Proceedings of the 6th WSEAS Int. Conf. on NEURAL NETWORKS, Lisbon, Portugal, June 16-18, 2005 (pp120-125)

Design of the Models of Neural Networks and the Takagi-Sugeno Fuzzy Inference System for Prediction of the Gross Domestic Product Development

VLADIMÍR OLEJ
Institute of System Engineering and Informatics
Faculty of Economics and Administration
University of Pardubice
Studentská 84, 532 10 Pardubice
CZECH REPUBLIC

Abstract: - The paper presents the possibility of the design of frontal neural networks and feed-forward neural networks (without pre-processing of inputs time series) with learning algorithms on the basis genetic and eugenic algorithms and Takagi-Sugeno fuzzy inference system (with pre-processing of inputs time series) in predicting of gross domestic product development by designing a prediction models whose accuracy is superior to the models used in praxis [1,2].

Key-Words: - Gross domestic product, frontal neural networks, feed-forward neural networks, Takagi-Sugeno fuzzy inference systems, genetic and eugenic algorithms, EuSANE algorithm.

1 Introduction

Neural networks (NNs) [3], genetic algorithms (GAs) [4], eugenic algorithms (EuAs) [4,5] and the Takagi-Sugeno fuzzy inference systems (FISs) [4,6] provide in comparison with econometric methods of prediction the advantage of superior processing of non-linear dependencies which may contribute to improvement of prediction accuracy. Nominal gross domestic product (GDP) is one of the most complex indicators of production activity in the economy. Expected and actual development of GDP is among the crucial factors of determining state economic policy, making decisions about actual financial investments of corporations and households, drawing up corporate production budgets etc. Gross domestic product represents the market value of the final goods produced by factors of production on the territory of the given economy state during the selected period.

In the second section of this paper are introduce the index of leading economic indicators (ILEIs) and the diffusive index of leading economic indicators (DILEIs) of the US economy during the years 1965-2004. The third section of this paper contains the basic notions of NNs with implicit presentation of time and basic notions of GAs and EuAs, which are used as learning algorithms for NNs and the basic notions of the Takagi-Sugeno FIS. The fourth section presents design the models of GDP prediction for the unconditional prediction of real GDP development of the USA on the basis NNs and the Takagi-Sugeno FIS. In this section the learning process of the designed models on the basis of frontal NNs and feed-forward NNs is realised by means of GAs and EuAs. A new learning algorithm for prediction model of GDP by means of frontal NNs and feed-forward NNs was proposed on the basis of advantages of GAs and EuAs. The fifth section includes the analysis of results of designed models with new learning algorithm Eugenic Symbiotic Adaptive Neuro-Evolution (EuSANE) and analysis of results of designed the Takagi-Sugeno FIS. A short conclusion of the results obtained in this paper and an overview of future research topics conclude the paper.

2 Modelling of the GDP Development

Even if business cycles are of different intensity, duration and are caused by various factors, there exist certain relations in the movement of economic variables (EVs), which describe business cycles. The term business cycle represents fluctuations in the aggregate economic activity. The aggregate economic activity is measured by a set of indicators, which include (except for GDP) national income, employment and intensity of foreign trade. Therefore, EVs used for the prediction of business cycles can be used for the prediction of GDP development. At the present time there are lots of EVs at disposal, which monitor the development of the economy, e.g. ILEI and DILEI. The DILEI measures the percentage of decreasing and increasing EVs included in ILEI, i.e. the proportion of EVs indicating the reversal of economic growth. Both ILEI and DILEI can be used to model the prediction of GDP development. Both indexes are used in praxis for the prediction of recession. Whereas ILEI represents aggregate development of ten leading anti-cyclical EVs, DILEI measures the percentage of decreasing or increasing EVs, i.e. the proportion of EVs signalling the reversal of economic development. The interconnection of these two indicators provides a reliable signal of recession. It is assumed that recession starts if the average value of ILEI growth falls below 2 [%] and at the same time the value of DILEI drops below the critical value of 0.5. The development of selected indexes of the US economy during the years 1965-2004 is shown in Fig.1 and Fig.2. The existing econometric models of the prediction of GDP development use different formulations of relations among them and various methods of estimating the parameters of specified models. The presented facts result in different degrees of accuracy. The prediction accuracy can be described by means of average error ε, mean average error σ and root mean squared error δ.

