computing@tanet.edu.te.ua
www.tanet.edu.te.ua/computing

ISSN 1727-6209
International Scientific
Journal of Computing

ADAPTIVE SELECTION OF NEURAL NETWORKS FOR A COMMITTEE DECISION

Arunas Lipnickas [1], Józef Korbicz [2]

1) Department of Control Technology, Kaunas University of Technology,
Studentų 48, 3031 Kaunas, Lithuania, E-mail: lipnick@soften.ktu.lt
2) Institute of Control and Computation Engineering, University of Zielona Góra,
ul. Podgórna 50, 65-246 Zielona Góra, Poland, E-mail: J.Korbicz@issi.uz.zgora.pl

Abstract: *To improve recognition results, decisions of multiple neural networks can be aggregated into a committee decision. In contrast to the ordinary approach of utilising all neural networks available to make a committee decision, we propose creating adaptive committees, which are specific for each input data point. A prediction network is used to identify classification neural networks to be fused for making a committee decision about a given input data point. The j^{th} output value of the prediction network expresses the expectation level that the j^{th} classification neural network will make a correct decision about the class label of a given input data point. The proposed technique is tested in three aggregation schemes, namely majority vote, averaging, and aggregation by the median rule and compared with the ordinary neural networks fusion approach. The effectiveness of the approach is demonstrated on three well known real data sets and also applied to fault identification of the actuator valve at one sugar factory within the DAMADICS RTN.*

Keywords: *Adaptive committees, neural networks, half&half sampling*

1. INTRODUCTION

It is well known that a combination of many different neural networks can improve classification accuracy. A variety of schemes have been proposed for combining multiple classifiers. The approaches used most often include the majority vote [1], the averaging [2], weighted averaging [3,4], the Bayesian approach [1], the fuzzy integral [4,5,6], the Dempster-Shafer theory [1], the Borda count [4], probabilistic aggregation, and aggregation by a neural network.

For some of the aforementioned approaches we can say that a combiner assigns weights of value to neural networks in one way or another. Aggregation weights assigned to neural networks or groups of them can be the same in the entire data space or can be different, data-dependent, in various regions of the space [7,8]. The use of data-dependent weights, when properly estimated, provides higher classification accuracy [7]. Space partitioning into regions can be defined in advance [9] or can be adaptive, given by k nearest neighbours of a data sample being considered, for example [7]. The adaptive space partitioning approach, however, is very expensive in computation time.

Two main approaches have prevailed in utilising neural networks, or ordinary classifiers, for building a committee. The most predominant one is to use all the networks available for making a committee decision. The alternative approach selects a single network, which is most likely to be correct for a given sample. In this case, aggregation weights are binary: $w_i \in \{0,1\}$, $i=1,\ldots,L$, where L is the number

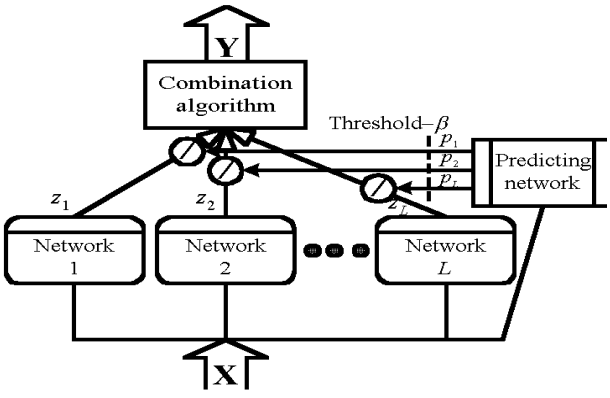

Fig. 1 - Architecture of the proposed combination scheme based on a dynamic neural network selection by a prediction MLP neural network.

of networks and $w_i = 1$ only for the most accurate network in the neighbourhood of a given sample.

Thus only the output of the selected network is considered in the final decision.

