DOI: https://doi.org/10.15276/hait.07.2024.18
UDC 004.942

Method of reference models for synthesis of intellectual systems of nonlinear dynamic objects identification

Oleksandr O. Fomin[1)]
ORCID: https://orcid.org/0000-0002-8816-0652; fomin@op.edu.ua. Scopus Author ID: 57103429400
Viktor O. Speranskyy[1)]
ORCID: https://orcid.org/0000-0002-8042-1790; speranskyy@op.edu.ua.Scopus Author ID: 54401618900
Andrii A. Orlov[1)]
ORCID: https://orcid.org/0000-0002-3256-5044; 9901020@stud.op.edu.ua. Scopus Author ID: 57202287340
Oleksii V. Tataryn[1)]
ORCID: https://orcid.org/0009-0005-3888-6569; otataryn@stud.op.edu.ua
Denys V. Kushchevyi[2)]
ORCID: https://orcid.org/0009-0008-2400-9989; biostar1992@gmail.com
[1)] Odesa Polytechnic National University, 1, Shevchenko Avenue, Odesa, 65044, Ukraine
[2)] Luxoft Ukraine, 9v, Kartamyshivska Street, Odesa, 65000, Ukraine

ABSTRACT

The paper is devoted to resolving the contradiction between the accuracy of modeling nonlinear dynamic objects and the speed of models building under conditions of limited computing resources. The purpose of the work is to reduce the time for building models of nonlinear dynamic objects with continuous characteristics while ensuring a given modeling accuracy. This goal is achieved by further developing the method of synthesing intelligent systems based on the superposition of pre-trained reference models in the form of neural networks reflecting the basic properties of the object. The scientific novelty of the work novelty consists in the development of a method for identifying nonlinear dynamic objects in the form of neural networks with time delays based on a set of pre-trained neural network models that reflect the basic properties of the subject area. In contrast to the traditional approach based on pre-trained neural networks the developed method allows building models of lower complexity and with shorter training time while ensuring the required accuracy. To determine the initial parameters of the model, expressions based on the superposition of reference models in the form of neural networks are proposed. The practical usefullness of the work consists in the development of an algorithm for the method of reference models for training neural networks with time delays in the tasks of identifying nonlinear dynamic objects with continuous characteristics, which can significantly reduce the training time of neural networks without losing the accuracy of the model. The value of the study lies in determining the area of effective use of the proposed method, namel y, in the availability of a sufficient amount of qualitative data for the building of reference models. Insufficient data or poor data quality can significantly reduce the accuracy of reference models and, as a result, significantly reduce the training time of the target model.

Keywords: Nonlinear dynamics; identification; neural networks with time delays; pre-training

***For citation*: Fomin O. O., Speranskyy V. O., Orlov A. A., Tataryn O. V., Kushchevyi D. V. "Method of reference models for synthesis of intellectual systems of nonlinear dynamic objects identification". *Herald of Advanced Information Technology.* 2024; Vol. 7 No.3: 262–274. DOI: https://doi.org/10.15276/hait.07.2024.18**

INTRODUCTION

In the context of continuous growth of the complexity of control objects, modeling plays a key role in solving a set of scientific and applied problems of control, diagnostics and management of complex objects. The current stage of development of modeling, which is mainly based on the use of intelligent technologies [1, 2], is marked by a number of requirements from practice both for the high accuracy of models and for the speed of their construction [3].

The achievement of high accuracy of modeling today is carried out through the use of machine learning methods, in particular, neural networks (NN) [4]. This apparatus is well suited to building models of objects with a high degree of internal complexity and interaction, especially multidimensional nonlinear dynamic objects [4, 5]. However, the use of such methods is often associated with high computational complexity, which leads to significant time spent on building models [4, 5], [6].

The problem of increasing the speed of simulation remains one of the most urgent, especially in areas related to model personalization [7, 8], [9], where solutions must adapt to changes in user behavior and the environment (e.g., authentication tasks, biomedical applications, human-machine systems), during rapid diagnostics and real-time functioning (e.g., assessment of the psychophysiological state using nonlinear dynamic models of oculo-motor human systems [8]).

Therefore, despite the success of machine learning, this area requires new approaches and methods that will significantly reduce the time for creating and training models while maintaining high modeling accuracy.

Based on the analysis of the advantages and disadvantages of existing methods of machine learning [4, 9], [10], the paper attempts to eliminate the contradiction between the accuracy of modeling complex objects and the speed of building models by developing a new method of synthesis of intelligent systems for the identification of nonlinear dynamic objects, which combines high accuracy of modeling and the speed of model training.

LITERATURE REVIEW

In recent years, the problem of increasing the speed of NN learning has become widespread among researchers [9, 10], [11]. One of the common approaches to solving this problem is to optimize the NN architecture, which involves reducing the number of model parameters without significantly losing their performance. Examples of such approaches include network simplification techniques such as pruning, quantization, and model compression) [12, 13]. These methods can significantly reduce the time required to train models, especially when used in limited computing environments.

The second relevant way to increase the speed of NN learning is the use of accelerated learning algorithms, such as stochastic gradient descent with momentum (SGD with momentum) and adaptive optimization methods (for example, Adam, RMSprop) [4, 10]. These methods make it possible to accelerate NN convergence by improving the strategy of updating weights and reducing the number of epochs required to achieve a given accuracy.

Another way to accelerate the NN learning process is transfer learning [14, 15]. The main advantage of this method is the ability to use models previously trained on a dataset from one subject area to solve target problems from another subject area. This approach is used in conditions where the amount of specific data is limited or there are no high-quality datasets for the target task. Training transfer is especially effective when dealing with large NN, such as deep convolutional networks, which require significant computing resources to train them [16, 17].

However, despite its advantages, the transfer of training has a number of disadvantages [18, 19]. First, migrated and adapted models may not provide sufficient accuracy on target data, especially if the overall and target data differ significantly in their characteristics. Second, transference of learning can contribute to the inheritance of systematic errors or biases contained in the model transferred to the target task. This has a negative impact on the results of modeling the target task.

A special case of transfer of training can be considered pre-training, in which the model is first trained on a large set of data of a general nature, and then retrained (fine-tuned) on more specific data of the target task [20, 21]. This approach allows for a significant reduction in the time and computing resources used during the retraining of the target model, compared to full training of the model on the data of the target task. As a result, pre-trained models converge faster and require fewer resources to achieve a given model quality [21].

However, like transference of learning, prior learning has its drawbacks [20, 21]. In particular, if the data for pre-training and the target task differ significantly, this can lead to incorrect initialization of the target model and a decrease in its accuracy. In addition, pre-training may require significant computing resources at the initial stage of training on a large dataset of a general nature, which limits its application in a resource-constrained environment.

Thus, although the methods of transfer and pre-training can significantly speed up the process of constructing NN, their application requires careful analysis and consideration of the specifics of the problem in order to minimize possible shortcomings and ensure high quality modeling.