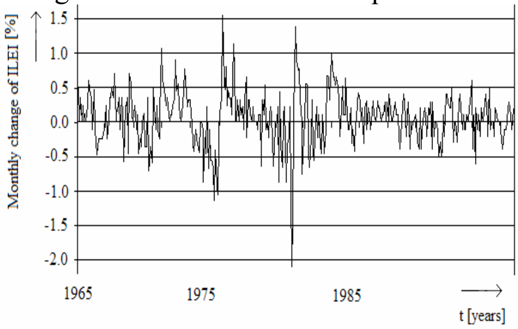

Fig.1 Development of ILEI of the US economy in the years 1965-2004

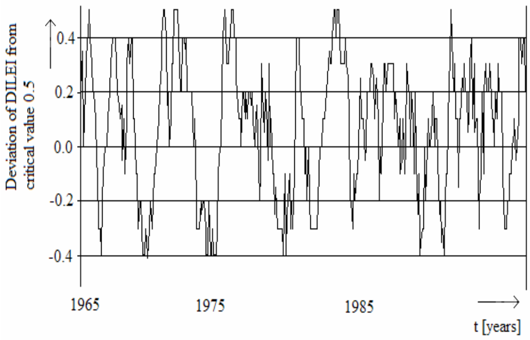

Fig.2 Development of DILEI of the US economy in the years 1965-2004

3 The Basic Notion

The output signal of NN becomes a function of time when making the prediction. The time dimension can be implemented into NNs implicitly and explicitly. Time becomes the proper output signal with regard to the explicit in two ways, namely presentation of time. The implicit presentation of time represents the incorporation of the time dimension into the structure of NN. With regard to the implicit presentation of time, so-called filters, which apply short-term memory and operation of dimensional summation, can be used in feed-forward NNs. Dimensional summation represents the summing-up of input signals. Short-term memory preserves the knowledge about states of NN surrounding, ultimately about individual neurons in the close past. There exist several possible structures of short-term memory from which the most commonly used is memory with a constant number of delays. This type of memory works with discrete time and with the operation of delay z^{-b}, where b is the number of delay time intervals. This operation works in the way that for the value of variable x in time t the value of variable x in time t-b returns. Every memory has its depth measured by parameter b. Deeper memory means that more information about past states of surrounding is preserved.

As far as feed-forward NNs there exists two types of architecture with the implicit presentation of time, i.e. frontal and distributed NNs. In frontal NNs short-term memory is situated in the beginning of NN as a fore-layer processing the time context. Classical back-propagation algorithms may learn such NNs or by another algorithm suitable for feed-forward NNs. Genetic algorithms are used as learning algorithms of NNs [4].

Genetic algorithms differ in the quality of epistasis processing. Epistasis represents the dependency of genotype fitness η on degree of interaction among alleles, eventually groups of alleles of genes. Epistasis occurs when a change of allele of a certain genotype or change of alleles of groups of genes causes an increase in the value of the objective function and at another time a decrease of this value, depending on alleles of other genes of the genotype. In this case it is not suitable to use GA, but the smart combination of alleles and groups of alleles is necessary. Eugenic algorithms enable the smart recombination dependent on the analysis of interaction among alleles and their groups [4,5]. Eugenic algorithms use principles of so-called eugenic evolution. Population develops according to the eugenic principles if genotypes with the smallest fitness η are gradually replaced by new genotypes generated on the basis of the smart recombination of genotypes of the entire population.

The general structure of FIS contains the fuzzification process by means of input membership functions, construction of base rules (BRs) or automatic extraction of rules from the input data, application of operators (AND, OR, NOT) in rules, implication and aggregation within rules and the defuzzification process of obtained outputs to the crisp values. Normalisation of the inputs and their transformation to the range of values of the input membership functions is realised during the fuzzification process. The inference mechanism is based on the operations of fuzzy logic and implication within rules [4,6]. Transformation of the outputs of individual rules to the output fuzzy set is realised on the basis of the aggregation process. Conversion of fuzzy values to expected crisp values is realised during the defuzzification process. There exists no universal method for designing shape, the number and parameters of the input and output membership functions. The input to the fuzzification process is the crisp value given by universe of discourse (reference set). The output of the fuzzification process is the membership function value.