In this paper, we propose an approach for building adaptive, data-dependent, committees, which are specific for each input data point. Depending on an input data point, different networks and a different number of them may be chosen to make a committee decision about the data point. A prediction network is trained to predict the behavior of committee members for each data point, and is further used to select neural networks to be fused for making the committee decision (Fig. 1, where *z* and *p* stand for the outputs of the networks).

Previous works on neural classifier committees design have shown that an efficient committee should consist of networks that are not only very accurate, but also diverse in the sense that the network errors occur in different regions of the input space. Clearly, there are no advantages in combining the networks, which generalise identically, no matter how ingenious a combination method is employed.

Bootstrapping [10], Boosting, and AdaBoosting are the most often used approaches for data sampling when training members of neural network committees. It has been recently shown that half&half bagging through majority voting is capable of creating very accurate committees of decision trees [11]. Data sampling by half&half bagging focuses on the most often misclassified data points from the training data set.

To obtain diverse networks comprising a committee, we use the bootstrapping and half&half sampling approaches [11] to collect data for training neural networks of the committee. Four real world problems have been used to evaluate the proposed approach. We compare the technique developed with the ordinary decision fusion scheme when all the networks available are utilised to make a committee decision.

The remainder of the paper is organised as follows. In the next Section, we briefly describe the data sampling approach used to collect data for training members of a neural network committee. The neural networks selection procedure proposed is presented in Section 3. Section 4 describes the aggregation schemes used. The databases used to test the approach proposed are briefly described in Section 5. Section 6 presents the results of the experimental investigations. Finally, conclusions of the work are given in Section 7.

2. DATA SAMPLING TECHNIQUES

Networks Diversity: One way to measure the diversity of neural networks is to construct κ-error diagrams. The diagrams display the accuracy and diversity of the individual networks. For each pair of networks, the accuracy is measured as the average error rate on the test data set, while the diversity is evaluated by computing the so-called a *degree-of-agreement* statistic κ. Each point in the diagrams corresponds to a pair of networks and illustrates their diversity and the average accuracy. The κ statistic is computed as:

$$\kappa = \frac{\Theta_1 - \Theta_2}{1 - \Theta_2}. \qquad (1)$$

with $\Theta_1 = \sum_{i=1}^{Q} c_{ii} \Big/ N$ and $\Theta_2 = \sum_{i=1}^{Q} \left\{ \sum_{j=1}^{Q} \frac{c_{ij}}{N} \sum_{j=1}^{Q} \frac{c_{ji}}{N} \right\}$

where Q is the number of classes, C is a QxQ square matrix with c_{ij} containing the number of test data points assigned to class i by the first network and into class j by the second network and N stands for the total number of test data. We used the statistic κ to evaluate the diversities of trained neural networks committee.

Half&Half sampling: The basic idea of the half&half sampling is very simple. It is assumed that the training set contains N data points. Suppose that k classifiers have been already constructed. To obtain the next training set, randomly select a data point **x**. Present **x** to that subset of k classifiers, which did not use **x** in their training sets. Use the majority vote to predict the classification result of **x** by the subset of classifiers. If **x** is misclassified, put it in set MC. If not, put **x** in set CC. Stop, when the sizes of both MC and CC are equal to M, where $2M \le N$. In [11], $M = N/4$ has been used. In this work, we investigate the effectiveness of the half&half sampling approach in creating accurate neural network committees for classification. We compare this approach with the bootstrapping technique

Bootstrapping: In bootstrapping, each training set is constructed by forming a bootstrap replicate of the original training set. In other words, given a training set of N data points, a new training set is constructed by randomly drawing N data points (with replacements) from the original data set. Each individual neural network in the committee is trained with a particular bootstrap replicate.

