). In 3) the population of networks is ranked in descending order by relative fitness. In 4) part of the fittest networks are replicated twice and inserted into the next population, the next part are replicated once and inserted. The bottom 20% is discarded from the population. This is then repeated until the criteria for stopping is met. Elitism is also implemented, such that one of the two replicas of the fittest individual Neural Network passed into the next generation does not have its weights perturbed.

5. TWO-LEVEL LEARNING MODEL

Using only a single ensemble at designing of a forecasting model often does not give the desired result, and leads to development that is repeated over and over again with varying architecture, varying the types and learning algorithms of neural networks [6]. This is due to the fact that some of the data make better use of one type or architecture, for others the opposite type, or architecture. It is proposed to use a two-level learning model [9]. This model allows realizing heterogeneity of the complex neural network (see Fig. 2).

The first level of structure is a set of ensembles of heterogeneous networks.

First, it can be different types of neural networks [3,10, 11].

Second, it may be similar networks (may have a similar architecture and parameters), but with different time-series analysis [4]. So the elements of this model may use different delays between the window and the predicted value of forecasting, to determine the delay in data being processed. Also, time series can be processed in a different scale, so each element will be fed data at a different time scale. This architecture may account for distant effects (see Fig. 3).

Third, it may be similar networks with different training parameters. This architecture can be used to find the optimal parameters presented neural network model.

The second level is a hybrid ensemble which is formed of the most successful elements of the ensembles of the first level after passing a certain number epochs of training.

Also, as a second level of the ensemble can be used the ensemble (or a single network) supervisor [4], processing the output values of all elements of the first level.

Fitness the first-level ensembles are analyzed for forming hybrid ensemble. If one of the ensembles is much more suited than others, the hybrid ensemble is formed only on this basis. In other cases, a hybrid ensemble made up of the fittest neural networks most accurate ENNs.

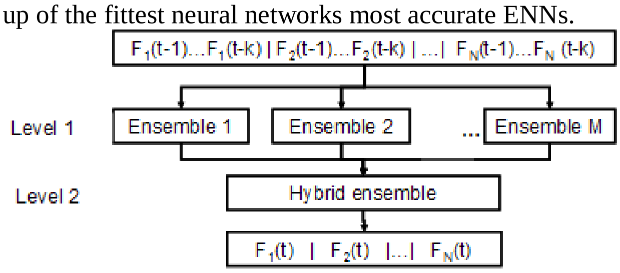

Fig. 2 – Two-level learning model.

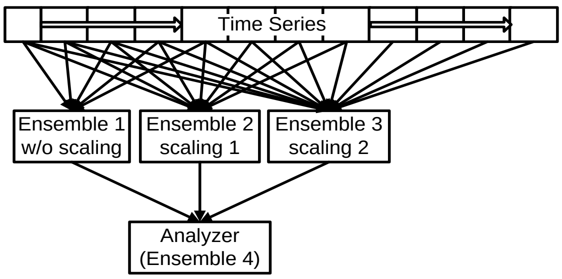

Fig. 3 - Predictive model with different time scales (for three ENN)

6. THE EXPERIMENTAL PART

The basic version of the test data for the experiments is the telemetry data, as well as time series with different characteristics (see Table 1.).

For forecasting, we used data from six aircraft sensors that measure the following parameters: distance, speed, heading angle on the first line of communication, pitching angle of the first line of communication, heading angle on the second line of communication and pitching angle according to the second line.
We also used a set of «The Santa Fe Time Series Competition Data», synthetic generated multivariate time series of states of objects.

Table 1. Description of data sets

Name	Description	Size
A	Generated by the laser, one-dimensional.	1000
B1,B2,B3	Physiological data, with intervals of 0.5 s.	34000
C	Price counts of the Swiss franc exchange rate on U.S. dollar; 08/07/1990 - 04/18/1991.	30000
Gen	Computer-generated data states of the objects, multi-dimensional.	28000
P39	Information about the state of the position sensors of the object in space.	2181
E	Astrophysical data. Set of measurements of the light curves star PG1159-035 in March 1989, with 10 second intervals.	27204

ENN application reduces the scatter of results, as the final relation is obtained by averaging for each element of the ensemble. This is clearly illustrated in Fig. 4. Here the schedule of convergence ENN has a smoother shape than the single neural network. The above learning algorithm allows to optimize the parameters of ENN. So if you initialize the ensemble of different parameters (learning rate, the size of the hidden layer), at the end of training we can get a set of models with parameters close to optimal.

