УДК 551.462:629.563

A.U. Zaiets

DETERMINING OPTIMAL PARAMETERS OF A NEURAL NETWORK FOR SEA ICE THICKNESS PREDICTION

In this paper we consider the problem of determining optimal neural network architecture in order to predict ice thickness. This paper puts forward an approach to determining the number of neurons in the intermediate layer.

Keywords: *ice loads; neural network; number of neurons; automated prediction; oceanengineering construction.*

Розглядається питання визначення оптимальної архітектури нейронної мережі для прогнозування товщини льоду. Пропонується підхід до визначення кількості нейронів на проміжному шарі.

Ключові слова: *льодовінавантаження, нейронні мережі, число нейронів, автоматичне прогнозування, океанотехнічні споруди.*

Рассматривается вопрос определения оптимальной архитектуры нейронной сети для прогнозирования толщины льда. Предлагается подход к определению количества нейронов на промежуточном слое.

Ключевые слова: *ледовые нагрузки, нейронные сети, число нейронов, автоматическое прогнозирование, океанотехнические сооружения.*

Offshore energy resources development refers to the most promising and important areas of economic development. Currently, operations in this direction are carried out by the State Joint-Stock Company (SJSC) «Chernomorneftegaz», the main facilities of which are located on the shelf of the Azov-Black Sea basin. Every year in winter those areas are covered with ice. In order to provide reliable and long-term exploitation of those facilities one has to estimate parameters of ice loads impact on them [1].

Ice properties produce great influence on ice loads. When calculating ice loads at early stages of designing one uses standard value of ice parameters, though there is no consistent approach to selecting such parameters [2; 3]. Some researchers suggest taking the highest possible values as calculated ones from the available line. Others recommend inputting the most probable values into calculations. While taking into consideration the physical picture [4] of the drift ice interaction with objects in the specific areas of the continental shelf the latter approach is the most reasonable.

The object of the article consists in selecting an optimal neural network architecture, which allows predicting of ice thickness. The main factor influencing ice loads value and is required for accounting in the oceanengineeringconstruction practice when automating calculation is the computative ice thickness [5].

Automated ice thickness prediction using information technology and neural networks in particular is one of the most up-to-date methods of estimating changing of ice cover and thickness.

When building a neural network one of the major problems is the one of selecting optimal neural network architecture. This article discusses selection of neural network architecture, in particular the selection of the number of neurons in the intermediate level.

The principle of building a model of a hybrid neural network. As a tool which allows predicting sea ice thickness in the Azov Sea for calculation of the load estimation, while analyzing time-varying specification, we select a neural network of direct propagation, which is a gradient descent algorithm.

A model of a neural network, which has three layers including an input layer, hidden layer and output layer has been worked out. The number of neurons in the each layer is *n*, *m* and *l* respectively.

An important stage when designing a neural network is a process of its training, which looks like this [6]. The initial weights W_1, W_2 and the thresholds Θ_1, Θ_2 of the neural network aregenerated by a random sequence mat-ching as being the given data. The common sigmoid function with a single slope parameter and range [0,1] is selected as an activation function between the layers of the neural network

$$f(x) = \frac{1}{1+\exp(-x)}. \quad (1)$$

Q groups of training samples

$$X_p = \{x_{p1}, x_{p2} .. x_{pn}\} (1 \le p \le Q)$$

should be introduced to the neural network, and then the total average error of the neural network is calculated

$$E_{MSE} = \frac{1}{2Q} \sum_{p=1}^{Q} \sum_{k=1}^{l} (T_k^P - V_{L_k}^p)^2, \quad (2)$$

wherein E_{MSE} is the overall average error from the Q groups of training samples, T_k^P and $V_{L_k}^p$ are the expected and the actual outputs of the neural network respectively.

Matching algorithm function is the following

$$F_v = \frac{1}{1+E^v{}_{MSE}} \quad (3)$$

wherein F_v is a matching to the v ' svalue in the sample, and $E^v{}_{MSE}$ is the average error of the v 's value $(1 \le v \le N)$. Considering the principle «survival of the fittest», we mention that the matching between the selection process, the selection probability p_s is calculated using (4), so that the selection process will be completed.

$$p_s(U^i) = \alpha p_f(U^i) + (1-\alpha) p_d(U^i) = \alpha \frac{f(U^i)}{\sum_{i=1}^{M} f(U^i)} + (1-\alpha)\frac{1}{N} e^{C_i/\beta}. \quad (4)$$

In the relation (4) $\alpha, \beta \in [0,1]$ are the constants, $f(U^i)$ is the matching function, C_i is the concentration of the initial population problem U^i.