3. PROCEDURES FOR SELECTING AND AGGREGATING NEURAL NETWORKS

The *neural networks selection* procedure is encapsulated in the following six steps:

1. Divide the available data into training, test, and cross-validation data sets.
2. Train L neural networks using the half&half sampling technique.
3. Classify the training data set by all networks of the committee.
4. For each training data vector $\mathbf{x}_i$ form a L-

dimensional target vector $t_i = [t_{i1}, ..., t_{iL}]^T$, with t_{ij}=1, if the $\mathbf{x}_i$ data vector was correctly classified by the j^{th} network and t_{ij}=0, otherwise.

5. Using the training data set and the target vectors obtained in Step 4 train a neural network to predict whether or not the classification result obtained from the L networks for an input data point **x** will be correct. The prediction network consists of L output nodes and n input nodes, where n is the number of components in **x**. Therefore, each output node stands for one particular network. The number of hidden nodes needs to be determined.
6. Determine the optimal threshold value β for including neural networks into a committee. The j^{th} network is included into a committee if $p_j > \beta$, where p_j is the j^{th} output of the prediction network. The value β is the value yielding the minimum cross-validation data set classification error obtained from a committee of the selected networks.

Having the threshold β determined, *data classification* proceeds as follows:

1. Present a data point x to the prediction network and calculate the output vector **z**.
2. Classify the data point by the networks satisfying the condition $p_j > \beta$.
3. Aggregate the outputs of the selected networks into a committee decision according to a chosen combination algorithm.

Note that the optimal threshold value is determined in the training phase and then fixed for the use in the classification phase. Note also that the committee build is specific for each input data point. This seems reasonable, since the L neural networks may have different accuracy in different regions of the input space.

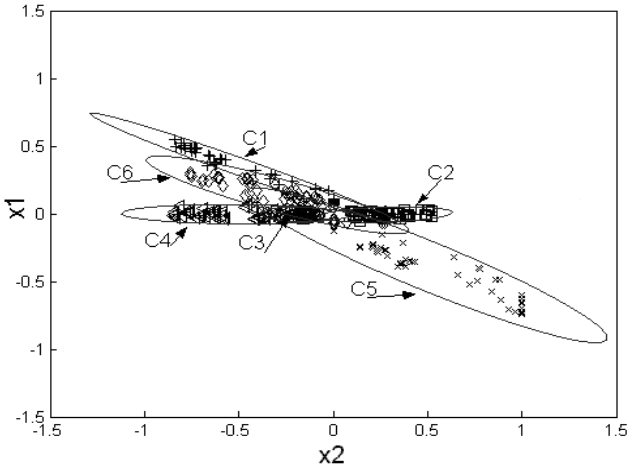

Fig. 2 - The *Damadics* data set.

We investigate three schemes for aggregating the outputs of the selected networks. In the context of the aggregation schemes used, we compare the proposed concept with an ordinary decision aggregation approach, when all the trained networks are utilised to make a committee decision.

4. AGGREGATION SCHEMES USED

To test the proposed approach we used three simple aggregation schemes that do not utilise any aggregation parameters, namely the *majority vote, averaging*, and the *median* aggregation rule. We now briefly describe the aggregation schemes used.

Majority vote. The correct class is the one chosen by the most neural networks. If all the neural networks indicate different classes, then the neural network with the overall maximum output value is selected to indicate the correct class. Ties can be broken randomly.

Averaging. This approach simply averages the individual neural network outputs. The output yielding the maximum of the averaged values is chosen as the correct class q:

$$q = \arg \max_{j=1,...,Q} \left(Z_j = \frac{1}{L} \sum_{i=1}^{L} z_{ji}(\mathbf{x}) \right) \qquad (2)$$

where Q is the number of classes, L is the number of neural networks, z_{ij} (**x**) represents the j^{th} output of the i^{th} network given an input pattern **x**, and Z_j(**x**) is the j^{th} output of the committee given an input pattern **x**.