Telemetric information about the object was considered in two ways (see Fig. 5). In the first case, the data from the sensors are analyzed directly, and then the resulting forecast determines predicted state. In this case, training and processing is done on the initial data.

In the second case the input data is classified into a number of states of the object, which is analyzed, so it significantly reduces the the volume of data being processed. Comparison of predicted and actual data is shown in Fig. 6.

In this case, we have not performed considerable pre-processing of data. But to get the best quality of results is also necessary to analyze the input data. It is necessary to analyze the correlation between different variables, to allocate the independent inputs, to determine the precursors, if they exist, to remove noise. Thus we can achieve significant reductions in processing information.

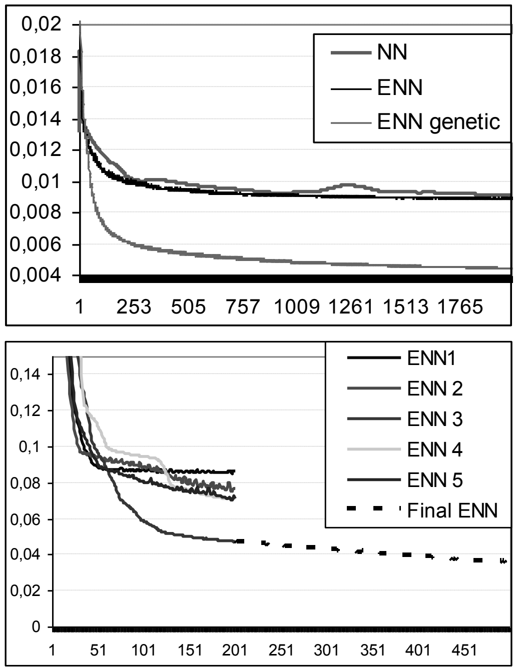

Fig.4 – Graphs of convergence: the convergence of different structures(up), the learning process of the proposed structure(down).

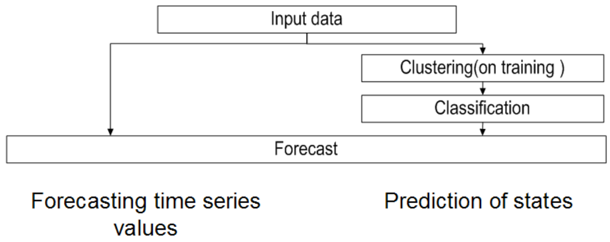

Fig. 5 - Two approaches states of the objects data processing.

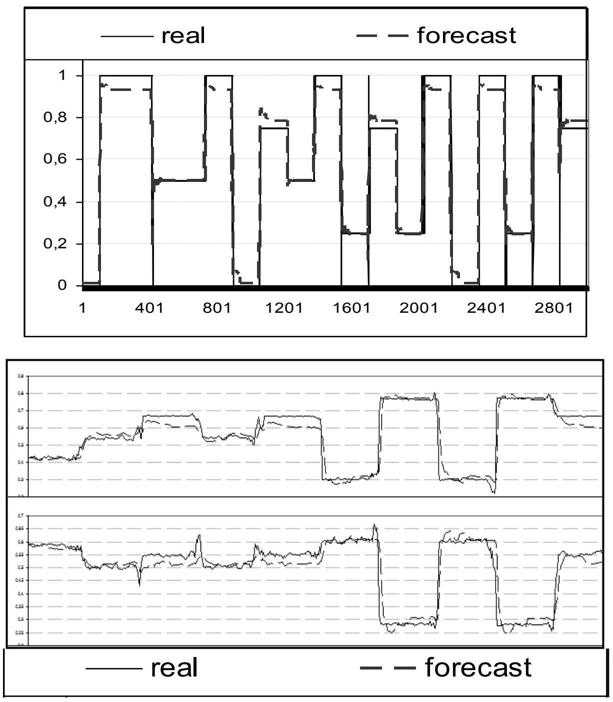

Fig.6 - the actual data and forecasts, time series and state.