However, technologies for transferring learning and pre-training NN have become an integral part of many research and development in the field of artificial intelligence. They have proven their effectiveness in natural language processing tasks (BERT Natural Language Processing Network, GPT Text Generation Network) [22], construction of computer vision systems (pre-trained DenseNet convolutional networks, VGG) [23], biomedical research, and human-machine interfaces [24].

This spread of the approach to building models based on previous NN training became possible due to its practical advantages: a significant reduction in the cost of training models and an increase in their efficiency in real-world applications, where not only accuracy but also speed of work is important. In addition, pre-trained models can be easily adapted to new tasks, making them indispensable tools in the face of dynamically changing requirements and tasks.

This direction also looks promising in the problems of identification of nonlinear dynamic objects. At the same time, there is a lack of work in the field of preliminary training of NN that simulate nonlinear dynamic properties of objects with continuous characteristics.

Based on the above, the paper develops an approach to the construction of NN on the basis of preliminary training, which is able to effectively cope with the requirements of modern modeling tasks, in particular the identification of nonlinear dynamic objects.

STATEMENT OF THE PROBLEM

The formal formulation of the problem of preliminary training of NN is as follows.

Let S is a domain containing problems of a general nature and for which there is a nested D_S dataset of sufficient *size* N_S:

$$D_S=\{(\mathbf{x}_i^S, y_i^S)\}, \quad (1)$$

where $\mathbf{x}_i^S$ is the vector of independent variables, y_i^S is the corresponding label (target variable), i=1,…, N_S.

Let $f_{\theta S}(D_S)$ is a general (rough) model with parameters θ_S that is trained on the D_S dataset.

Let T_k is a specific problem from the set of objective problems **T** defined in the domain S (k=1,...,p, p is the size of the set of problems **T**) for which there is a labeled *dataset* D_{Tk} of limited size N_{Tk}:

$$D_{Tk}=\{(\mathbf{x}_j^{Tk}, y_j^{Tk})\}, \quad (2)$$

where $\mathbf{x}_j^{Tk}$ is the vector of independent variables, y_j^{Tk} is the corresponding label (target variable), j=1,…, N_{Tk}.

Let $f_{\theta Tk}(\theta_S, D_{Tk})$ is the target (exact) model for the problem T_k with parameters θ_{Tk}, which is obtained by training *the* $f_{\theta S}$ model on the D_{Tk} dataset.

The task of pre-training is to determine the following parameters θ_S of the general model $f_{\theta S}(D_S)$, when used as initial values $\theta_{Tk0}=\theta_S$ in the tasks of retraining each of the p target models $f_{\theta Tk}(\theta_S, D_{Tk})$, a given level of accuracy (permissible error $E_{\theta T}$) is achieved for a minimum average period of time:

$$\begin{cases} \theta_S = \underset{\theta}{arg\ min}\ t_{\theta T}(f_{\theta Tk}(\theta, \mathbf{x}_j^{Tk})) \\ L_T(f_{\theta Tk}(\theta_S, \mathbf{x}_j^{Tk}), y_j^{Tk}) < E_{\theta T} \end{cases}, \quad (3)$$

where $t_{\theta T}$ is the average duration of additional training among p target models $f_{\theta Tk}(\theta_S, D_{Tk})$; the average number of epochs of model training can be used as the value of $t_{\theta T}$; L_T is the loss function adopted for the set of target models $f_{\theta Tk}(\theta_S, D_{Tk})$.

When conditions (3) are met, *the* $f_{\theta S}(D_S)$ model is said to be pre-trained. As a *loss* function of L_T, the following metrics are commonly used: mean absolute error (*mae*) and root mean square error (*mse)* [25, 26]:

$$mae = \frac{1}{N_{Tk}} \sum_{j=1}^{N_{Tk}} |\Delta f_{\theta Tk}|, , \quad (4)$$

$$mse = \frac{1}{N_{Tk}} \sum_{j=1}^{N_{Tk}} (\Delta f_{\theta Tk})^2 . \quad (5)$$

where $\Delta f_{\theta Tk} = y_j^{Tk} - f_{\theta Tk}(\theta_S, \mathbf{x}_j^{Tk})$ is the difference between the exact value of the target variable and the value predicted by the model.

To assess the accuracy of the target models $f_{\theta Tk}(\theta_S, D_{Tk})$, the paper also uses the *Huber Loss* function, which combines the properties of *mae* and *mse* metrics, namely: resistance to emissions and sensitivity to small errors.

Huber Loss is defined as follows [26]:

$$h = \begin{cases} \dfrac{1}{2N_{Tk}} \sum_{j=1}^{N_{Tk}} (\Delta f_{\theta Tk})^2, |\Delta f_{\theta Tk}| \le \delta \\ \dfrac{1}{N_{Tk}} \sum_{j=1}^{N_{Tk}} \delta(|\Delta f_{\theta Tk}| - \dfrac{\delta}{2}), |\Delta f_{\theta Tk}| > \delta \end{cases}, \quad (6)$$

where δ is a hyperparameter that defines the switching threshold between *mae* and *mse*.

For the overall assessment of the quality of the learning process of the target model $f_{\theta Tk}(\theta_S, D_{Tk})$, the indicator of learning efficiency $P_{\theta Tk}$ on the target dataset D_{Tk} is used in the form of the following metric [21]:

$$P_{\theta Tk}=L_{Tk}/t_{\theta Tk}, \quad (7)$$

where $t_{\theta Tk}$ is the duration of retraining of the target model $f_{\theta Tk}(\theta_S, D_{Tk})$, the number of epochs of model training can be used as the value of $t_{\theta Tk}$; L_{Tk} is the loss function of the target model $f_{\theta Tk}(\theta_S, D_{Tk})$.

ISSN 2663-0176 (Print)
ISSN 2663-7731 (Online)

PURPOSE AND OBJECTIVES OF THE STUDY

The purpose of the work is to reduce the time for constructing models of nonlinear dynamic objects with continuous characteristics while ensuring the specified accuracy of modeling by further developing the method of synthesis of intelligent systems based on the superposition of previously trained reference models in the form of NN reflecting the basic properties of the object.

To achieve the purpose, the following objectives are established**.**

1. Development of the method of synthesis of intelligent systems for identification of nonlinear dynamic objects on the basis of preliminary training of reference models in the form of NN.

2. Construction of reference models in the form of NN with basic nonlinear and dynamic characteristics.

3. Study of the speed of identification of complex objects using the method of reference models when working with test objects containing nonlinear and dynamic characteristics.

SYNTHESIS OF AN INTELLIGENT SYSTEM BASED ON REFERENCE MODELS

1. Pre-training of reference models

The approach to building NN based on prior learning in practice faces a number of significant limitations. In particular, its application does not always provide the expected reduction in the training time of the target model and may even reduce the accuracy of the simulation.

One of the reasons for the decrease in the effectiveness of the pre-training process is the too general nature and large volume of the original D_S training dataset. Such a situation arises when trying to describe the behavior of objects in the widest possible range of external conditions, modes of operation and under the influence of different input signals. As a result, when solving target tasks that cover a much smaller part of the subject area, the rough model $f_{\theta S}(D_S)$ and the exact models $f_{\theta Tk}(\theta_S, D_{Tk})$ turn out to be excessively complex. This leads to a complication of the process of additional training and a decrease in the effectiveness of the target model.