Median rule. In some cases, when a classifier in a combined group is very sensitive to outliers, then the group decision could lead to an error. It is well known that a robust estimate of the averaging is the median. The median combination leads to the following rule:

$$q = \arg \max_{j=1,...,Q} \left(Z_j = \operatorname*{med}_{i=1}^{L} \left(z_{ji}(\mathbf{x}) \right) \right). \qquad (3)$$

5. DATA

The ESPRIT Basic Research Project Number 6891 (*ELENA*) provides databases and technical reports designed for testing both conventional and neural classifiers. All the databases and technical reports are available via anonymous *ftp: ftp.dice.ucl.ac.be* in the directory *pub/neural-net/ELENA/databases*. From the ELENA project we have chosen the two real data sets, *Phoneme* and *Satimage*. The *Thyroid* database has been taken from a collection called PROBEN 1, which represents a medical diagnosis task. The additional, *Damadics* database was generated from the Lublin (Poland) sugar factory process data [13].

The multi-disciplinary and complementary Research Training Network project "DAMADICS" is focussed on development and application of methods for actuator diagnosis in industrial control systems.

During the dynamic sugar production process, there is a possibility that there will occur one of the 19 faults in the actuator valve block, with four types of strength i.e.: abrupt {*small, medium, big*} and incipient. The early detection and isolation of those faults minimise damages in the industrial line.

For the fault detection the *model-based fault diagnosis* approach [14, 15] is used. The idea of such fault detection is to compare output signals of the model and the process, thereby generating the residual or an output error, witch is used to make the decision about the state of process. But some of the *Damadics* faults are undetectable by *model-based* approach and cannot be isolated. Therefore, in this paper, not all faults were investigated, only "big abrupt".

The data set consist of two dimensional data describing the state of actuator's valve, i.e.: the residuals of the rod displacement ($x1$) and the juice flow ($x2$), where co-ordinates close to $x1=0$ and $x2=0$ express "no fault" situation. The data with faults is presented in Fig. 2.

Observing the Fig. 2 it is possible to distinguish six less overlapped data clusters. Data from those clusters were assigned to the classes, for example faults F13 and F18 are highly overlapped and cannot be isolated separately, therefore they have the same class label (Table 1).

The data sets used are summarised in Table 2. The benchmark (BM) errors presented in Table 2 are taken from the ELENA project and the PROBEN 1 database. In the ELENA project, the errors presented are the average errors obtained when using an MLP with two hidden layers of 20 and 10 units, respectively. To solve the *Thyroid* task, an MLP with two hidden layers of 16 and 8 units, respectively, has been employed.

Table 1. Summary of considered faults from the *Damadics* benchmark

Class Label	Description
C1	F1 – Valve clogging
C2	F2 – Valve plug or valve seat sedimentation
C3	F19 – Flow rate sensor fault
C4	F13, F18 – Rod displacement sensor fault and fully or partly opened bypass valve
C5	F15 – Positioner feedback fault
C6	F10, F16 – Servomotor's diaphragm perforation and positioner supply pressure drop

Table 2. Summary of the data sets used

Data Set	# of classes	# of features	#of samples	BM error
Phoneme	2	5	5404	16.4
Satimage	6	5	6435	11.9
Thyroid	3	21	7200	1.31
Damadics	6	2	10000	-

6. EXPERIMENTAL TESTING

All comparisons between the different aggregation schemes presented here have been performed by leaving aside 10% of the data available as a *Cross-Validation* data set and then dividing the rest of the data into *Training* and *Test* sets of equal size. In all the tests, one hidden layer MLPs with 10 sigmoidal hidden units served as committee members. This architecture was adopted after some experiments. The experiments showed that the network used in the *ELENA* project was too large. Since we only investigate different aggregation schemes, we have not performed expensive experiments for finding the optimal network size for each data set used.

We run each experiment ten times, and the *mean* errors and *standard deviations* of the errors are calculated from these ten trials. In each trial, the data set used is randomly divided into *Training, Cross-Validation,* and *Test* parts. The size of the data sets *MC* and *CC*, used in the half&half sampling, were set to $M = N_i/4$, where N_i is the size of the learning set.