Adaptive crossing probability p_c and the change probability p_m are calculated using (5) and (6) respectively, the crossing processes and changes will be completed, so that the new generation values are generated.

$$p_c = \begin{cases} k_1 \dfrac{f_{\max} - f_m}{f_{\max} - f_{MSE}}, f_m \ge f_{MSE} \\ k_2, f_m < f_{MSE} \end{cases}, \quad (5)$$

wherein $k_1, k_2 \in [0,1]$ are the constants, $f_{\max}$ is the highest matching in the population crossing, f_m is the highest matching of two crossed individuals and f_{MSE} is the average matching of the current population.

$$p_m = \begin{cases} k_3 \dfrac{f_{\max} - f_l}{f_{\max} - f_{MSE}}, f_l \ge f_{MSE} \\ k_4, f_l < f_{MSE} \end{cases}, \quad (6)$$

wherein $k_3, k_4 \in [0,1]$ are the constants, f_l is the individual's mutation.

The calculation is in progress until the average error from the optimal value E_{MSE} satisfies $E_{MSE} < \varepsilon_1$, wherein ε_1 is the global search accuracy. An adaptive algorithm is used for the neural network weight and threshold optimization using (7), until the average error of the optimal value E_{MSE} satisfies $E_{MSE} < \varepsilon_2$, wherein ε_2 is the final accuracy.

$$\Delta W = \eta \frac{\partial E_{avg}}{\partial W} \tag{7}$$

wherein ΔW is the weight adjustment, η is the training rate and is calculated using (8)

$$\eta(k+1) = \begin{cases} 1.05\eta(k), & if \ \ E_{MSE}(k+1)<E_{MSE}(k) \\ 0,7\eta(k), & if \ \ E_{MSE}(k+1)>1,04E_{MSE}(k) \\ \eta(k), & another \end{cases}. \tag{8}$$

After all operations trained weights W_1, W_2 and thresholds Θ_1, Θ_2 of the neural network preserve and are ready for the testing samples.

Selecting the neural network architecture. When building a neural network one of the major problems is the one of selecting the optimal neural network architecture. Typically multilayered, and most often three layered neural networks of direct propagation are used when predicting the timing series. The external view of such network is presented in the figure 1.

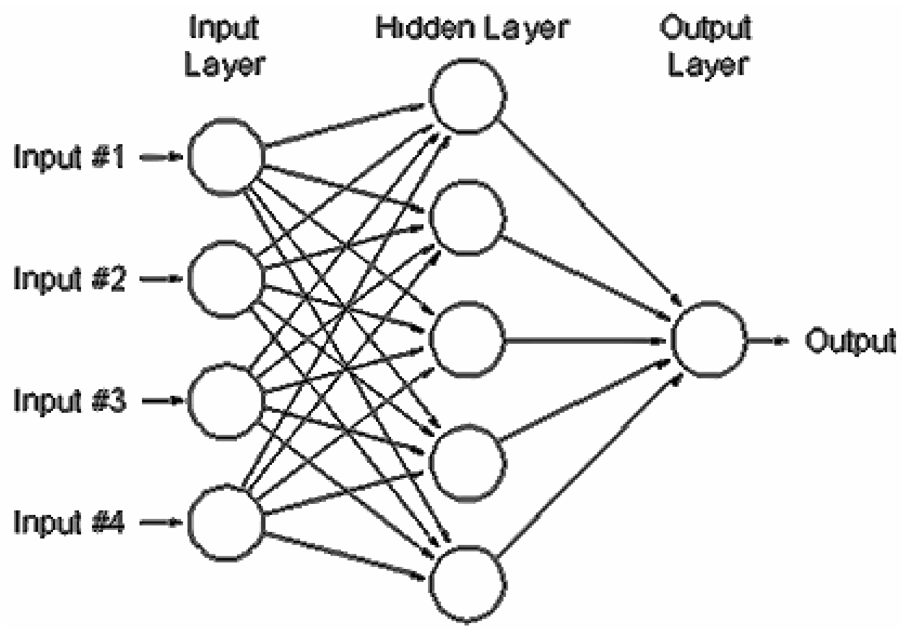

Fig.1. The external view of the direct propagation neural network

And the standardized sample in this case will take values in the range [0-1]. Standardizing allows improving the rate and quality of neural network training. Let us consider ice thickness predicting problem. Experimental data are given in the table 1.

Table 1

№	Windspeed, mps	Water Temperature, °C	Depth, m	Saltiness, %	Icethickness, sm
1	11	-0,2	3,73	1,01	35
2	10	-0,2	3,69	1,03	40
3	10	-0,4	4,0	2,34	35
4	11	-0,4	5,82	6,05	40
5	8	-0,4	7,0	6,66	30
6	10	-0,7	8,08	9,4	20
7	10	-0,7	8,29	11,51	20
8	17	-0,7	8,79	11,22	25
9	0	-0,8	12,00	11,89	40
10	9	-0,8	12,48	11,83	50
11	8	-0,8	13,00	11,86	40
12	6	-0,8	12,28	11,12	30
13	6	-0,8	12,00	11,41	40
14	4	-0,8	11,00	11,70	30
15	6	-0,3	6,61	11,60	25
16	0	-0,2	7,43	11,53	25
17	0	-0,5	6,89	11,72	30
18	17	-0,8	11,27	11,64	40
19	15	-0,8	11,00	11,35	40
20	10	-0,7	11,00	11,24	50