To solve the problem of excessive complexity and increase the efficiency of training target models, two most general approaches can be distinguished.

Formation of separate specialized training datasets D_{Sr} (r=1,..., q, q is the number of specialized training datasets, $q \le p$). This approach is used in most cases and consists in creating a rough model $f_{\theta Sr}(D_{Sr})$ for each target problem, which describes the general properties of the D_{Sr} domain, followed by additional training on the target dataset D_{Tk}. Under $q=p$, the pre-training problem degenerates into q of individual target problems.

The application of the approach based on the formation of separate specialized training datasets largely negates the advantages of previous training, since for each new task it is necessary to create a separate dataset and train a new model. This leads to a significant investment of time even at the stage of setting the problem.

Using a set of reference datasets D_{Sv} (v=1,..., g, g is the number of basic characteristics of the subject area). This approach used in the article consists in the use of g reference datasets, each of which describes a separate basic property of the study area. On the basis of these datasets, pre-trained $f_{\theta Sv}(D_{Sv})$ reference models are built with parameters θ_{Sv}. By combining and adapting the appropriate reference models, a rough model $f_{\theta Sv}(D_{Sv})$ is constructed, which is characterized by a certain set of characteristics (nonlinear and dynamic) of the object of study. The target model $f_{\theta Tk}(\theta_{Sv}, D_{Tk})$ of objects with certain characteristics is constructed by training the rough model on D_{Tk} dataset.

This approach allows you to maintain the advantages of previous training, since reference models obtained once can be reused for different subject areas and target tasks, significantly reducing the total time and resources for training models without collecting additional data.

A structural diagram of the learning process based on reference models is presented in Fig. 1.

The paper proposes a method of synthesis of models of nonlinear dynamic objects, based on the use of a set of reference pre-trained NN that reflect the basic properties of the subject domain.

The algorithm of the proposed method is to perform the following steps.

Step 1. Selection of basic domain properties and formation of a set of D_{Sv} datasets that reflect the selected properties.

Step 2. Determination of the structure θ_{Sv} of reference models $f_{\theta Sv}(D_{Sv})$ in the form of NN corresponding to the established basic properties of the subject area and preliminary training of reference models on the generated D_{Sv} datasets.

Step 3. Determination of the properties of the objective task from the set of basic properties of the domain and construction of a rough model $f_{\theta S}(D_S)$ based on the superposition of the corresponding reference models obtained in *Step* 2.

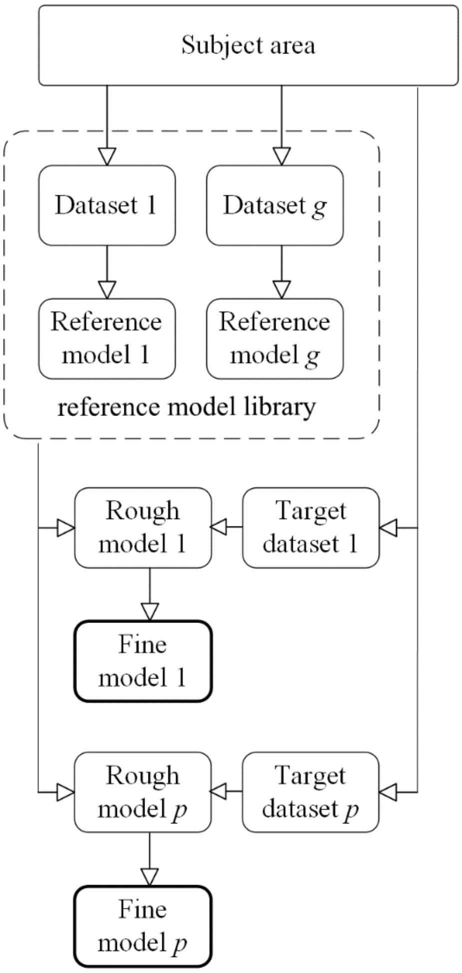


Fig. 1. **Structural diagram of the pre-training process based on reference models**
Source: **compiled by the authors**

Step 4. Training an accurate NN model $f_{\theta Tk}(\theta_{Sv}, D_{Tk})$ based on the rough model $f_{\theta S}(D_S)$ obtained in *Step* 3.

Step 5. Determination of accuracy indicators (4)-(6), $t_{\theta T}$ training time, and model training process efficiency indicator $P_{\theta T}$ (7). In case of unsatisfactory quality indicators of the target model, the transition to *Step* 2 is carried out to adjust the structure θ_{Sv} of the reference models $f_{\theta Sv}(D_{Sv})$, and, if necessary, to *Step 1* to adjust a set of basic domain properties and a set of D_{Sv} datasets that reflect the selected properties.

2. Selection of basic properties of the domain for the formation of a dataset collection

The basic properties of the subject area in the work are understood as the characteristics of objects that reflect the essential aspects of their behavior in different conditions. These properties may include physical, technical, social, or other parameters that are important for solving the model's problem.

The procedure for selecting the basic properties of the domain and forming a set of D_{Sv} datasets reflecting the selected properties is based on the analysis of the domain and is as follows.

1. Determination of the range of tasks that are solved in the subject area, analysis of the properties of the subject area, which are essential for the objects of the subject area, significantly affect the results of modeling and should be reflected in the formation of the D_S dataset.

2. Determination of the types of signals (e.g., periodic, random, pulsed) and operating conditions that best reflect the dynamics of the objects under study and the reaction of the object to which should be included in the D_S dataset. Determination of signal parameters for each type of signal, such as amplitude, frequency, phase, pulse duration, etc., corresponding to the typical operating conditions of domain objects.

3. Formation of the D_S dataset based on the list of basic properties of the subject area established in point 1 and the set of input signals and reactions of the object ($\mathbf{x}_i^S$, y_i^S) formed in point 2.

4. Segmentation of the D_S dataset into separate D_{Sv} datasets (v=1,..., g) in accordance with the defined list of basic properties of the subject area.

This procedure is iterative and involves the adjustment of a set of basic properties of the subject area and types of input signals and corresponding datasets, provided that unsatisfactory quality indicators of the target model are obtained at *Step 5* of the developed method of reference models of NN synthesis.

3. Determination of the structure of reference models and their preliminary training

As mentioned earlier, reference models for the identification of nonlinear dynamic objects are constructed in the form of NN. Today, there are several common methods for modeling nonlinear dynamical objects using NN: dynamic Wiener-type DNN, Dynamic Neuro-SM and TDNN.

Among the above methods, it is worth noting TDNN – NN structures consisting of several layers with direct signal propagation [28, 29]. Due to its simplicity and versatility, TDNNs have become the most widely used in modeling problems of nonlinear dynamic objects. In practice, the most commonly used structure is TDNN, consisting of three layers: input, hidden, and output [29].