In the first set of experiments, we investigated the ability of the half&half sampling technique to create diverse and accurate neural networks. We compared the half&half technique with the bootstrapping sampling approach. Two training techniques have been employed in these tests: 1) the Bayesian inference technique [12] to obtain regularised networks and 2) the standard backpropagation training technique without regularisation, which was run for a high number of training iterations. The size of the committees was varied from 2 to 20 members.

These tests have shown that, on average, the half&half sampling technique outperformed the bootstrapping approach by creating more accurate neural network committees. Fig. 3 illustrates the *Test* set classification error of the committees for the different databases as a function of the committee size. Aggregation by the majority vote rule has been used in these experiments. Fig. 4 present the κ-error diagrams for the *Phoneme* data set illustrating the diversity of the networks of the committees made of 20 members. The following notations are used in the figures. BS_R stands for the bootstrapping sampling and regularised training case, BS stands for the bootstrapping sampling and training without regularisation, $H\&H_R$ means the

half&half sampling approach and regularised training, and H&H stands for the half&half sampling approach and training without regularisation.

As can be seen from the κ-error diagrams, the networks created by bootstrapping form a much tighter cluster than they do with the half&half sampling approach. This is expected, since with the bootstrapping technique each network is trained on a sample drawn from the same distribution. This explains why half&half sampling outperforms bootstrapping. The lower accuracy of networks produced by the half&half sampling approach is well compensated by the increased diversity. The same pattern of accuracy and diversity was observed across the other data sets. For both sampling techniques the regularised committees, on average, were more accurate than the non-regularised ones.

We can, therefore, conclude that the half&half sampling technique is capable of creating diverse and sufficiently accurate neural networks. A more thorough comparison of the bootstrapped and half&half sampling committees can be found in [6,8].

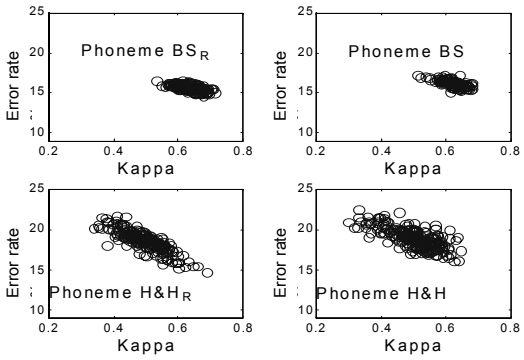

Fig. 4 -κ-Error diagrams for the *Phoneme* data set using Bootstrapped (top) and half&half sampled (bottom) committees.

In the next set of experiments, we investigated the effectiveness of the adaptive networks selection technique in creating accurate neural network committees. The regularised training for both sampling techniques have been employed in these tests. In the ordinary decision aggregation approach, without the proposed neural networks selection procedure, we utilised committees consisting of 20 members. The actual average size of the committees created by the

Table 3. The test data set classification error rate obtained from the bootstrapped neural network committees fused by the Majority Vote, Averaging, and Median aggregation rules

Without selection								
	The best		Majority		Averaging		Median	
Database	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Phoneme	14.27	0.33	12.23	0.28	12.26	0.32	12.08	0.31
Satimage	11.80	0.15	11.27	0.20	11.22	0.17	11.23	0.15
Thyroid	1.51	0.06	1.16	0.09	1.19	0.09	1.14	0.09
Damadics	7.25	0.09	6.95	0.13	6.91	0.11	6.84	0.13
With proposed selection								
Phoneme	14.27	0.33	10.48	0.16	10.75	0.18	10.38	0.16
Satimage	11.80	0.15	10.11	0.21	10.16	0.21	10.13	0.20
Thyroid	1.51	0.06	0.83	0.06	0.87	0.06	0.83	0.06
Damadics	7.25	0.09	6.90	0.18	6.74	0.16	6.78	0.18

