HYBRID SYSTEM OF COMPUTATIONAL INTELLIGENCE
BASED ON BAGGING AND GROUP METHOD
OF DATA HANDLING

Ye. BODYANSKIY, O. KUZMENKO, He. ZAICHENKO, Yu. ZAYCHENKO

Abstract. The paper considers the problem of short- and middle-term forecasting in
the financial sphere. To solve this problem, a hybrid system of computational intel-
ligence based on the group method of data handling (GMDH) and bagging, as well
as an algorithm for its training, is proposed. The odd stacks of the hybrid system are
formed by ensembles of parallel connected subsystems. ARIMA and the GMDH-
eo-fuzzy hybrid network were chosen as such subsystems. The proposed system
does not require a large training data set, automatically determines the number of
stacks during training, and provides online operation. The experimental investiga-
tions were conducted using the proposed hybrid system, as well as separately using
ARIMA and GMDH-neo-fuzzy. The accuracy of the predictions obtained is compared,
based on which the feasibility of using the proposed hybrid system is substantiated.

Keywords: hybrid system, bagging, hybrid GMDH-neo-fuzzy network, ARIMA,
short- and middle-term forecasting.

INTRODUCTION

Today deep neural networks (DNN) are widely used for solving a large class of
Data mining problems and, above all, classification (pattern recognition, image
processing of various nature) and extrapolation (time series forecasting, natural
language text processing, fault detection), due to their universal approximating
properties, based mainly on the G. Cybenko [2] theorem. At the same time, DNNs
are rather cumbersome constructs that contain too many configurable synaptic
weights, in turn, requiring too much training data, which is not always available
for solving specific real-world problems. In addition, the learning process of DNN
requires quite a lot of time so that when solving Data Stream mining problems,
especially when processing nonstationary data streams, the use of these systems
encounters some difficulties.

It is possible to overcome these difficulties using ensemble methods [2–7],
while the ensemble can contain a variety of computational intelligence systems
the simplest type of elementary perceptron of F. Rosenblatt and N – Adaline [8]
to the most complex DNNs.
The main problem when using the ensemble approach is to combine the results of individual members of this ensemble in order to obtain the optimal initial result in the sensing of the accepted learning criterion. For this purpose, Bagging procedures [9] can be used, minimizing the RMS error on the training sample, modified to work online.

The resulting ensemble approach and bagging system may be too complex from a computational point of view. To simplify its implementation, it should be decomposed into a number of simpler subsystems, while such decomposition can be implemented quite simply using the Group Method of Data Handling (GMDH) [10; 11].

It is interesting to note that J. Schmidhuber [12] believes that just based on GMDH the first deep learning systems were built. Subsequently, based on GMDH, neural networks and neuro-fuzzy systems were proposed, which demonstrated their accuracy and speed in solving a number of problems [13–17].

In our opinion, it is appropriate to introduce into consideration a hybrid system of computational intelligence (HSCI), built based on an ensemble approach and bagging, which would increase its architecture in the learning process based on GMDH ideas, would not require significant amounts of training selections and would be quite simple from a computational point of view.

ARCHITECTURE OF HSCI ON THE BASE OF BAGGING AND GMDH

In the Fig. 1 the architecture of the proposed system is presented.

![Fig. 1. Hybrid system of computational intelligence based on bagging and GMDH](image)

The architecture of the system contains 2S sequentially-connected stacks, while odd stacks are formed by ensembles of parallel-connected subsystems that solve the same problem (recognition, prediction, etc.) and even ones are essentially learning metamodels that generalize the output signals of ensembles and form optimal results in the sense of the accepted criterion. The output signal of the first metamodel is the generalized optimal signal \( y^{*1}(k) \) and \((n-1)\) output signals \( y_{j_i}^{[1]}(k), i = 1, 2, \ldots, n-1 \) “best members of the ensemble”. At their core, metamodels function as selection units in traditional GMDH systems, but not only select the best results from the previous stack, but also form the optimal solution based on these results.