The size of the layers in this TDNN structure is defined as follows:

– the input layer consists of M neurons and is responsible for the memory (dynamic characteristics) of the model,

– the hidden layer consists of K neurons and is responsible for the nonlinear characteristics of the model,

– the output layer contains the number of neurons Y, which is equal to the number of outputs of the model.

To determine the structure of the reference models, which are three-layer TDNNs, it is enough to find the number of neurons M in the input layer and K in the hidden layer.

Determine the memory size of the Model M. The number of M neurons in the input layer is chosen in such a way as to best reflect the dynamic properties of the object.

In the case where transient information is available, determining the size of the model's memory can be reduced to determining the duration of that process. This time corresponds to the moment when the system ceases to “remember” its initial conditions and stabilizes.

The following algorithm quantifies the memory size M of the model.

1. Definition of the transient process: the input *signal* $x(t)$ is generated in the form of a step function and the response of the object $y(x)$ to the input signal is recorded.

2. Determination of the sampling step by time Δt, which will be used to level continuous data. The step Δt is chosen small enough to reflect the dynamic properties of the model, but not too small to avoid an excessively large amount of data. At the same time, the observation time of the signal is $n\Delta t$.

3. Determination of the transient time: obtaining the time interval T_s during which the reaction of the object $y(x)$ stabilizes (enters and remains in a certain range around the stationary value *of ys*). In practice, this range is 1-5 % of the constant value *of* y_s.

4. Determining the size of memory: the number of neurons in the input layer of the NN is calculated using the expression $M=T_s/\Delta t$.

Determine the size of the hidden layer of the K model. The number of K neurons is chosen to best reflect the nonlinear characteristics of the object. In practice, the value of K is chosen empirically through additional experiments.

The block diagram of the TDNN obtained taking into account the number of neurons M in the input and K in the hidden layers is shown in Fig. 2. The value of None in the data dimension vector in the figure means the variable number of rows of the dataset.

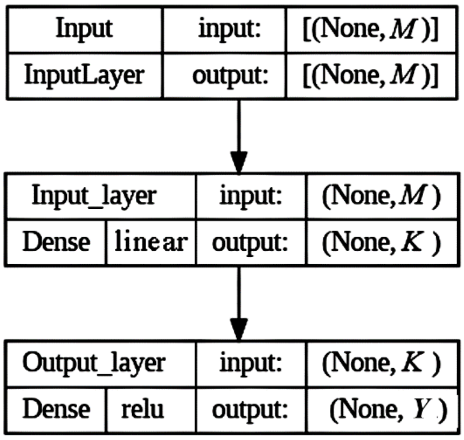


Fig. 2. **TDNN block diagram with *M* Inputs, *K* Hidden Neurons, and *Y* Outputs**
***Source:* compiled by the authors**

Reference models in the form of NN of a certain structure are trained according to the following algorithm.

1. Initialization of the model: setting the initial values of weights and shifts of NN, which will be optimized in the learning process.

2. Input Preparation: For TDNN training, the generated D_{Sv} dataset is converted in a special way, where each element of the input sequence corresponds to one point in time.

$$\mathbf{x}_n^S=[x(t_n),\ x(t_{n-1}),\ \dots\ ,\ x(t_{n-M-1})],\ t_n=n\Delta t,\ n=1,2,\dots. \quad (8)$$

Thus, at each time interval $i\Delta t$ ($1 < i < n$), neurons take into account not only the current input data, but also the data of the previous step $(i-1)\Delta t$ (previous state). This allows the network to "remember" information about past states, which is important for processing the dynamics of an object.

3. Building a reference model: the NN is trained by the backpropagation method of error with updating the network parameters by the Levenberg-Marquardt method to ensure high modeling accuracy and rapid convergence of the learning process.

4. Constructing a rough model based on a composition of relevant reference models

After the pre-training process of the set of reference models is completed, a rough model $f_{\theta S}(D_S)$ is built on their basis. This model consists of a set of pre-trained reference models $f_{\theta Sv}(D_{Sv})$ corresponding to the existing basic characteristics of the object (v=1, 2, ..., b, where b is the number of basic characteristics of the object, $b \le p$).

The developed method involves the construction of a rough model $f_{\theta S}(D_S)$, and subsequently, an exact model $f_{\theta Tk}(\theta_S,\ D_{Tk})$, in the form of NN with the same structure (dimension of the parameter vector θ_S), which have the reference models $f_{\theta Sv}(D_{Sv})$:

$$dim(\theta_S) = dim(\theta_{Sv}) = dim(\theta_{\theta Tk}), \quad (9)$$

which ensures the simplicity of an accurate model.

Due to this condition, the definition of a rough model is reduced to simple operations on the parameter vectors of the reference models θ_{Sv}. Thus, to determine the parameter vector θ_S of a rough model, an operation can be introduced based on the calculation of the arithmetic mean of the corresponding components of the parameter vectors of the reference models θ_{Sv}:

$$\theta_S^i = \frac{1}{b}\sum_{v=1}^{b}\theta_{Sv}^i, \quad (10)$$

where i is the index of the corresponding elements of the parameter vectors of the coarse θ_S and the reference θ_{Sv} models.

Another way to determine the parameter vector θ_S of a rough model is to find the maximum value among the corresponding parameter vector components of the reference models θ_{Sv}:

$$\theta_S^i = max(\theta_{Sv}^i), v = \overline{1,b}. \quad (11)$$

Thus, another advantage of forming a rough model using the reference model method is the absence of a training procedure, which significantly speeds up the process of building a rough model.

5. Training a target neural network model based on a rough model

After constructing a rough model $f_{\theta S}(D_S)$, an accurate NN model is trained. For the exact model, the same structure is chosen as for the rough model (9). The initial vector of the parameters of the tone model θ_{Tk} is the vector θ_S, obtained from one of the expressions (10), (11). Further, the exact model in the form of NN is trained by the method of backpropagation of the error with updating the network parameters by the method Levenberg-Marquardt to ensure high accuracy of modeling and rapid convergence of the learning process. The learning process continues until the selected loss function (4)–(6) reaches a minimum or a stop condition is met (e.g., achieving a given precision or a maximum number of epochs).

This approach allows you to retain the knowledge gained on the common D_S dataset and effectively adapt the $f_{\theta Tk}(\theta_S, D_{Tk})$ model to the target task.

Approbation of the developed method of synthesis of intelligent systems for identification of nonlinear dynamic objects on the basis of reference models is carried out on the task of modeling a test object with continuous characteristics.

EXPERIMENT SETUP

Simulation model of the test object

The study of the accuracy of the reference model method is carried out on the example of a test object. A simulation model of a test object in the form of a sequence of a nonlinear link with saturation and a dynamic link of the first order is shown in Fig. 3. Transient characteristics of the test object, demonstrating its nonlinear and dynamic characteristics, are shown in Fig. 4.

