OPEN ACCESS
International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-3, February 2019

Mathematical Simulation of Dynamics on the Basis of Analysis of Multidimensional Time Series with Consideration for Lagged Influence of Factors Using Neural Networks

Oleg Yakovlevich Kravets, Evgeny Efimovich Krasnovskiy, Irina Nikolaevna Kryuchkova, Evgeniya Vitalievna Bolnokina, Vladimir Dmitriyevich Sekerin

***Abstract*: This article investigates into the models and methods of neural simulation of dynamics on the basis of analysis of multidimensional time series with consideration for lagged influence of significant factors. Mathematical formulation of the neural network construction for nonzero lag is presented, peculiarities of lag optimization are described for one independent variable (input), simulation and forecast database is specified as well as neural algorithms of data processing, algorithmization of multivariate regression analysis is carried out with optimization of lag vector for significant factors.**

***Index Terms*: mathematical simulation, neural networks, lag.**

I. INTRODUCTION, FORMULATION OF THE PROBLEM

The problem of forecast of time series in neural logical basis is considered in the following formulation [1–3]. There are M time series X1, X2,…XM, each of N observations. There is a considered variable Y which is similar time series in terms of parameters depending on observations. It is required to construct and to train a neural network which would provide mapping of Y=φ(X1, X2,…XM) with preset accuracy on the array of initial data comprised of observations at times ti, (i=0,1…..N-1). The trained network is used for prediction of Y(t +ΔY) on the basis of X(tN-ΔX).

Training of double-layer neural network with existing time lag is comprised of searching for solution of minimization of target function known as error function in the following form:

$$\min_{\forall(w_i,\Delta_j)} E=\frac{1}{2}\sum_{p=1}^{N}(y_p-f_p(\sum_{i=1}^{K}w_{ki}^{(2)}\cdot(f(\sum_{j=0}^{M}w_{ij}^{(1)}\cdot(x_j(t-\Delta_j))))))^2, \quad (1)$$

where N is the number of training examples, M is the number of inputs, wij(1) is the weight of the i-th neuron of hidden layer, wki(2) is the weight of the i-th neuron of output layer, K is the number of neurons in hidden layer, X0 = 1, w0(1) = 1, w0(2) = 1, Δj is the time lag for the j-th input, xj(t - Δj) is the independent variable biased by time lag.

Manuscript published on 28 February 2019.
* Correspondence Author (s)
Evgeny Efimovich Krasnovskiy*, Bauman Moscow State Technical University, Baumanskaya 2-ya St., 5/1, Moscow, 105005, Russia.
Irina Nikolaevna Kryuchkova, Voronezh State Technical University, 20 years of October St., 84, Voronezh, 394006, Russia.
Evgeniya Vitalievna Bolnokina, Voronezh State Technical University, 20 years of October St., 84, Voronezh, 394006, Russia.
Vladimir Dmitriyevich Sekerin, Kuban State Agrarian University, Kalinina Str., 13, Krasnodar, 350044, Russia
V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Profsoyuznaya St., 65, Moscow, 117997, Russia.

According to the conclusion of the Kolmogorov–Arnold–Hecht-Nielsen theorem [4], for any set of pairs of input–output vectors of arbitrary dimensions there exists homogeneous double-layer neural network with sequential links, with sigmoidal functions of activation [5], and with finite number of neurons; for each input vector this network forms corresponding output vector. Hence, sigmoid function is sued as the function of neuron activation:

$$f(x)=\frac{1}{1+e^{-x}}. \quad (2)$$

Let us define constraints for Δ_j.

1. The time lag Δ_j should be nonnegative, which is defined by physical essence of the problem. Indeed, variation of independent variable at time t + 1 cannot effect dependent variable at time t, hence $\Delta_j \geq 0$.

2. Displacement of observation point towards earlier times should not result in violation of lower boundary by initial time series since the phenomenon described by this time series can be absent as such at that time, and the extrapolated values will have no physical sense. Therefore, $x_j(t - \Delta_j) \in X_j$.

Addition of the variable Δ_j into the target function excludes from training all standard algorithms based on selection of weight coefficients, since existence of a time lag leads to uncertainty in selection of training vectors. Therefore, taking into account the impossibility of simultaneous determination of optimum time lag and network training, it is necessary to consider determination of multidimensional lag as an individual optimization problem.