Further, the output signals of the first metamodel are fed to the inputs of the second ensemble, which is completely similar to the first. The outputs of the second ensemble \( y_{j}^{[2]}(k), y_{j_2}^{[2]}(k), \ldots, y_{j_q}^{[2]}(k) \) come to the second metamodel, which calculates the optimal signal \( y^{*2}(k) \) and \((n-1)\) \( y_{j_i}^{[2]}(k) \) “closest” to it. The last \( S \)-th ensemble is similar to the first two, and the output of the last \( S \)-th metamodel is: \( y^{*S}(k) \).
Hybrid system of computational intelligence based on bagging and group method...
Minimization (2) by \( w^{[*][1]} \) leads to the result

\[
w^{[*][1]} = \lim_{\delta \to 0} w^{[1]}(\delta) = w^{LS[1]} + P[[1]](N) \frac{1 - I_q^T w^{LS[1]} I_q}{I_q^T P[[1]](N) I_q},
\]

where \( w^{LS[1]} \) is a standard LSM estimate:

\[
\]

The same result may be obtained using the indefinite Lagrange multipliers. Introducing into consideration meta-model error

\[
e^{[1]}(k) = y(k) - y^{[*][1]}(k) = y(k) - \hat{y}^{[1]T}(k) w^{[*][1]} = w^{[*][1]} I_q y(k) - w^{[*][1]} \hat{y}^{[1]T}(k) = \]

\[
= w^{[*][1]} (I_q y(k) - \hat{y}^{[1]}(k)) = w^{[*][1]} E^{[1]}(k),
\]

write Lagrange function in the form

\[
L(w^{[*][1]}, \lambda) = \\
= \sum_{k=1}^{N} w^{[*][1]} E^{[1]}(k) E^{[1]T}(k) w^{[*][1]} + (I_q^T w^{[*][1]} - I_q) = w^{[*][1]} R(N) w^{[*][1]} + \lambda (I_q^T w^{[*][1]} - I_q) - 1,
\]

where \( \lambda \) is indefinite Lagrange multiplier and solving Kuhn–Tucker equations system

\[
\begin{cases}
F^{[1]T}(N) L(w^{[*][1]}, \lambda) = 2R((N) w^{[*][1]} + \lambda I_q = \bar{0}, \\
\partial L(w^{[*][1]}, \lambda) / \partial \lambda = I_q^T w^{[*][1]} - I_q = 0,
\end{cases}
\]

obtain the final result

\[
w^{[*][1]} = R^{-1}(N) I_q (I_q^T R^{-1}(N) I_q)^{-1},
\]

\[
\lambda = -2 I_q^T R^{-1}(N) I_q.
\]

It was proved [3; 4] that the use of score (3) leads to results that are not inferior in accuracy to the best of the members of the first ensemble.

If observations from the training sample are processed sequentially online, it is advisable to use the least squares recurrent method in the form

\[
P[[1]](k + 1) = P[[1]](k) - \frac{P[[1]](k) \hat{y}^{[1]}(k + 1) \hat{y}^{[1]T}(k + 1)}{1 + \hat{y}^{[1]T}(k + 1) P[[1]](k) \hat{y}^{[1]}(k + 1)},
\]

\[
w^{LS[1]}(k + 1) = w^{LS[1]}(k) + P[[1]](k + 1) \{y(k + 1) - \hat{y}^{[1]T}(k + 1) w^{LS[1]}(k) \hat{y}^{[1]}(k + 1),
\]

\[
w^{[*][1]}(k + 1) = w^{LS[1]}(k + 1) + P[[1]](k + 1) (I_q^T P[[1]](k + 1) I_q)^{-1} (1 - I_q^T w^{LS[1]}(k + 1) I_q),
\]

or if a training sample is non-stationary we may use exponentially weighted recurrent LSM method

\[
P[[1]](k + 1) = \frac{1}{\alpha} \left[ P[[1]](k) - \frac{P[[1]](k) \hat{y}^{[1]}(k + 1) \hat{y}^{[1]T}(k + 1) P[[1]](k)}{\alpha + \hat{y}^{[1]T}(k + 1) P[[1]](k) \hat{y}^{[1]}(k + 1)} \right],
\]