Table 4. The test data set classification error rate obtained from the half&half sampled neural network committees fused by the Majority Vote, Averaging, and Median aggregation rules

Without selection								
	The best		Majority		Averaging		Median	
Database	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Phoneme	15.07	0.86	11.20	0.20	11.10	0.11	11.06	0.18
Satimage	11.87	0.21	10.37	0.19	10.23	0.18	10.34	0.22
Thyroid	1.46	0.19	0.74	0.09	0.72	0.11	0.72	0.12
Damadics	7.58	0.16	6.65	0.11	6.61	0.10	6.61	0.08
With proposed selection								
Phoneme	15.07	0.86	10.21	0.24	10.36	0.20	10.11	0.22
Satimage	11.87	0.21	9.86	0.15	9.76	0.23	9.85	0.26
Thyroid	1.46	0.19	0.53	0.06	0.53	0.07	0.53	0.07
Damadics	7.58	0.16	6.57	0.18	6.57	0.20	6.59	0.14

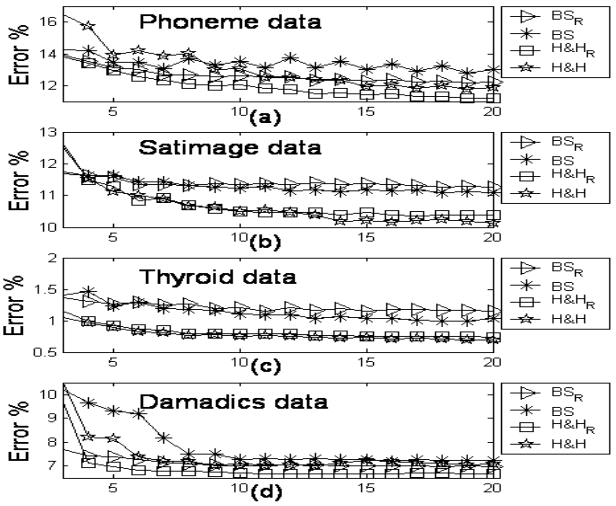

Fig. 3 - Classification error as a function of the committee size for: a) the *Phoneme* data set, b) the *Satimage* data set, c) the *Thyroid* data set, and d) the *Damadics* data set.

procedure proposed was considerably smaller. The prediction network was found to have 25 nodes in the hidden layer. Table 3 and Table 4 summarise the *Test* data set classification error obtained in these tests. The following notations are used in the tables: *Mean* stands for the percentage of the average test set classification error, *Std* is the standard deviation of the error, and *The best* means the single neural network with the best average performance.

As can be seen from the tables, there is an obvious improvement in classification accuracy when combining networks. While all the three aggregation schemes yielded approximately the same performance, the proposed neural networks selection and aggregation technique is superior to the ordinary aggregation approach when all the networks available are aggregated to make a committee decision. As can be seen, decrease in the classification error rate is observed for both training techniques. The decrease of the classification error rate obtained from the use of the bootstrapped committees is not so noticeable, since this training technique creates less diverse neural networks.

Table 5 provides the average number of neural networks included into a committee from the 20 available for the aggregation by majority voting rule and databases. The table also presents the average value of the optimal neural network selection threshold β found for the different cases. The value of $\beta = 0$ implies using all the networks available to make committee decisions. The average number of neural networks selected is far below the 20 available. Therefore, the technique proposed allows reducing both classification error and computational time by removing unreliable classifiers.

Table 5. The average number of selected neural networks from the 20 available and the average value of the optimal selection threshold found for the aggregation by majority vote rule

bootstrapping		
Database	# Selected NN	Threshold β
Phoneme	9.5	0.44
Satimage	6.8	0.41
Thyroid	12.2	0.48
Damadics	13.1	0.57
half&half sampling		
Database	# Selected NN	Threshold β
Phoneme	13.0	0.43
Satimage	10.9	0.29
Thyroid	13.6	0.36
Damadics	13.4	0.28

Fig. 5 plots the *Test* data set classification error rate of the committee for the *Damadics* data set as a function of the neural network selection threshold β. In this experiment, the majority vote rule has been used to aggregate the selected networks into a committee. The graphs presented show the strong dependence between the threshold value and the classification error rate.