Evaluation of the accuracy of the model is also an important component of testing. For this purpose, a series of experiments was performed. The different single neural network, different ENN, and the proposed model with different sets of ENN were trained on the same data. The results are shown in Table. 2. : the average error on time series (Er. 1), and on classified data (Er. 2).

Table 2. Estimation of accuracy

Model	Er. 1, %	Er. 2, %
Single ANN	6,7-24	5,5-22
ENN	5,6-17,2	2,7-14
The proposed model	5,6-8,3	2,7-7,7

7. CONCLUSION

In this paper, a generalization of existing models of complexes based on the ENN has been submitted. We also proposed the use of two-level training algorithm of heterogeneous ensembles, using training with elements of an evolutionary strategy for processing telemetry data of the states of small airborne objects.

The software module was developed that implements the proposed method and a series of experiments performed to determine the characteristics of software modules and models. The results showed that two-level learning algorithm allows to build neural network model for the given ENN close to optimal in the specified range of parameters. There is no need for a set of experiments to get a good result.

Disadvantages include increased training time, by using a set of single neural networks.

8. REFERENCES

[1] Patterson Dan W. Artificial neural networks: Theory and applications / Dan W. Patterson. –New York: Prentice Hall. – 1996. – 477 p.

[2] Emelyanov Yu.G. Neural orientation angles and distance of the spacecraft sensor control system / Yu.G. Emelyanov, K. A. Konstantinov, S. V. Pogodin, etc. // Software Systems: Theory and Applications. - 2010. - № 1 (1). - C. 45-59. (in russian)

[3] Chitra A. An Ensemble Model of Multiple Classifiers for Time Series Prediction / Chitra A., Uma S. // International Journal of Computer Theory and Engineering. . – June, 2010 – Vol. 2, No. 3.

[4] Shugai Yu.S. A Neural Network Algorithm for the Prediction of Events in a Multidimensional Time Series and Its Application to the Analysis of Data in Cosmic Physics / Yu.S. Shugai, S.A. Dolenko, I.G. Persiantsev, Yu.V. Orlov. // Pattern Recognition and Image Analysis. – 2006. – V.16, №1. – P. 79-81.

[5] Dolenko S.A. Neural Network Segmentation of Time Series / S.A.Dolenko, Yu.V.Orlov, I.G.Persiantsev, Yu.S.Shugai // Pattern Recognition and Image Analysis. – 2003. – Vol.13, No.3. – P.433-440.

[6] Persiantsev I.G. Adaptive construction of hierarchical neural network systems for classification and for Segmentation of time series / I.G. Persiantsev, Lectures on neuroinformatics MSU. – M., 2006.

[7] Dziuba D. A. Time series forecasting based on the individual neural networks and committees of neural networks: A comparative experiment / A. N. Chernodub, D.V. Novitsky, D. A. Dziuba // Scientific session of the MEPHI - 2011. Collection of scientific papers. - 2011. - Part 2. - S. 192-201.

[8] Fieldsend J.E. Optimizing forecast model complexity using multi-objective evolutionary algorithms / J.E. Fieldsend, S. Singh // in Coello CAC, Lamont GB (eds) Applications of Multi-Objective Evolutionary Algorithms, World Scientific. – 2004. – P. 675-700.

[9] Marushko Y. Forecasting Multivariate Time Series Using Ensembles of Neural Network / Y. Marushko // PRIP'2011 -The Eleventh International Conference on Pattern Recognition and Information Processing May 18-20, Minsk, Belarus. – Minsk, 2011. – P. 172-176.

[10] Gheyas Iffat A. A Neural Network Approach to Time Series Forecasting / Iffat A. Gheyas, Leslie S. Smith // Proceedings of the World Congress on Engineering 2009. – Vol 2. – July 1 - 3, 2009.

[11] Fieldsend J.E. Pattern Matching and Neural Networks based Hybrid Forecasting System / J.E. Fieldsend, S. Singh // ICAPR 2001, Lecture Notes in Computer Science,. – P. 72–82.