\[
w^{LS[1]}(k + 1) = w^{LS[1]}(k) + \frac{P[[1]](k) (y(k + 1) - \hat{y}^{[1]T}(k + 1) w^{LS[1]}(k)) \hat{y}^{[1]}(k + 1)}{\alpha + \hat{y}^{[1]T}(k + 1) P[[1]](k) \hat{y}^{[1]}(k + 1)},
\]

\[
w^{[*][1]}(k + 1) = w^{LS[1]}(k + 1) + P[[1]](k + 1) (I_q^T P[[1]](k + 1) I_q)^{-1} (1 - I_q^T w^{LS[1]}(k + 1) I_q),
\]

\[
w^{[*][1]}(0) = q^{-1} p = 1, 2, ..., q
\]
where \( 0 < \alpha \leq 1 \) — forgetting factor.

To the parameters of the metamodel can be given meaning the levels of fuzzy membership to the optimal output signal by introducing additional restrictions on the non-negative values of these parameters, that is, in addition to the configurable parameters \( w^{*[1]} \) we can also calculate the levels of this membership \( \mu_p^{[1]} \geq 0, \quad p = 1,2,\ldots,q \).

To do this, we introduce into consideration the extended Lagrange function

\[
L(\mu^{[1]}, \lambda, \rho) = (Y(N) - \bar{Y}^{[1]}(N)\mu^{[1]})^T (Y(N) - \bar{Y}^{[1]}(N)\mu^{[1]}) + \lambda (I_q^T \mu^{[1]} - 1) - \rho^T \mu^{[1]},
\]

where \( \rho \) — \((q \times 1)\) is vector of non-negative indefinite Lagrange multipliers.

Using the equations system by Kuhn–Tucker

\[
\left\{ \begin{array}{l}
\nabla L(\mu^{[1]}, \lambda, \rho) = 0, \\
\frac{\partial L(\mu^{[1]}, \lambda, \rho)}{\partial \lambda} / \rho \lambda = 0,
\end{array} \right.
\]

it’s not difficult to get the solution in the form

\[
\mu^{[1]} = p^{[1]}(N)(\bar{y}^T(N)Y(N) - 0,5\lambda I_q + 0,5\rho),
\lambda = \frac{I_q^T \rho^{[1]}(N)\bar{y}^T(N)Y(N) - 1 + 0,5I_q^T \rho^{[1]}(N)\rho}{0,5I_q^T \rho^{[1]}(N)I_q}
\]

For finding vector of non-negative Lagrange multipliers, it’s reasonable to apply Arrow–Hurwitz–Uzawa procedure

\[
\mu^{[1]}(k+1) = w^{LS[1]}(k+1) - p^{[1]}(k+1) - 0,5I_q^T \rho^{[1]}(k+1)I_q + 0,5I_q^T \rho^{[1]}(k+1)\rho(k) + 0,5p^{[1]}(k+1)\rho(k),
\]

\[
\rho(k+1) = P_r(\rho(k) - \eta_{\rho}(k)\mu^{[1]}(k+1)),
\]

where \( P_r(.) \) is a projector to positive orthant, \( \eta_{\rho}(k) \) — learning rate, parameter.

First expression (4) after non-complex transformations may be presented in a more compact form

\[
\mu^{[1]}(k+1) = w^{*[1]}(k+1) - p^{[1]}(k+1) - 0,5I_q^T \rho^{[1]}(k+1)I_q + 0,5p^{[1]}(k+1)\rho(k) =
\]

\[
= w^{*[1]}(k+1) + 0,5 \left( \frac{I_q^T \rho^{[1]}(k+1)I_q}{I_q^T \rho^{[1]}(k+1)I_q} \right) p^{[1]}(k+1)\rho(k)
\]

where instead of least squares estimates \( w^{LS[1]}(k+1) \), the parameters of the metamodel \( w^{*[1]}(k+1) \) are used, which simplifies the process of configuring it.