7. CONCLUSIONS

In this paper, we have used the bootstrapping and half&half sampling techniques to collect data sets for training neural network committees. In all the tests performed, the half&half data sampling approach outperformed the bootstrapping technique.

A new approach to create adaptive neural network committees for classification was proposed. The approach banks on the idea of having a committee specific for each input data point. Different networks and a different number of them may be adaptively selected and fused into a committee to make a decision about different input data points. The networks utilised are determined by those outputs of a prediction network, the output value at which exceeds a particular selection threshold. The j^{th} output value expresses the expectation level that the j^{th} classification neural network will make a correct decision about the class label of a given input data point.

The effectiveness of the proposed approach in creating accurate neural network committees for classification was investigated using four real data sets. The approach proposed was compared with the ordinary neural networks fusion scheme. The comparisons were made for three neural networks aggregation approaches, namely *majority vote*, *averaging*, and aggregation by the *median rule*.

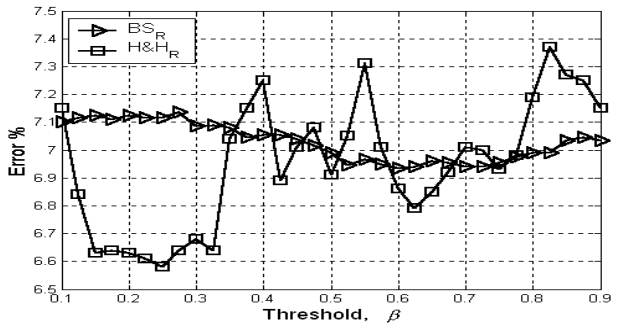

Fig. 5 - The test data set classification error rate of the committee for the *Damadics* data set as a function of the neural network selection threshold β for the committees created by the bootstrapping and the half&half sampling.

In all the tests performed, the proposed way of generating neural network committees was superior to the ordinary decision fusion scheme when all the networks available are utilised to make a committee decision.

8. ACKNOWLEDGEMENTS

This work was performed at *Institute of Control and Computation Engineering*, University of Zielona Góra (Poland) and was funded by EU FP 5 Research Training Network project DAMADICS: *Development and Application of Methods for Actuator Diagnosis in Industrial Control Systems.*

9. REFERENCES

[1] L. Xu, A. Krzyzak, C.Y. Suen, Methods for combining multiple classifiers and their applications to handwriting recognition. *IEEE Trans. Systems, Man, and Cybernetics* 22(3), pp. 418-435, 1992.

[2] V. Tresp and M. Taniguchi, *Combining estimators using non-constant weighting functions*. In G. Tesauro, D. S. Touretzky, and T. K. Leen, editors, Advances in Neural Information Processing Systems 7, MIT Press, 1996.

[3] S. Hashem, Optimal linear combinations of neural networks. *Neural Networks* 10(4), pp. 599-614, 1997.

[4] A. Verikas, A. Lipnickas, K. Malmqvist, M. Bacauskiene, A. Gelzinis, Soft combination of neural classifiers: A comparative study. *Pattern Recognition Letters* 20, 429-444, 1999.

[5] M. Grabisch and J.-M. Nicolas, Classification by fuzzy integral: Performance and tests. *Fuzzy Sets and Systems* 65, 255-271, 1994.

[6] A. Verikas, A. Lipnickas, M. Bacauskiene, K. Malmqvist, *Fusing neural networks through fuzzy integration*. In H. Bunke, A. Kandel, editors, Hybrid Methods in Pattern Recognition, World Scientific, 2002.