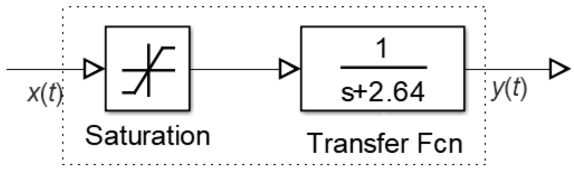


Fig. 3. **Simulation model of the test object**
***Source:* compiled by the authors**

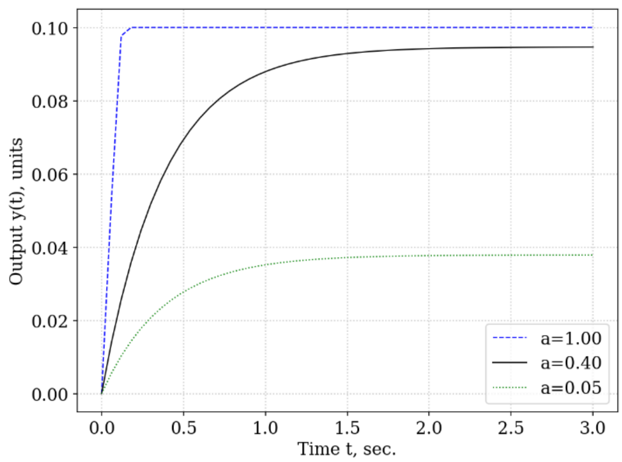


Fig. 4. **Transient characteristics of the test object**
***Source:* compiled by the authors**

As typical properties of the subject area for the formation of the description of the test object, a nonlinear characteristic in the form of saturation and a dynamic link of the first order were chosen.

A labeled dataset $D_S=\{(\mathbf{x}_i^S, y_i^S)\}$ was generated for the test object based on the $x(t)$ signals at the object's input and $y(t)$ responses at its output. Pulsed $x(t)=a\delta(t)$, stepped $x(t)=a\Theta(t)$, linear $x(t)=at$, and harmonic $x(t)=a\ sin(t)$ signals of different amplitude $a \in (0, 1)$.

ISSN 2663-0176 (Print)
ISSN 2663-7731 (Online)

On the basis of the D_S dataset, two reference datasets were formed:

– input step signals $x(t)=a\Theta(t)$ and *responses* $y(t)$ to an object with nonlinearity in the form of saturation $D_{T1}=\{(\mathbf{x}_j^{T1}, y_j^{T1})\}$,

– input step signals $x(t)=a\Theta(t)$ and *responses* $y(t)$ to the object in the form of a first-order dynamic link $D_{T2}=\{(\mathbf{x}_j^{T2}, y_j^{T2})\}$.

The experiment consists in studying the learning speed of an exact model of a test object, built by various methods:

– training based on a pre-trained model on a common D_S dataset;

– training on the basis of individual reference models previously trained on D_{T1} and D_{T2} datasets;

– training by the method of reference models.

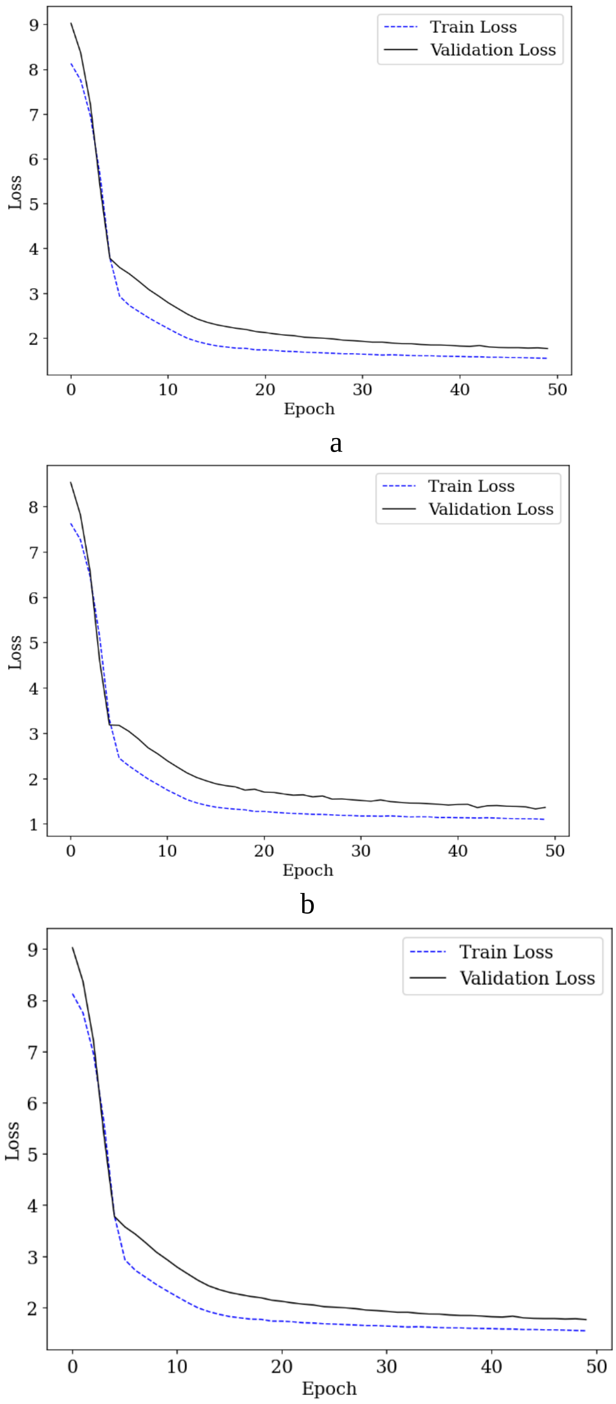
Building a Rough Model

To determine the structure of *the modedi* $f_{\theta S}(D_S)$, which is a three-layer TDNN, according to the results of additional studies, the number of neurons $M = 30$ in the input layer and $K = 30$ in the hidden layer was taken. The D_S input for NN training is converted according to expression (8). The model is trained by the backpropagation method of error propagation with updating network parameters by the Levenberg-Marquardt method. Prior training is limited 50 epochs to prevent overlearning and preserve adaptability.

Fig. 5 shows the dependencies of loss functions (*mse, mae, Huber Loss*) on the number of epochs of learning.

Structure of the exact model $f_{\theta T}(\theta_S, D_T)$ based on the previously trained rough model $f_{\theta S}(D_S)$, models $f_{\theta Tk}(D_{Tk})$ based on separate reference models of *datasets* D_{T1} and D_{T2} and model $f_{\theta T}(D_T)$ on the basis of the superposition of the reference models is chosen to be identical to the rough model $f_{\theta S}(D_S)$ in the form of a three-layer TDNN. The input data for NN training is transformed according to expression (8). The NN is trained by the backpropagation method of error with updating network parameters by the Leeuwenberg-Marquardt method. Training of an accurate model is carried out over a period of 50 epochs.