As a result of learning the first metamodel, the optimal signal \( y^{*[1]}(k) \) is formed at its output, as well as \( q \) signals \( y_{p,q}^{*[1]}(k) \) from which we choose \( n-1 \) (if \( q \geq n \)) with the highest levels of fuzzy membership \( \mu_p^{[1]} \), which subse-
quenty in the form of \((n \times 1)\) — vector are fed to the input of the second ensemble, the outputs of which go to the inputs of the second metamodel, and so on. The process of increasing the number of ensembles and metamodels continues until the required accuracy of the last metamodel with the output \(y^{[*1]}(k)\) is achieved, or the value of the criterion minimized for the bagging model begins to increase, i.e. \(\overline{e}^2(y^{[*1]}(k)) \geq \overline{e}^2(y^{[*1]}(k))\).

**EXPERIMENTAL INVESTIGATIONS**

The experimental investigations of bagging based on GMDH were performed at the problems of short-term and middle-term forecasting of Dow–Jones Industrial average index. The dynamics of DJIA is presented in the Fig. 2.

![Dynamics of Dow–Jones Average](image)

*Fig. 2. Dynamics of Dow–Jones Average*

The data of DJ index was taken since 05.07.22 till 03.07.23. The correlation function of process Dow–Jones index was calculated and the correlogram was built presented in the Fig. 3.

![Correlogram](image)

*Fig. 3. Correlogram*

Analyzing this correlogram we can see strong correlation between values of DJIA.
As a first model used in bagging ensemble is ARIMA. ACF function plot is presented in the Fig. 4.

Using differencing the process DJIA was transformed to stationary one: Dickey–Fuller test was performed: $P$-value: $1.3051439086544856 \times 10^{-28} < 0.05$. The experimental investigations were performed at different forecasting intervals: 1, 3, 5 (short-term) and 10, 20 days (middle-term) forecasting. Flow-chart for 1 day forecast is presented in the Fig. 5.

Flow chart of forecast by ARIMA for 5 days is shown in the Fig. 6.
Next model used in ensemble is Hybrid GMDH-neo-fuzzy network. It was optimized by parameters. After that the experiments on forecasting with different forecasting intervals 1, 3, 5, 10, 20 days were performed. Some of results are presented below. Flowchart of forecast for 3 days is shown in the Fig. 7 and for 10 days in the Fig. 8.

After that the algorithm of bagging based on GMDH was implemented and experiments at short-term and medium-term forecasting were performed. The tables with criteria MSE and MAPE are presented in the Tables 1 and 2.

**Table 1.** Average MSE values for different intervals

<table>
<thead>
<tr>
<th>Interval</th>
<th>ARIMA</th>
<th>GMDH-neo-fuzzy</th>
<th>HSCI-GMDH-bagging</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17422.752</td>
<td>27420.394</td>
<td>17382.425</td>
</tr>
<tr>
<td>3</td>
<td>43434.022</td>
<td>59202.99</td>
<td>49332.635</td>
</tr>
<tr>
<td>5</td>
<td>66993.011</td>
<td>78330.284</td>
<td>58153.202</td>
</tr>
<tr>
<td>10</td>
<td>235427.989</td>
<td>108358.616</td>
<td>104324.0</td>
</tr>
<tr>
<td>20</td>
<td>696291.974</td>
<td>253693.345</td>
<td>241508.146</td>
</tr>
</tbody>
</table>

_Fig. 7. Flow chart of forecast for 3 days interval by GMDH-neo-fuzzy network_

_Fig. 8. Flow chart of forecast by GMDH-neo-fuzzy network for 10 days_
**Table 2.** Average MAPE values for different intervals

<table>
<thead>
<tr>
<th>Interval</th>
<th>ARIMA</th>
<th>GMDH-neo-fuzzy</th>
<th>HSCI-GMDH-bagging</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.83</td>
<td>1.049</td>
<td>0.828</td>
</tr>
<tr>
<td>3</td>
<td>1.416</td>
<td>1.63</td>
<td>1.398</td>
</tr>
<tr>
<td>5</td>
<td>1.616</td>
<td>1.903</td>
<td>1.577</td>
</tr>
<tr>
<td>10</td>
<td>3.134</td>
<td>2.208</td>
<td>2.108</td>
</tr>
<tr>
<td>20</td>
<td>5.579</td>
<td>3.433</td>
<td>3.308</td>
</tr>
</tbody>
</table>

Average MSE values for different intervals are presented in the Fig. 9 and MAPE values — in the Fig. 10.