[7] K. Woods, W.P. Kegelmeyer, K. Bowyer, Combination of multiple classifiers using local accuracy estimates. *IEEE Trans. Pattern Analysis and Machine Intelligence* 19(4), 405-410, 1997.

[8] A. Verikas and A. Lipnickas, *Fusing neural networks through space partitioning and fuzzy integration*, *Neural Processing Letters*, Kluwer Academic Publishers, 2002.

[9] A. Lipnickas, Classifiers Fusion With Data Dependent Aggregation Schemes, Proceedings of the 7th *International Conference on Information Networks, System and Technologies,* ICINASTe'2001, Minsk, Belarus, October 2-4, pp.:147-154, 2001.

[10] B. Efron, R. Tibshirani, *An introduction to the bootstrap*, London, Chapman and Hall, 1993.

[11] L. Breiman, Half&Half bagging and hard boundary points. *Technical report 534*, Statistics Departament, University of California, Berkeley, 1998. (www.stat.berkley.edu /users/breiman).

[12] D.J. MacKay, Bayesian interpolation, *Neural Computation*, 4, 415-447, 1992.

[13] EC FP5 Research Training Network DAMADICS: *Development and Application of Methods for Actuator Diagnosis in Industrial Control Systems,* (http://www.eng.hull.ac.uk/research/control/damadics1.htm).

[14] R. J. Patton, P. M. Frank, R. N. Clark (ed): *Issues of Fault Diagnosis for Dynamic Systems*. – London: Springer Verlag, 2000.

[15] J. Chen, R. J. Patton, *Robust model-based fault diagnosis for dynamic systems*, Kluwer Academic Publisher, 1999.

Arunas Lipnickas *received the B.S. and M.S. degrees in Electrical Engineering from the Kaunas University of Technology, Lithuania, in 1992 and 1998, respectively.*

In 1998 he joined the Neural Networks and Image Processing research group at Kaunas University of Technology, Lithuania, where he has prepared the PhD thesis.

In 2002 he defended PhD thesis "Decision fusion in multiple classifier system" and gained the degree of Doctor of Science in Informatics Engineering from Kaunas University of Technology, Lithuania.

Afterwards, for 6 months, he has joined the DAMADICS project at University of Zielona Góra, Institute of Control and Computation Engineering, Poland. The head of the institute and co-ordinator of the project was Prof. Józef Korbicz.

Recently, Arunas Lipnickas is the scientific worker at Kaunas University of technology, Lithuania, Department of Control technology,

Electric drives group. His current research interests are neural fuzzy systems, intelligent control, electric drives, system modelling, and fault detection and isolation (FDI) problems.

Józef Korbicz *was born in Wysoka, Poland in 1951. He received his M.Sc. degree in 1975 and Ph.D. degree in 1980, both in automatic control from Kiev University of Technology, (KPI), Ukraine. In 1986 he received his D.Sc. degree (habilitation) from Kiev University of Technology, as well. Since 1975 he has been at University of Zielona Góra, Poland, where since 1994 he is a full professor of automatic control. Since 1992 he is the director of the Institute of Control and Computation Engineering. Moreover in 1991 he founded the International Journal of Applied Mathematics and Computer Science,(AMCS) and up to now he is the editor-in-chief.*

His main fields of research interests include computational intelligence (artificial neural networks, fuzzy and neuro-fuzzy systems), fault detection and isolation (analytical methods, soft computing) and control theory (distributed parameter systems, sensor and actuator locations). He is the author or co-author about 220 journal and conference papers, and nine books. He is co-editor (with J.M. Kościelny, Z. Kowalczuk and W. Cholewa) of the book Fault Diagnosis. Models, Artificial Intelligence, Applications (Springer-Verlag, 2004).

Professor Korbicz is the vice-chairman of the Automatics and Robotics Committee of Polish Academy of Sciences since 2003, and the senior member of IEEE since 2002.