The construction of BRs can be realised by extraction of rules from historical data, if they are at one's disposal. Various optimisation methods for number of rules in BRs are presented in [4,6]. One of the possibilities for optimising BRs from historical data is the so-called Adaptive Neuro-Fuzzy Inference System (ANFIS) method. The essence of this method is the neuro-adaptive learning process, on the basis of which it is possible to derive the parameters of membership functions and to extract BRs. The input data are mapped to the output data by means of this method, whereas parameters of individual membership functions are gradually changed during the learning process, so that relations between the space of the input variables and the output variables are described in the best possible way.

4 The Models of GDP Prediction

The first model of the prediction is presented by NN with the use of the filters applying short-term memory with a constant number of delays. The formulation of frontal NN with the same memory depth of all filters, linear filter and linear output neuron is as follows

$$Y = \sum_{k=1}^{K} \alpha_k d\left(\sum_{j=1}^{J} \beta_{jk} \sum_{i=1}^{b} \chi_{ijk} X_{ijk} \right), \text{ where} \tag{1}$$

Y is the output of NN, α is synapse weight vector among neurons in the hidden layer and output neuron, β is synapse weight vector among filters and neurons in the hidden layers, χ is synapse weight vector inside the filter, k is index of neuron in the hidden layer, K is the number of neurons in the hidden layer, d is activation function, j is index of the filter, J is the number of filters per neuron in the hidden layer, b is short-term memory depth of the filter, X is the input vector of NN. Structures of the filter of frontal NN and frontal NN are shown in Fig.3 and Fig.4.

The number of filters in the input layer of frontal NN is dependent on the number of EVs used in the model and their time delay. The algorithm selects the inputs of frontal NN. It would be necessary to allow the algorithm to choose the length of the input vector. Thus the number of input neurons becomes the parameter of the model. The model uses the linear combination of the inputs, because time series of EVs consist of significant linear elements. Therefore the activation function should be sufficient to cover the non-linearity. Neurons used in the hidden layers consist of frequently applied logistic functions with slope parameter equal to 1. The output neuron is linear.

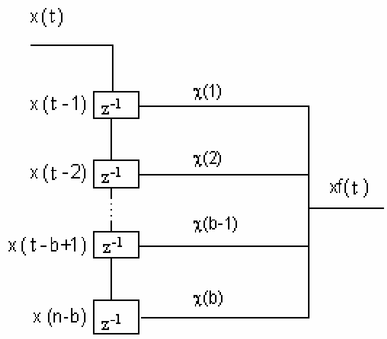

Fig.3 The filter of frontal NN

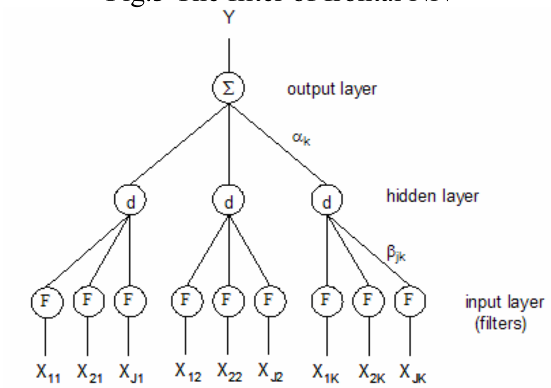

Fig.4 Frontal NN

The learning algorithm looks for such combinations of EVs, time delays and weights which lead to the least prediction error as regards tendency and size of GDP change. Thus the algorithm realises not only the prediction of GDP, but also the selection of suitable EVs, including time delays. The advantage of the proposed solution is the independence on the validity of statistical assumptions. In this case the space of all possible solutions consists of all possible combinations of EVs, their time delay and all synapse weights of NN. Considering that the required degree of accuracy is big, the first found local optimums of objective function will probably not be sufficient. Therefore, it is necessary to design the learning algorithm, which enables to leave local minimums and to continue to search in various segments of the space of all possible solutions. One of the suitable alternatives is the application of a GA. A GA can be considered as a faster algorithm than a EuA, because a GA does not compute epistasis. However, in this case epistasis analysis appears to be useful because of dependency of statistical significance of the additional EV of the model on the EV already included in the model, therefore application of the EuA would be justified too. The EuSANE algorithm, which allows taking advantage of both algorithms, is presented in [4,5]. A new learning algorithm has been proposed for the model of frontal NN on the basis of advantages of the EuSANE algorithm. Except for its presentation of neurons it differs from in the sense that it does not work in two but in three grades. In the first grade the population of P0 filters associated with the input neurons is developed by means of a GA. In the second grade the population of P1 neurons in the hidden layer is also developed by means of a GA. In the third grade the population of P2 hidden layers is developed by means of a EuA.