![Fig. 9. Average MSE values for different forecasting intervals](image)

![Fig. 10. Average MAPE values for different forecasting intervals](image)

Analyzing the presented results, one may conclude that bagging procedure based on GMDH has the best results as compared with separate models in ensemble, the second place takes GMDH-neo-fuzzy network for all the intervals at the exception 1, 3 and 5 days. And for short-term forecasting 1, 3 and 5 days ARIMA appears to be better than hybrid GMDH-neo-fuzzy network.
In a whole the obtained results well comply with theoretical statements as for properties of bagging procedure and the proposed bagging based on GMDH in HSCI system appeared to be very efficient procedure which demand minimum calculations due to application of GMDH.

CONCLUSION

The architecture and learning algorithms of hybrid computational intelligence system, which is built based on Group Method of Data Handling (GMDH) method and bagging approach, are proposed.

The system consists of a sequence of stacks, while odd stacks are essentially ensembles formed by parallel connected different subsystems that solve the same problem, while even ones are metamodels that implement the bagging procedure and calculate the levels of fuzzy membership of each member of the ensemble to the optimal result. The process of increasing the number of stacks is based on GMDH principles until the desired accuracy of the final results is achieved. The proposed system does not require large volumes of training samples, provides online work and automatically determines the number of its layers — stacks in the learning process.

Experimental investigations confirm that Hybrid system of computational intelligence based on bagging and GMDH is effective for short- and middle-term forecasting in the financial sphere and has better metrics than ARIMA and GMDH neo fuzzy.

REFERENCES


**INFORMATION ON THE ARTICLE**

Yevgeniy V. Bodyanskiy, ORCID: 0000-0001-5418-2143, Kharkiv National University of Radio Electronics, Ukraine, e-mail: yevgeniy.bodyanskiy@nure.ua

Oleksii V. Kuzmenko, ORCID: 0000-0003-1581-6224, Educational and Research Institute for Applied System Analysis of the National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Ukraine, e-mail: oleksii.kuzmenko@ukr.net

Helen Yu. Zaichenko, ORCID: 0000-0002-4630-5155, Educational and Research Institute for Applied System Analysis of the National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Ukraine, e-mail: syncmaster@bigmir.net

Yuriy P. Zaychenko, ORCID: 0000-0001-9662-3269, Educational and Research Institute for Applied System Analysis of the National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Ukraine, e-mail: zaychenkoyuri@ukr.net

**ГІБРИДНА СИСТЕМА ОБЧИСЛЮВАЛЬНОГО ІНТЕЛЕКТУ НА ОСНОВІ БЕГГІНГУ ТА МЕТОДУ ГРУПОВОГО УРАХУВАННЯ АРГУМЕНТІВ / Є.В.Бодянський, О.В. Кузьменко, Ю.П. Зайченко, О.Ю. Зайченко**

Анотація. Розглянуто проблему короткострокового та середньострокового прогнозування у фінансовій сфері. Для її вирішення запропоновано гібридну систему обчислювального інтелекту на основі методу групового урахування аргументів (МГУА) та бетгінгу, а також алгоритм її навчання. Непарні стеки гібридної системи сформовані ансамблями паралельно з’єднаних підсистем. Як такі підсистеми обрано ARIMA та гібридну мережу МГУА-нео-фаззі. За- пропонована система не потребує великої обсягу навчальної вибірки, автоматично визначає кількість стеків у процесі навчання та забезпечує роботу у режимі online. Проведено експериментальні дослідження з використанням запропонованої гібридної системи, а також окремо ARIMA та МГУА-нео-фаззі. Порівняно точність прогнозів, отриманих експериментальним шляхом, на основі чого обґрунтовано доцільність застосування запропонованої гібридної системи.

**Ключові слова:** гібридна система, бетгінг, гібридна мережа МГУА-нео-фаззі, ARIMA, короткострокове та середньострокове прогнозування.