Parameter values are fed to the input, and in output we get the prediction value. Mathematical formulation of the problem can be described as follows. Suppose we are given a sample $\{x_i, y_i\}_{i=1}^{n}$ wherein x_i is the parameters vector describing the precedent, y_i is the value of the dependent variable in the i precedent, n is the number of precedents. We need to build a neural network that allows getting the prediction $\hat{y} = f(x)$. For the case of multifactor dependency the input parameters can be described in the matrix form X_i, and the prediction needed as following $\hat{y} = f(X)$.

The problem solution of building the predicting network can be carried out in several stages:

- Selection of the network structure:
- Input data preparation.

The preparation of data that will be used for the network training is in progress at the second stage. Most often such preparation comes down to scaling or standardization of data. The following functions may be used for it

$$a_i^{'} = \frac{a_i - Ma}{\sqrt{Da}}; \qquad Ma = \frac{1}{n}\sum_{i=1}^{n} a_i; \qquad Da = \frac{1}{n-1}\sum_{i=1}^{n}(a_i - Ma)^2$$

or

$$a_i^{'} = \frac{a_i - a_{i,\min}}{a_{i,\max} - a_{i,\min}}.$$

We will use the first 15 rows for the network training, and the last 5 to verify the functionality.

Network parameters calculation (with one hidden layer):

- We determine the number of inputs $N_x = 4$;
- We determine the number of outputs $N_y = 1$;
- We determine the number of trained samples $Q = 15$;
- We calculate the estimate of the required number of the synaptic weights N_w based on the inequality [7]

$$\frac{N_y Q}{1 + \log_2(Q)} \le N_w \le N_y \left(\frac{Q}{N_x} + 1\right)(N_x + N_y + 1) + N_y. \tag{9}$$

The estimate of the number of neurons in the hidden layer is calculated using the relation [4]

$$N = \frac{N_w}{N_x + N_y} \tag{10}$$

It is possible to use another formula for estimating the neurons number [4]

$$\frac{Q}{10} - N_x - N_y \le N \le \frac{Q}{2} - N_x - N_y \tag{11}$$

Using (9) and (10) we get $0{,}6 \le N \le 5{,}9$. Using the inequality (11) gives the following result $0 \le N \le 2{,}5$. Thus let us take the neurons number in the hidden layer as 2.

Conclusion. We have considered the problem of determining the number of neurons in order to select the optimal network architecture that allows predicting ice thickness in the Azov-Black Sea basin. Using the suggested algorithm will allow selecting the number of neurons with regard to the data amount and required data accuracy.

The calculations showed that the optimal number of neurons in the intermediate layer for the network considered is two neurons. We plan to conduct numerical experiments using the gotten network architecture for getting predicted values of ice thickness and verification with field data.

REFERENCES

1. *Воздействие поверхностных гравитационных волн на прибрежные океанотехнические сооружения / В.М. Кушнир, В.Р. Душко, В.А. Крамарь // Восточно-Европейский журнал передовых технологий. −№ 6/5 (66). − 2013. − С.38-41.*
2. *Лосет С. Воздействие льда на морские и береговые сооружения: Учебн. пособие для вузов / С. Лосет, К.Н. Шхинек, О. Гуместад, К. Хойланд. – СПб.: Изд-во Лань, 2010. – С.109-187.*
3. *Афанасьев В.П. Разработка и совершенствование методов расчета ледовых нагрузок на вертикальные сооружения шельфа от морских ледовых образований: дис. ... док.техн. наук: 05.23.07: защищена 2007, утв. 2007/ Афанасьев Владимир Петрович. – М., 2007. – 329 с.*
4. *Боровская Р.В. О возможности использования спутниковой информации в исследовании Азово-Черономорского бассейна // Сб. научн. трудов ЮгНИРО. – Керчь: Изд-во ЮгНИРО, 2008. – 250 с.*
5. *Заец А.Ю. Применение нейронных сетей в задачах прогнозирования толщины морского льда // Севастополь: СевНТУ, 2014. − С. 228-233.*
6. *Hong Lin. A hybrid neural network model for sea ice thickness forecasting // Journal of Hydrologyvol. − (375). − 2009. − P. 450-458.*
7. *Callan R. The Essence Of Neural Networks // Prentice Hall Europe. − 1999. −248 p.*

Стаття надійшла до редакції 25.11.15

Рецензенти:

кандидат технічних наук, доцент, професор Одеського національного морського університету **О.Б. Ляшенко**

доктор технічних наук, професор, зав. кафедрою «Теоретична і прикладна механіка» Одеського національного морського університету **В.О. Гришин**