The second model of the prediction is presented by NN which input vector consists of various values of EVs with various time delays. The formulation of feed-forward NN (Fig.5) is as follows

$$Y = \sum_{k=1}^{K} \alpha_k d\left(\sum_{j=1}^{J} \beta_{jk} X_{jk}\right), \text{ where} \qquad (2)$$

Y is the output of NN, α is synapse weight vector among neurons in the hidden layer and output neuron, β is synapse weight vector among the input neurons and the neurons in the hidden layer, k is index of neuron in the hidden layer, K is the number of neurons in the hidden layer, d is activation function, j is index of the input neuron, J is the number of the input vectors per neuron in the hidden layer, X is the input vector of NN. The number of neurons in the input layer of NN is dependent on the number of EVs used in the model and their time delays. Neurons used in the hidden layers consist of frequently applied logistic functions with slope parameter equal to 1. The output neuron is linear.

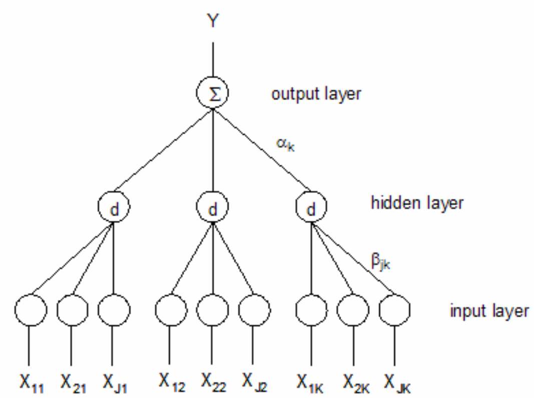

Fig.5 Feed-forward NN

The learning algorithm works in two grades. In the first grade it develops by means of GA the population of P1 potential neurons in the hidden layer of NN. In the second grade it develops by means of EuA the population of P2 hidden layers of NN. By doing this it tries to find such neurons in the hidden layer form P1 which are not only qualitative itself (the quality of neurons is given by their fitness η), but also neurons that cooperate well within hidden layer. The EuSANE algorithm starts to function by random initialisation of both population and finishes its function by finding of the hidden layer with predefined fitnes η. A band-new learning algorithm of NN model which input vectors are created by time delays of EVs was proposed on the basis of advantages of the given EuSANE algorithm. This algorithm is similar to [4,5] algorithm with the difference that neurons in the hidden layer are presented not only by synapse weights, but also by reference to the values of EVs, which are used by the given neuron in the input vector.

The population initialisation of proposed EuSANE algorithms is random. The end criterion of algorithms is the maximal number of the generations. Prediction quality should have two dimensions: deviation size of predicted value of GDP growth from real value and number of errors when predicting the tendency of GDP change. These criteria influence the construction of the objective function h. A suitable form is as follows

$$h = \frac{1}{1 + O a_1} \frac{1}{1 + \delta a_2}, \text{ where} \qquad (3)$$

O is the number of errors when predicting the tendency of GDP change, δ is the root mean squared error of the prediction, a is the parameter of the objective function. This function is maximised by means of the newly-designed EuSANE algorithms to find the suitable economic hypothesis and the suitable parameter values of the prediction model. The aim goal of the algorithm is to find the combination of the independent variables and parameters of the model to obtain higher accuracy of the prediction in comparison with the Federal Reserve Bank of San Francisco (FRBSF) model [1].