Fig. 6 shows the dependencies of loss functions (*mse, mae, Huber Loss*) on the number of epochs of learning for exact models based on a rough model, based on individual reference models $f_{\theta T1}(D_{T1})$ and $f_{\theta T2}(D_{T2})$, based on the superposition of reference models.



c

Fig. 5. **Dependencies of loss functions during training of a rough model on the number of epochs of learning:**
a – *mse;* b – *mae;* c – *Huber Loss*
***Source:* compiled by the authors**

Building an accurate model.

Fig. 6 shows the advantage of using pre-trained reference NN in the identification of nonlinear dynamic objects, namely, a significant reduction in the training time of the TDNN model (by 4.6 times) compared to the traditional approach of constructing an accurate model based on a pre-trained rough model with comparable accuracy of both models (*mse*=4.6; *mae*=2.5; *h*=2.5). The use of individual reference models as rough models can also provide a reduction in the training time of an accurate model (by a factor of 1.8) with comparable accuracy of both models (*mse*=4.6; *mae*=2.5; *h*=2.5).

DISCUSSION OF THE RESULTS

The obtained modeling results indicate that the use of TDNN models to identify nonlinear dynamic objects with continuous characteristics allows obtaining high-accuracy models that depict both nonlinear and dynamic characteristics of the objects under study. Such models demonstrate a high ability to adapt to different target tasks in the subject area. Building TDNN on the basis of reference models can significantly reduce the training time of NN without losing the accuracy of modeling.

According to the results of the experiment, the fundamental possibility and efficiency of replacing the process of preliminary training of a rough model with a faster process of obtaining a rough model in the form of a composition of reference models reflecting individual properties of the subject area inherent in the target task are proved.

The advantages of the proposed approach to identifying models of nonlinear dynamic objects on the basis of reference models in the form of TDNN are the ability to quickly adapt to changing operating conditions, high speed of building an accurate model while ensuring a given modeling accuracy. In addition, the developed method allows improving the efficiency of model training in the absence of labeled data for the target task.

The disadvantages of the proposed approach, inherited from methods based on previous learning, are the dependence of modeling results on the quantity and quality of dataset data. Firstly, the common and target datasets should not differ significantly in the distribution of parameters in order to prevent the Domain Shift problem. Secondly, insufficient data for fine-tuning can lead to the problem of overfitting or insufficient training of the model.

The practical limitations of the application of the proposed approach are the a priori need for reference models built on a sufficient amount of qualitative data. Insufficient data or poor data quality can significantly reduce the accuracy of the reference points and, as a result, significantly reduce the training time of the model.

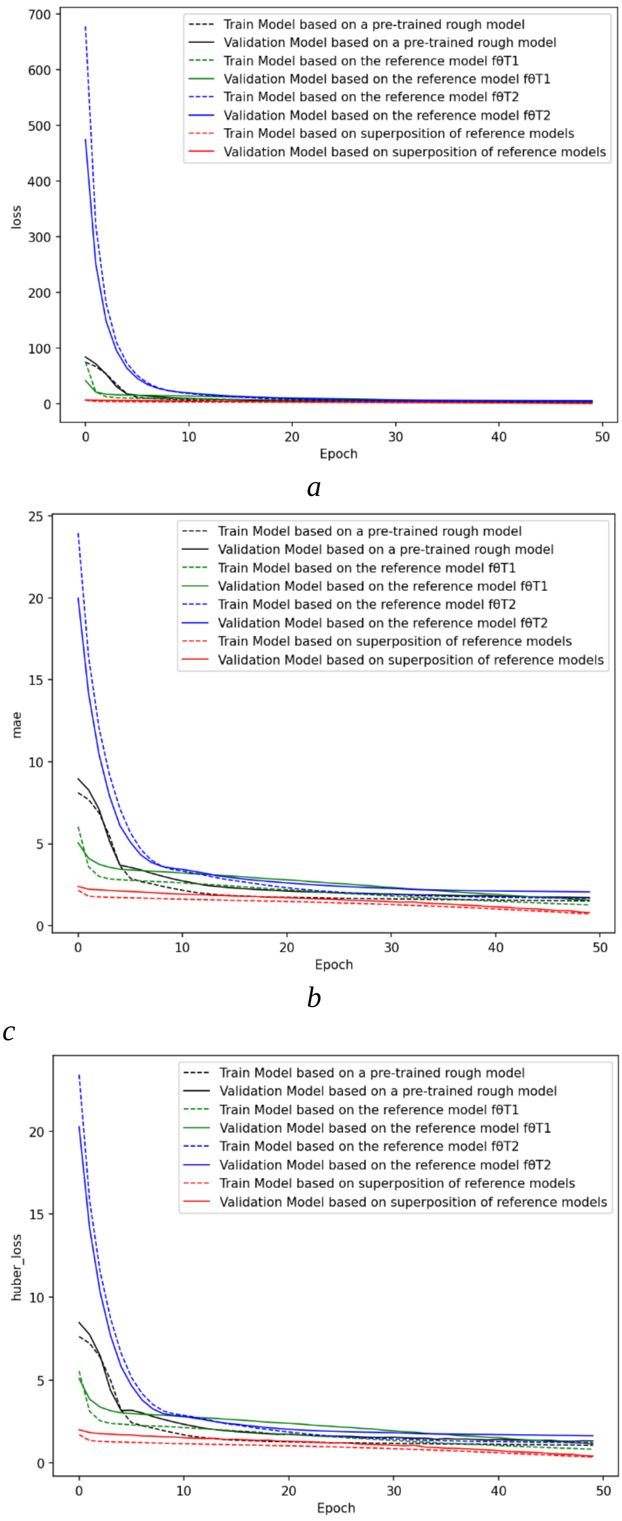


Fig. 6. **Dependencies of loss functions during training of exact models based on a rough model, based on individual reference models, based on the superposition of reference models on the number of epochs of learning:**
a – *mse;* b – *mae;* c – *Huber Loss*
***Source:* compiled by the authors**

Thus, the area of effective application of the proposed method has been allocated: lack of labeled data of the target task in the presence of a common dataset of sufficient size; absence of significant discrepancies between the characteristics of the general and target dataset.

CONCLUSIONS

The paper successfully solves the problem of reducing the time of building models of nonlinear dynamic objects with continuous characteristics while ensuring the specified accuracy of simulations. To resolve the contradiction between the accuracy of modeling nonlinear dynamics and the speed of model construction, the method of synthesis of intelligent systems based on the superposition of reference models in the form of NN reflecting the basic properties of the object was further developed.

The efficiency of the developed method of identification of nonlinear dynamic objects is proved in solving the problem of identification of a test nonlinear dynamic object. The experiment demonstrates a 4.6-fold reduction in the time it takes to build an accurate TDNN model using reference models compared to the full training procedure with comparable accuracy of both models.

Integration of the method of reference models for the identification of nonlinear dynamic objects into existing systems to increase the speed of modeling makes sense when solving a large number of target problems in a sufficiently large subject area, where a one-time set of reference models can be used repeatedly to model a set of different objects.