II. PROPOSED METHODOLOGY.

A. Vector of Optimum Time Lags

Construction of forecasting neural network with accounting for lags is reduced to the problem of construction of vector of optimum time lags and its application for presetting of bias values of independent variables $x_j(t - \Delta_j)$ in Eq. (1). Then the problem presented as Eq. (1) can be solved as a standard problem of training of neural network using a known method, for instance, backpropagation [6] implemented nearly in all softwares of neural simulation. Let the considered previous values of the independent variable X_i are described by the subset of inputs.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

$S_i = \{x_i(t), x_i(t-1), \ldots x_i(t-\Delta_s), \ldots x_i(t-h_i)\}$ of double-layer neural network, multilayer perceptron, of the structure $M^* - K - 1$, where K is the number of neurons in hidden layer, Δ_s is the time lag, h_i is the sampling depth in terms of the i-th factor, M^* is the number of inputs of neural network determined as follows:

$$M^* = \sum_{i=1}^{M}(1+h_i), \qquad (3)$$

where M is the number of considered factors. This equation makes it possible to determine the composition of inputs at arbitrary immersion depth for ach factor. At h = 0 the values of independent variables at previous times are not included into the input vector [2].

Then, the sensitivity of neural network to elimination of input corresponding to the value of independent variable $X_i(t-\Delta_s)$ can be described by absolute value of forecast error as follows:

$$E_{\Delta is} = \left| y_p - f\left(\sum_{r=1}^{K} w_r^{(2)} \left(f\left(\sum_{j=0}^{M^*} w_{rj}^{(1)} \cdot x_j^{is} \cdot \lambda_{sj}^{i}\right)\right)\right) \right|, \qquad (4)$$

where $E_{\Delta is}$ is the forecast error upon elimination of input corresponding to the i-th factor at the time $t - \Delta_s$; $w_r^{(2)}, w_{rj}^{(1)}$ are weights of neurons of trained network; X_j is the value of the j-th input of neural network, y_p is the observed value of forecasted parameter at the time t.

In Eq. (4), λ^i_{sj} is the element of matrix determined as follows:

$$\lambda_{sj}^{i} = \begin{cases} 0,\ j = s + \sum_{k=1}^{i}(1+h_i); \\ 1,\ j \neq s + \sum_{k=1}^{i}(1+h_i). \end{cases} \qquad (5)$$

The elimination matrix λ^i_{sj} is added in order to provide zeroing of input link weights corresponding to the lag Δ_s.

For the set of inputs x_j^{is}, which is a subset of input vector $(x_0, \ldots x_j)$ describing the history of variation of the i-th factor, the following is valid:

$$x^{is} = x^{i}(t-\Delta_s), \qquad (6)$$

where Δ_s is the time lag, $s = 0, \ldots h_i$.

We will consider such value of the time lag Δ_s as optimum when elimination of respective input leads to maximum forecast error [3]:

$$\begin{cases} E_{\Delta is} \xrightarrow[x^{is} \in X_i]{} \max, \\ x^{is} = x^{i}(t-\Delta_s). \end{cases} \qquad (7)$$

The maximum depth of historic sampling in terms of each factor should not exceed half-length of initial time series. The time lag Δ_s should be nonnegative, hence:

$$0 \leq \Delta_s \leq N/2. \qquad (8)$$

In final form the problem of optimization of multidimensional lag with accounting for Eqs. (2), (5), (6), and (8) is as follows:

$$\begin{cases} E_i = \left| y_p - f\left(\sum_{r=1}^{K} w_r^{(2)} \left(f\left(\sum_{j=0}^{M^*} w_{rj}^{(1)} \cdot x_j^{is} \cdot \lambda_{sj}^{i}\right)\right)\right) \right| \xrightarrow[x_j^{is} \in X_i]{} \max, \\ M^* = \sum_{i=1}^{M}(1+h_i), h_i \leq \frac{N}{2}, i = \{1, \ldots M\}, y_p \in Y. \end{cases} \qquad (9)$$

Nonlinear pattern of Eq. (2) in Eq. (9), uncertainty of selection of sampling depth for each factor, dimensions of the problem increasing with the number of considered factors and sampling depths require for development of dedicated neural algorithm accounting for the mentioned constraints.