The third model of the prediction, the Takagi-Sugeno FIS to predict of the GDP development can be realised by data pre-processing. The ILEI and DILEI can be pre-processed by means of indicators of technical analysis. The calculated indicators can be used as the inputs to the Takagi-Sugeno FIS. The output of FIS is predicted quarterly growth of real GDP development (PGDP, i.e. Δt = 3 months). Thus indicators of technical analysis in time t (the inputs to FIS) are assigned to quarterly growth of real GDP development in time $t + \Delta t$ (the output of FIS). Fuzzy inference system appears to be the most suitable with the following input variables: Running Median (RM), Exponentially Equalization (EE), Weighted Moving Average (WMA), Center Moving Average (CMA), Moving Average (MA11), 11-monthly MA, Moving Average (MA5), 5-monthly MA. The data pre-processing for example for EE_{ILEI} is presented in Fig.6. The BRs consists of rules. The rules are used for the construction of conditional terms, which create the basis of FIS.

Let $x_1, x_2, \ldots, x_i, \ldots, x_n$ be the input variables defined in the reference sets $X_1, X_2, \ldots, X_i, \ldots, X_n$ and let y be the output variable defined in the reference set Y. Then FIS has n input variables and one output variable. Every set X_i, i = 1, ... ,n, can be divided into p_j, j = 1, ... ,m, the fuzzy sets $\mu_1^{(i)}(x)$, $\mu_2^{(i)}(x)$, ... ,$\mu_{pj}^{(i)}(x)$, ... ,$\mu_m^{(i)}(x)$. The individual fuzzy sets $\mu_1^{(i)}(x)$, $\mu_2^{(i)}(x)$, ... ,$\mu_{pj}^{(i)}(x)$, ...

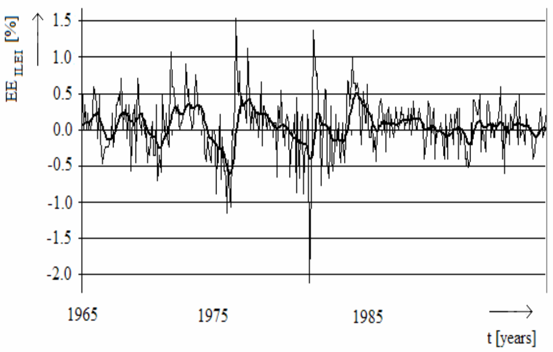

Fig.6 Development of ILEI of the US economy in the years 1965-2004 by EE_{ILE} pre-processing

,$\mu_m^{(i)}(x)$, i = 1, ... ,n; j = 1, ... ,m represent assignment of linguistic variables relating to sets X_i. The set Y is also divided into p_k, k = 1, ... ,o the fuzzy sets $\mu_1(y)$, $\mu_2(y)$, ... ,$\mu_{pk}(y)$, ... ,$\mu_o(y)$.The fuzzy sets $\mu_1(y)$, $\mu_2(y)$, ... ,$\mu_{pk}(y)$, ... ,$\mu_o(y)$ represent assignment of linguistic variables for the set Y. Then rule of the Mamdani FIS can be written as follows

IF x_1 is $A_1^{(i)}$ AND x_2 is $A_2^{(i)}$ AND ... AND x_n is $A_{pj}^{(i)}$ THEN y is B, (4)

where $A_1^{(i)}$, $A_2^{(i)}$, ... ,$A_{pj}^{(i)}$ represent linguistic variables corresponding to the fuzzy sets $\mu_1^{(i)}(x)$, $\mu_2^{(i)}(x)$, ... ,$\mu_{pj}^{(i)}(x)$, ..., $\mu_m^{(i)}(x)$, i = 1, ... ,n; j = 1, ... ,m and B represents the linguistic variable corresponding to the fuzzy sets $\mu_1(y)$, $\mu_2(y)$, ... ,$\mu_{pk}(y)$, ... ,$\mu_o(y)$, k = 1, ... ,o.

The Takagi-Sugeno FIS can be acquired by modification of the Mamdani FIS. These two types of FISs differ in the specification of the output membership functions. These membership functions are constant, linear or polynomial in the case of the Takagi-Sugeno FIS. The division of sets X_i, i = 1, ... ,n, into the fuzzy sets $\mu_1^{(i)}(x)$, $\mu_2^{(i)}(x)$, ... ,$\mu_{pj}^{(i)}(x)$, ... ,$\mu_m^{(i)}(x)$, i = 1, ... ,n; j = 1, ... ,m is the same in both types of FISs. Then rule of the Takagi-Sugeno FIS can be written in the following way