REFERENCES

1. Wirtz, B. W., Weyerer, J. C. & Geyer, C. "Artificial intelligence and the public sector – applications and challenges". *International Journal of Public Administration.* 2019; 42 (7): 596–615, https://www.scopus.com/record/display.uri?eid=2-s2.0-85050557949&origin=resultslist. DOI: https://doi.org/10.1080/01900692.2018.1498103.

2. Hamm, P. & Klesel, M. "Success factors for the adoption of artificial intelligence in organizations: A literature review". *AMCIS Proceedings*. 2021.

3. Chisty, N. M. A. & Adusumalli, H. P. "Applications of artificial intelligence in quality assurance and assurance of productivity". *ABC Journal of Advanced Research.* 2022; 11 (1): 23–32. DOI: https://doi.org/10.18034/abcjar.v11i1.625.

4. Sen, J. "Machine learning – algorithms, models and applications". *Artificial Intelligence*. *IntechOpen.* London, United Kingdom. 2021. DOI: https://doi.org/10.5772/intechopen.94615.

5. Mitrea, C. A., Lee, C. K. M. & Wu, Z. "A comparison between neural networks and traditional forecasting methods: case study". *International Journal of Engineering Business Management*, 2009; 1 (2): 19–24, https://www.scopus.com/record/display.uri?eid=2-s2.0-84997830620&origin=resultslist. DOI: https://doi.org/10.5772/6777.

6. Kutyniok, G. "The mathematics of artificial intelligence". *Proc. Int. Cong.* Math. 2022; 7: 5118–5139. DOI: https://doi.org/10.4171/ICM2022/141.

7. Zhou, H. "Intelligent personalized content recommendations based on neural networks". *International Journal of Intelligent Networks*. 2023; 4: 231–239. DOI: https://doi.org/10.1016/j.ijin.2023.09.001.

8. Pavlenko, V. D., Shamanina, T. V. & Chori, V. V. "Estimation psychophysiological state via nonlinear dynamic integral models". *Applied Aspects of Information Technology*. 2023; 6 (2): 117–129. DOI:https://doi.org/10.15276/aait.06.2023.8.

9. Kariri, E., Louati, H., Louati, A. & Masmoudi, F. "Exploring the advancements and future research directions of artificial neural networks: A text mining approach". *Applied Sciences*. 2023; 13: 3186. DOI: https://doi.org/10.3390/app13053186.

10. Islam, M., Chen, G. & Jin, S. "An overview of neural network". *American Journal of Neural Networks and Applications*. 2019; 5 (1): 7–11. DOI: https://doi.org/10.11648/j.ajnna.20190501.12.

11. Elsken, T., Metzen, J. H. & Hutter, F. "Neural architecture search: A survey". *In: Hutter, F., Kotthoff, L., Vanschoren, J. (eds) Journal of Machine Learning Research*. Springer. Cham. 2019; 20, https://www.scopus.com/record/display.uri?eid=2-s2.0-85068557085&origin=resultslist. DOI: https://doi.org/10.1007/978-3-030-05318-5_3.

12. Karampiperis, P., Manouselis, N. & Trafalis, T. B. “Architecture selection for neural networks”. *Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02*. Honolulu, HI, USA, 2002; 2: 1115–1119. DOI: https://doi.org/10.1109/IJCNN.2002.1007650.

13. Kästner, C. & Kang, E. “Teaching software engineering for AI-enabled systems”. *The 42nd International Conference on Software Engineering (ICSE). Software Engineering Education and Training*. 2020. – Available from: URL: https://arxiv.org/abs/2001.06691. – [Accessed: Dec. 2023].

14. Hosna, A., Merry, E., Gyalmo, J. et al. “Transfer learning: a friendly introduction”. *J Big Data*. 2022; 9, (1): 102, https://www.scopus.com/record/display.uri?eid=2-s2.0-85140596814&origin=resultslist. DOI: https://doi.org/10.1186/s40537-022-00652-w.

15. Kamath, U, Liu J. & Whitaker, J. “Transfer learning: Domain adaptation”. *In: Deep Learning for NLP and Speech Recognition*. *Springer,* 2019. p. 495–535. DOI: DOI:10.1186/s40537-022-00652-w.

16. Hussain, M., Bird, J. J. & Faria, D. R. “A study on CNN transfer learning for image classification”. *In: UK Workshop on Computational Intelligence*. *Springer.* 2018. p. 191–202, https://www.scopus.com/record/display.uri?eid=2-s2.0-85052244260&origin=resultslist. DOI: https://doi.org/10.1007/978-3-319-97982-3_16.

17. The. V. T., Tien. N. T. K., Thanh. T. K. “A survey on deep learning based face detection”. Applied Aspects of Information Technology. 2023; 6 (2): 201–212. DOI: https://doi.org/10.15276/aait.06.2023.15.

18. Arsirii, O. O. & Petrosiuk, D. V. “Pseudo-labeling of transfer learning convolutional neural network data for human facial emotion recognition”. *Herald of Advanced Information Technology*. 2023; .6 (3): 203–214. DOI: https://doi.org/10.15276/hait.06.2023.13.

19. Tercan, H. et al. “Transfer-learning: bridging the gap between real and simulation data for machine learning in injection molding”. *Procedia CIRP*. 2018; 72: 185–90. DOI: https://doi.org/10.1016/j.procir.2018.03.087.

20. Lu, Y., Jiang, X., Fang, Y. & Shi. Ch. “Learning to pre-train graph neural networks”. *Proceedings of the AAAI Conference on Artificial Intelligence*. 2021; 33 (5): 4276–4284. DOI: https://doi.org/10.1609/aaai.v35i5.16552.

21. Fomin, O. O. & Orlov, A. A. “Modeling nonlinear dynamic objects using pre-trained time delay neural networks”. *Applied Aspects of Information Technology*. 2024; 7 (1): 24–33. DOI: https://doi.org/10.15276/aait.07.2024.2.

22. Hasin, R et al. “Evaluation of GPT and BERT-based models on identifying proteinprotein interactions in biomedical text.” *ArXiv*. 2023. DOI: https://doi.org/10.48550/arXiv.2303.17728.

23. Zhou, Y., Chang, H., Lu, Y., Lu, X. & Zhou, R. “Improving the Performance of VGG through different granularity feature combinations”. *In IEEE Access*. 2021; 9: 26208–26220. DOI: https://doi.org/10.1109/ACCESS.2020.3031908.

24. Pavlenko, V., Shamanina, T. & Chori, V. “Application of nonlinear dynamic models of the oculo-motor system in diagnostic studies in neurosciences”. *Journal of Neuroscience and Neurological Disorders*. 2023; 7 (2); 126–133.

25. Tian, J. “Quality-evaluation models and measurements”. *IEEE Software*. 2004; 21 (3): 84–91. DOI: https://doi.org/10.1109/MS.2004.1293078.

26. Trzęsiok, M. “Measuring the quality of multivariate statistical models”. *Acta Universitatis Lodziensis. Folia Oeconomica*. 2019; 6 (339): 99–100. DOI: https://doi.org/10.18778/0208-6018.339.06.