B. ANALYSIS OF SENSITIVITY

Possibility in principle to apply the analysis of sensitivity in order to determine optimum multidimensional time lag was confirmed in the course of computational experiment.

The sensitivity was analyzed using test data acquired by calculation functions of several variables with known lags for certain variables. In the example below the function is as follows:

$$Y(t) = 10 + \frac{(X_1(t) + \frac{X_2(t)}{3} - \ln(X_4(t)) \cdot X_3(t_3))}{X_6(t) + 2{,}5X_5(t_5)} - 0{,}44X_7(t_7) + X_8, \qquad (10)$$

where $t_3 = t-1$, $t_5 = t-3$, $t_7 = t-2$.

Relative positions of series of initial data $X_1, \ldots X_8$ and depending function Y used in one experimental sequence are illustrated in Fig. 1.

Variations of the parameters in Eq. (9) provided the set of neural networks with the generalized structure illustrated in Fig. 2.

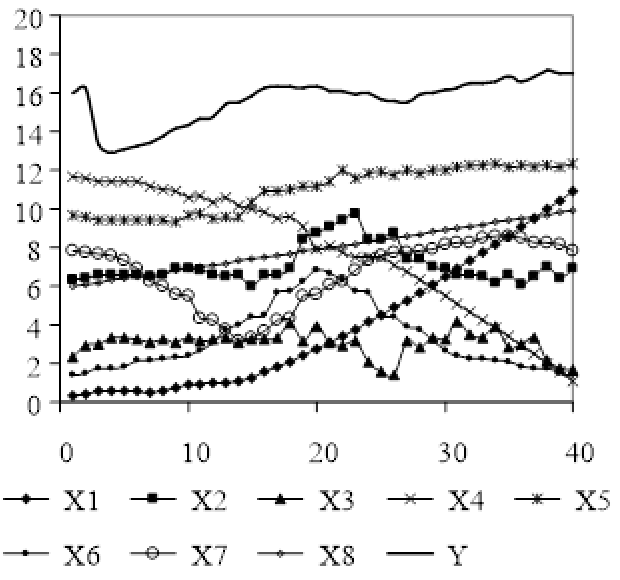

Fig. 1. Eight series of independent variables X1(t),… X8(t) and dependent variable Y.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

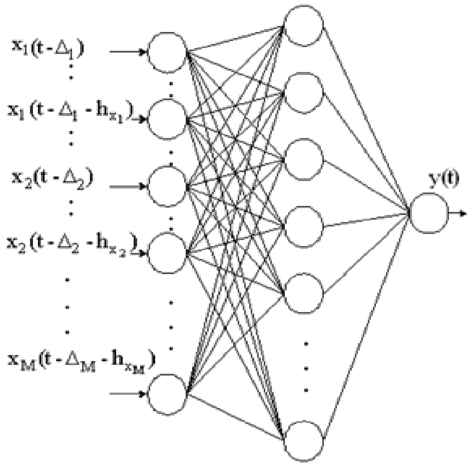

Fig. 2. Generalized structure of neural network after optimization of time lag vector

Table 1 summarizes the errors of training, generalization, and forecasting obtained during computational experiment.

Table 1. Quality of neural network training and forecast Y

Initial data set	S.D. Ratio (for training set)	S.D. Ratio (for reference set)	MAPE, %
$X_1(t)$,… $X_8(t)$	0.240844	0.184005	2.78
$X_1(t)$, $X_2(t)$, $X_3(t-1)$, $X_4(t)$, $X_5(t-3)$, $X_6(t)$, $X_7(t-2)$, $X_8(t)$	0.09965	0.199094	1.37

Some results of sensitivity analysis of neural networks are illustrated in Fig. 3. In the series without initial time lag, for instance, $X_4(t)$, the highest sensitivity was observed for the inputs $X_i(t)$ (Fig. 3, a). For the series $X_3(t)$, $X_5(t)$, $X_7(t)$ the most sensitive were the inputs $X_i(t-\Delta)$ ($\Delta_3=1$, $\Delta_5=3$, $\Delta_7=2$) (Fig. 3, b).Analysis of the errors of training, generalization, and forecasting for initial and bias series shown in Table 1 made it possible to obtain the conclusion of significant decrease in errors of training and forecasting for biased series at nearly constant error of generalization, which evidenced efficiency of the proposed algorithm and absence of structural effects in variation of forecast quality.