IF x_1 is $A_1^{(i)}$ AND x_2 is $A_2^{(i)}$ AND ... AND x_n is $A_{pj}^{(i)}$ THEN y = h, (5)

where h is the constant. In this case the output membership functions are singletons. A FIS containing rules defined by relation (5) is known as the Takagi-Sugeno FIS a zero-order. If the output membership functions are linear then rules of the Takagi-Sugeno FIS are in the following form

IF x_1 is $A_1^{(i)}$ AND x_2 is $A_2^{(i)}$ AND ... AND x_n is $A_{pj}^{(i)}$ THEN y = f(x_1, ... ,x_n), (6)

where f(x_1, ... ,x_n) is the linear function. A FIS containing rules defined by relation (6) is known as the Takagi-Sugeno FIS a first-order. In the case that f(x_1, ... ,x_n) is a polynomial function, it is the Takagi-Sugeno FIS a second-order.

A FIS is shown in Fig.7. It is the Takagi-Sugeno FIS with 12 input variables, 3 rules (which are the result of BRs extraction from the historical data within the optimisation process of FIS) and 1 output variable.

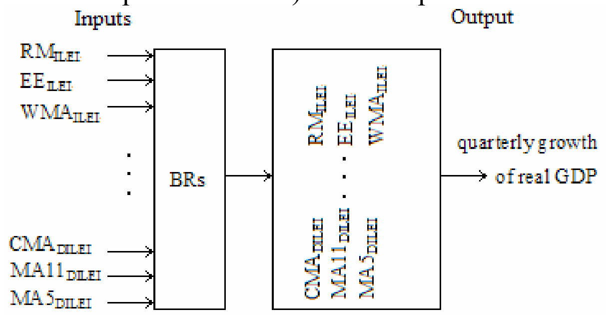

Fig.7 The Takagi-Sugeno FIS

The input variables for ILEI (RM_{ILEI}, EE_{ILEI}, WMA_{ILEI}, CMA_{ILEI}, $MA11_{ILEI}$, $MA5_{ILEI}$) and for DILEI (RM_{DILEI}, EE_{DILEI}, WMA_{DILEI}, CMA_{DILEI}, $MA11_{DILEI}$, $MA5_{DILEI}$) in time t is represented by means of 3 membership functions. They are bell membership functions. The individual membership functions are described by means of the linguistic variables, for example: decrease_EE_{ILEI}, zero_EE_{ILEI}, increase_EE_{ILEI}.

The other input variables of FIS, i.e. (RM_{ILEI}, WMA_{ILEI}, CMA_{ILEI}, $MA11_{ILEI}$, $MA5_{ILEI}$) and (RM_{DILEI}, EE_{DILEI}, WMA_{DILEI}, CMA_{DILEI}, $MA11_{DILEI}$, $MA5_{DILEI}$) are similarly represented. The output variable (PGDP in time t + Δt = 3 months) can be described by means of 3 membership functions. These membership functions are linear, because the FIS is the Takagi-Sugeno FIS. The coefficients of the output membership functions of the designed FIS are optimised by the ANFIS method. The BRs of the designed FIS consists of 3 rules extracted from the historical data. The individual rules are weighted by value 1 which means that all rules have the same influence on the output variable.

5 Analysis of the Results

The designed the frontal NN, the feed-forward NN and the Takagi-Sugeno FIS can be tested on the historical data. The development of GDP and PGDP in time t + Δt is shown in Fig.8.

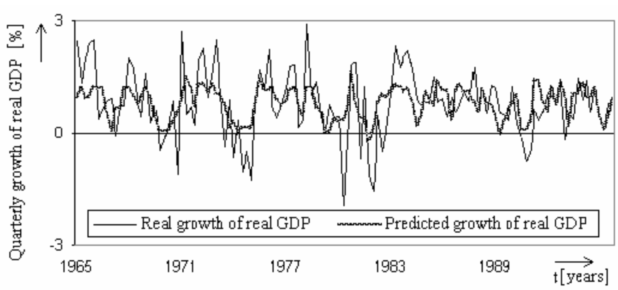

Fig.8 Prediction of quarterly growth of real GDP by means of the frontal NN, the feed-forward NN and the Takagi-Sugeno FIS

Tables 1, Table 2 and Table 3 consists of the means of average error ε, mean average error σ and root mean squared error δ of the prediction of the reference values, learning set and testing set.