27. Tong, H. “Functional linear regression with Huber loss”. *Journal of Complexity*. 2023; 74: 101696. DOI: https://doi.org/10.1016/j.jco.2022.101696.

28. Wenyuan, L., Zhu, L., Feng, F. at al “A time delay neural network based technique for nonlinear microwave device modelling”. *Micromachines*. Basel, 2020; 11 (9): 1–14.

29. Fomin, O. et al. “Interpretation of dynamic models based on neural networks in the form of integral-power series”. In: Arsenyeva, O., Romanova, T., Sukhonos, M., Tsegelnyk, Y. (eds). *Smart Technologies in Urban Engineering. Lecture Notes in Networks and Systems*. *Springer*. Cham. 2022; 536: 258–265, https://www.scopus.com/record/display.uri?eid=2-s2.0-85050654902&origin=resultslist. DOI: https://doi.org/10.1007/978-3-031-20141-7_24.

Conflicts of Interest: the authors declare no conflict of interest

Received 19.07.2024
Received after revision 13.09.2024
Accepted 19.09.2024

DOI: https://doi.org/10.15276/hait.07.2024.18
УДК 004.942

Метод опорних моделей синтезу інтелектуальних систем ідентифікації нелінійних динамічних об'єктів

Фомін Олександр Олексійович[1)]
ORCID: https://orcid.org/0000-0002-8816-0652; fomin@op.edu.ua. Scopus Author ID: 57103429400
Сперанський Віктор Олександрович[1)]
ORCID: https://orcid.org/my-orcid?orcid=0000-0002-8042-1790; speranskyy@op.edu.ua.Scopus AuthorID: 54401618900
Орлов Андрій Андрійович[1)]
ORCID: https://orcid.org/0000-0002-3256-5044; 9901020@stud.op.edu.ua. Scopus Author ID: 57202287340
Татарин Олексій Васильович[1)]
ORCID: https://orcid.org/0009-0005-3888-6569; otataryn@stud.op.edu.ua
Кущевий Денис Вікторович[2)]
ORCID: https://orcid.org/0009-0008-2400-9989; biostar1992@gmail.com
[1)] Національний університет «Одеська політехніка», проспект Шевченка, 1. Одеса, Україна
[2)] Luxoft Україна, вул. Картамишевська, 9,в. Одеса, 65000, Україна

АНОТАЦІЯ

Робота присвячена вирішенню протиріччя між точністю моделювання нелінійної динаміки і швидкістю побудови моделей в умовах обмежених обчислювальних ресурсів. Метою роботи є скорочення часу побудови моделей нелінійних динамічних об'єктів із безперервними характеристиками при забезпеченні заданої точності моделювання. Ця мета досягається шляхом шляхом подальшого розвитку метода синтеза інтелектуальних систем на основі суперпозиції попередньо навчених опорних моделей у вигляді нейронних мереж, що відбивають базові властивості об'єкта. Наукова новизна роботи полягає у розвитку методу ідентифікації нелінійних динамічних об'єктів у вигляді нейронних мереж з часовими затримками на основі набору попередньо навчених нейромережевих моделей, що відбивають базові властивості предметної області. На відміну від традиційного підходу до попереднього навчання, розроблений метод дозволяє будувати моделі меншої складності та з меншим часом навчання при забезпеченны необхыдноъ точност. Для визначення початкових параметрів моделі запропоновано вирази на основі суперпозиції опорних моделей у вигляді нейронних мереж. Практична користь роботи полягає в розробці алгоритму методу опорних моделей для навчання нейронних мереж із часовими затримками в задачах ідентифікації нелінійних динамічних об'єктів з безперервними характеристиками, що дозволяє суттєво скоротити час навчання нейронних мереж без втрати точності моделі. Цінність проведеного дослідження полягає у визначенні області ефективного використання запропонованого методу, а саме в наявності достатнього об'єму якісних даних для побудови опорних моделей. Недостатня кількість даних або низька якість даних може суттєво знизити точність опорних моделей і як наслідок, суттєво знизити час навчання цільової моделі.

Ключові слова: нелінійна динаміка; ідентифікація; нейронні мережі з часовими затримками; попереднє навчання

ABOUT THE AUTHORS

Oleksandr O. Fomin - Doctor of Engineering Sciences, Professor. Department of Computerized Systems and Software Technologies. Odesa Polytechnic National University, 1, Shevchenko Ave. Odess, 65044, Ukraine
ORCID: https://orcid.org/0000-0002-8816-0652; fomin@op.edu.ua. Scopus Author ID: 57103429400
Research field: Math modeling; information systems;intelligent control systems

Фомін Олександр Олексійович - доктор технічних наук, професор кафедри Комп'ютеризованих систем та програмних технологій. Національний університет «Одеська політехніка», проспект Шевченка, 1, Одеса, Україна

Viktor O. Speranskyy - Candidate of Engineering Sciences, Associate Professor. Department of Computerized Systems and Software Technologies. Odesa Polytechnic National University, 1, Shevchenko Avenue. Odesa, 65044, Ukraine
ORCID: https://orcid.org/ 0000-0002-8042-1790; speranskyy@op.edu.ua. Scopus Author ID: 54401618900
Research field: Tonlinear systems; modeling; identification; software development; neural networks

Сперанський Віктор Олександрович - кандидат технічних наук, доцент кафедри Комп'ютеризованих систем та програмних технологій. Національний університет «Одеська політехніка», проспект Шевченка, 1, Одеса, Україна

Andrii A. Orlov - graduate student, Department of Computerized Systems and Software Technologies, Odesa Polytechnic National University, 1, Shevchenko Ave, Odesa, Odessa, Ukraine
ORCID: https://orcid.org/0000-0002-3256-5044; 9901020@stud.op.edu.ua. Scopus Author ID: 57202287340
Research field: Information technologies; artificial neural networks

Орлов Андрій Андрійович - аспірант кафедри Комп'ютеризованих систем та програмних технологій, Національний університет «Одеська політехніка», проспект Шевченка, 1, Одеса, Україна

Oleksiy V. Tataryn - graduate student, Department of Computerized Systems and Software Technologies. Odesa Polytechnic National University, 1, Shevchenko Ave. Odesa. Odessa, Ukraine
ORCID: https://orcid.org/0000-0002-3256-5044; otataryn@stud.op.edu.ua.
Research field: Information technologies; artificial neural networks

Татарин Олексій Васильович - аспірант кафедри Комп'ютеризованих систем та програмних технологій. Національний університет «Одеська політехніка», проспект Шевченка, 1, Одеса, Україна

Denys V. Kushchevyi - Senior Java Developer, Luxoft Ukraine, 9v, Kartamyshivska Street. Odessa, 65000, Ukraine
ORCID: https://orcid.org/0009-0008-2400-9989; biostar1992@gmail.com.
Research field: Navigational technologies; geo-informational systems; artificial neural networks

Кущевий Денис Вікторович - провідний інженер, Luxoft Україна, вул. Картамишевська, 9,в. Одеса, 65000, Україна

ISSN 2663-0176 (Print)
ISSN 2663-7731 (Online)