C. Algorithm

ALGORITHMIZATION OF HETEROGENEOUS AND REGRESSION ANALYSIS

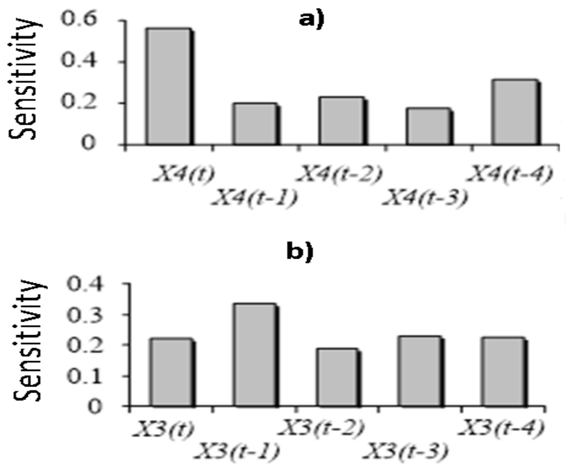

Fig. 3. Selected results of analysis of input sensitivity reflecting variations of independent variables.

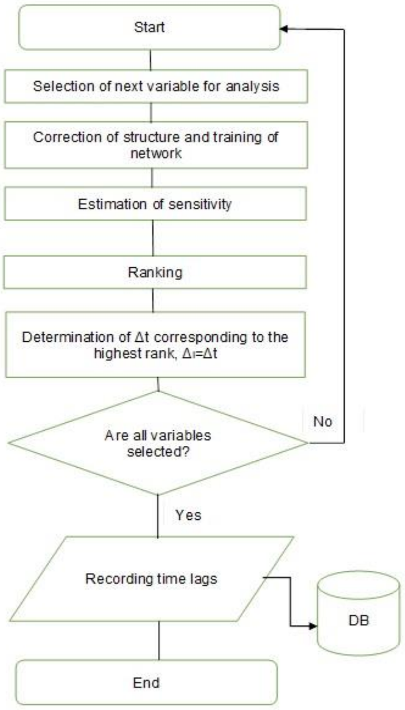

Fig. 4. Algorithm of determination of optimum lag vector.

Then, algorithms of heterogeneous and regression analysis were developed facilitating construction of vector of optimum time lags of significant independent variables. The following subtasks were solved during consideration of the main forecast problem:

- determination of optimum lag vector;
- biasing the initial time series by the determined lags.

In order to determine optimum time lag for each time factor, the procedure was proposed for estimating neural network sensitivity to the values of independent variable in various time points upon increase in immersion depth in terms of the considered factor. The maximum immersion depth is determined as $N_i/2$, where N_i is the length of time series of the i-th independent variable. The algorithm is illustrated in Fig. 4.

D. FLOW CHART

The sensitivity has been analyzed as follows:

Stage 1. Determination of maximum immersion depth.

Stage 2. Construction and training of neural networks with various composition of inputs: immersion depth of the considered series is set to maximum, the other series are presented by one input per series.

Stage 3. Determination of time lag for each series based on the analysis of neural network sensitivity to elimination of input.

After determination of optimum time lags, a new combination of data was obtained used for training of neural network by backpropagation and determination of forecasted value of dependent variable. The obtained values of optimum time lags were used for bias of time series of independent variables and construction of

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

adjusted set of initial data according to the following sequence:

1). determination of maximum optimum lag;
2). biasing each series by optimum lag of certain independent variable;

selection of tactics to continue the process: reduction of series of initial data by maximum lag or complementing missed data.