Table 1 Values of the average error indicators of the frontal NN prediction

Indicator of the prediction error / Determination of time interval for which average error of the prediction is calculated	All years (1965-2004)	Learning set frontal NN (1965-2004)	Testing set frontal NN (1985-2004)	Years allowing comparison with FRBSF model (1985-2004)
ε	0.0679	- 0.4313	0.4307	0.4555
σ	1.2815	1.1816	1.3751	0.8088
δ	1.7738	1.6099	1.9148	1.1298

Table 2 Values of the average error indicators of the feed-forward NN prediction

Indicator of the prediction error / Determination of time interval for which average error of the prediction is calculated	All years (1965-2004)	Learning set feed-forward NN (1965-2004)	Testing set feed-forward NN (1985-2004)	Years allowing comparison with FRBSF model (1985-2004)
ε	0.4327	0.0142	0.8251	0.8359
σ	1.3557	1.2965	1.4113	1.0389
δ	1.8002	1.8096	1.7915	1.3401

Table 3 Values of the average error values indicators of the Takagi-Sugeno FIS prediction

Indicator of the prediction error / Determination of time interval for which average error of the prediction is calculated	Learning set of the Takagi-Sugeno FIS (1965-2004)	Testing set of the Takagi-Sugeno FIS (1985-2004)	Reference values of OECD model (1985-2004)	Reference values of FRBSF model (1985-2004)
ε	-	-	0.23	zero
σ	0.7567	0.9378	1.05	1.8451
δ	0.9251	1.2539	1.21	2.5111

It can be concluded on the basis of the data in Table 1, Table 2 and Table 3 that a mean prediction error of approximately zero indicates the absence of systematic errors in the model. The value of indicator δ reveals the existence of several larger individual errors of the prediction. The proposed model achieves, within the indicator of mean errors, superior accuracy (with reference to the FRBSF model or OECD model). Two times better accuracy was achieved by the designed frontal neural network when taking into account the δ criterion, which puts emphasis on large individual deviations from the prediction.

Economic variables were problematic in the years 1981, 1982 and 1991. In these years the ILEI had insufficient explanation ability for the changes of GDP, because the group of EVs included in the given index did not consist of all-important EVs, which were connected with GDP development in selected years. These factors explain the prediction errors in the years 1981, 1982 and 1991. The prediction errors in 1965, 1974 and 1984 are connected either with the parameter estimates of the model or with the mathematical formulation of the model.

6 Conclusion

The aim of the paper has been to show the possibilities of utilisation of frontal NNs, feed-forward NNs (without pre-processing of inputs time series) with learning by GAs and EuAs and the Takagi-Sugeno FIS (with pre-processing of inputs time series) in the prediction of the GDP development by designing a prediction model, which provides superior accuracy to the existing reference model. Models of the GDP development formulated as frontal NN, feed-forward NN was designed to fulfil this goal, whereas parameter estimates and selection of EVs were realised by the newly-designed three-grade EuSANE algorithm. A software tool was developed in the C^{++} programming language and run under Microsoft Windows. The Takagi-Sugeno FIS is designed in a MATLAB environment and is optimized by means of the ANFIS method. The approach of extraction of rules from the historical data is employed when constructing BRs of the designed FIS. The presented results of the designed models prediction of the GDP are presenting typical results of the experiments.

References:

[1] R. Ingenito, T. Bharat T, Using Monthly Data to Predict Quarterly Output, *FRBSF, Economic Review*, No.3, 1996.

[2] OECD Working Papers, *BIS Working Papers*, Vol.51, 1997.

[3] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd edition, Prentice-Hall, Inc., New York, 1999.

[4] V. Olej, *Modelling of Economics Processes on the Basis Computational Intelligence,* Miloš Vognar-M&V, Hradec Králové, Czech Republic, 2003.

[5] D. Polani, R. Miikkulainen, Fast Reinforcement Learning Through Eugenic Neuro-Evolution, *Technical Report AI 99-277*, Artificial Intelligence Laboratory, The University of Texas at Austin, Texas, 1999.

[6] L. I. Kuncheva, *Fuzzy Classifier Design*, A Springer Verlag Company, Germany, 2000.