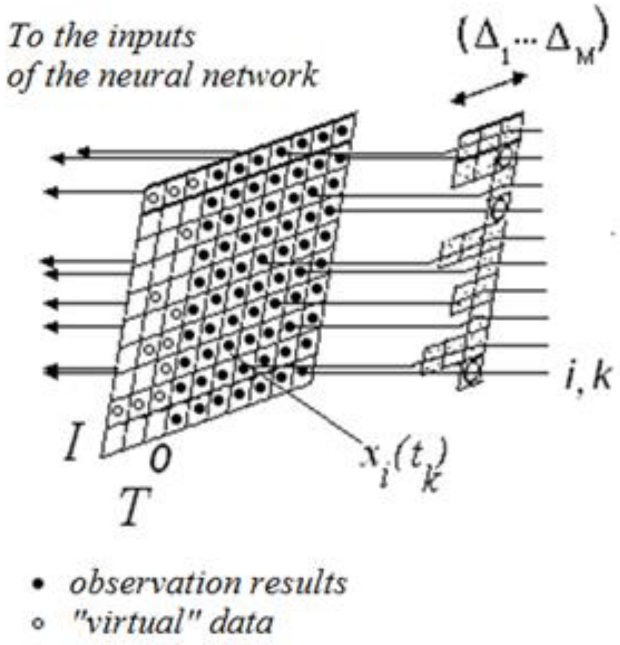

Fig. 5. Operation principle of lag network.

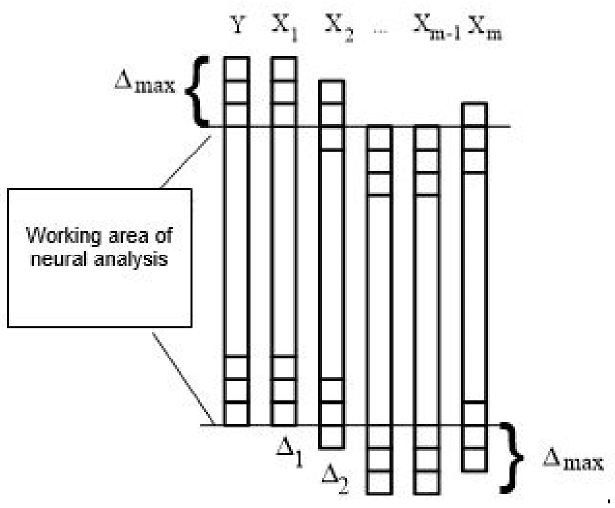

Fig. 6. Generation of reduced set of initial data

The series of initial data are biased by means of delay line, its operation principle is illustrated in Fig. 5, where $x_i(t_k)$ is the value of the i-th independent variable in the k-th time point, $\Delta(\Delta_1 \dots \Delta_M)$ is the vector of optimum time lags.

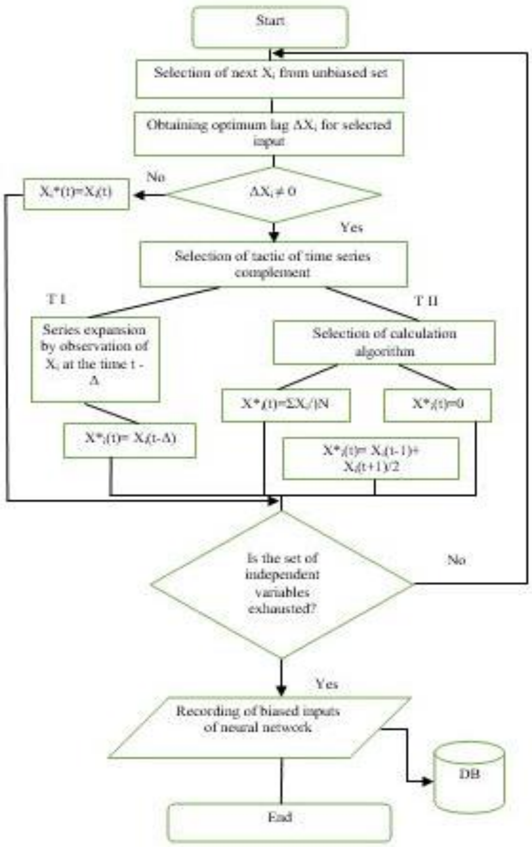

Fig. 7. Expanded algorithm of calculation of biased inputs of neural network.

III. RESULT ANALYSIS

In order to obtain adjusted sampling for training and forecast, the following algorithm is proposed:

Step 1. The first parameter X1 is selected.

Step 2. The corresponding time lag Δ1 is selected.

Step 3. If the time lag Δi is zero, then the values of the i-th independent variable at selected times are transferred to the resulting sampling without variations. Otherwise, the series is biased.

If Xi at the time t-Δ is absent, then the tactics of reduction or complementation of time series are selected, and according to the selected variant either reduced (Fig. 6) or complemented observation series are obtained.

Step 4. Assignment of a new value of Xi in each time point. The missed data were complemented using one of the three tactics of obtaining virtual values.

Tactics I – the time series is complemented by direct observations in this point.

Tactics II – the time series is complemented by calculated data according to the selected algorithm (average for the series, average for adjacent values, zero filling).

Figure 6 illustrates construction of reduced matrix of initial data. This method of process continuation is restricted by necessity to maintain the number of training examples at the level sufficient for training neural network.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Expanded algorithm of construction of adjusted sampling with complemented missed values is illustrated in Fig. 7.

The tactics of maintaining imply the existence of missed data in the initial series which should be filled prior to training and forecasting. While selecting the tactics of complementing time series, it is necessary to account for peculiarities of subject area of the considered problem.

Therefore, the developed optimization algorithm of multidimensional lag based on sensitivity analysis using delay time at neural network input makes it possible to consider for delayed influence of important factors on forecast and to eliminate uncertainty in selection of initial data vector upon neural network training and testing, and for obtaining forecast.

IV. CONCLUSION

1. The authors studied the necessity to consider for lagged influence of independent variables on forecast. It was demonstrated that the existence of unaccounted time lag did not permit to obtain correct training examples and deteriorated forecast quality.
2. Mathematical formulation of the construction of forecasting neural network was obtained for the case of nonzero lag for one and several independent variables.
3. The method of determination of optimum time lag was proposed based on searching for maximum absolute error upon elimination of input corresponding to certain lag value.
4. The proposed algorithm was verified by sensitivity analysis of neural network constructed for approximation of function with predetermined time lags. The results evidence that the error of training on biased time series decreases significantly as a consequence of presetting unique relation of examples ($x_i(t)$, $y(t)$) and construction of correct training set which more completely characterizes real dependences between variables, and forecast error decreases nearly twice.
5. An additional stage of preprocessing of initial data that comprised biasing the series of initial data by optimum time lags, was proposed.
6. Optimization algorithm of time lag vector was developed based on the analysis of neural network sensitivity to elimination of input corresponding to optimum time lag for each independent variable.
7. Procedure of adjustment of neural network structure was developed with decrease in the depth of historical sampling determined by the plots of autocorrelation functions, which is especially important upon a lack of training examples upon studying systems with short observation history. The structure of delay line was proposed permitting to account for the determined time lags for each independent variable and to form correct sets for neural network training and testing, and for obtaining adjusted forecast.

REFERENCES

1. V.M. Avdeeva, O.Ya. Kravets, “Teoreticheskie osnovy prognozirovaniya nalogovykh postuplenii na osnove krosskorrelyatsionnogo analiza mnogomernykh vremennykh ryadov” [Theoretical foundations of forecast of tax revenue on the basis of cross correlation analysis of multidimensional time series], *Sistemy upravleniya i informatsionnye tekhnologii*, 1.2(23), 2006, p. 212-216.
2. V.M. Avdeeva, O.Ya. Kravets, I.N. Kryuchkova, “Territorial'noe prognozirovanie nalogovykh postuplenii s primeneniem mnogomernykh krosskorrelyatsionnykh tekhnologii” [Territorial forecast of tax revenue using multidimensional cross correlation technology], *Innovatsionnyi Vestnik Region*, 3(9), 2007, p. 31-36.
3. V.M. Avdeeva, I.N. Kryuchkova, “Issledovanie tekhnologii neirosetevogo prognozirovaniya nalogovykh postuplenii territorii s primeneniem tekhniki mnogomernogo krosskorrelyatsionnogo analiza” [Studying neural network forecast of territorial tax revenue using multidimensional cross correlation technology], *Territoriya nauki*, 4(5), 2007, p. 428-436.
4. R., Hecht-Nielsen, “Theory of the Backpropagation Neural Network”, *Neural Networks*, 1(1), 1989, p. 593 - 605.
5. M.N. Gibbs, “Variational Gaussian process classifiers” *IEEE Transactions on Neural Networks*, 11(6), 2000, p. 1458–1464.
6. I. Goodfellow, Y. Bengio, A. Courville, “*Deep Learning*”, MIT Press, 2016, p. 196.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)