Adaptive Mixture of Local Expert

4 authors, including:

Robert Jacobs
University of Rochester
131 PUBLICATIONS **13,613** CITATIONS

Michael Jordan
University of California, Berkeley
768 PUBLICATIONS **128,525** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project — Abstract representations of object directed action in the left inferior parietal lobule. View project

Project — Spectral clustering: perturbation, approximation and fast computation View project

Adaptive Mixtures of Local Experts

Robert A. Jacobs
Michael I. Jordan
Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology,
Cambridge, MA 02139 USA

Steven J. Nowlan
Geoffrey E. Hinton
Department of Computer Science, University of Toronto,
Toronto, Canada M5S 1A4

We present a new supervised learning procedure for systems composed of many separate networks, each of which learns to handle a subset of the complete set of training cases. The new procedure can be viewed either as a modular version of a multilayer supervised network, or as an associative version of competitive learning. It therefore provides a new link between these two apparently different approaches. We demonstrate that the learning procedure divides up a vowel discrimination task into appropriate subtasks, each of which can be solved by a very simple expert network.

1 Making Associative Learning Competitive ──────────────

If backpropagation is used to train a single, multilayer network to perform different subtasks on different occasions, there will generally be strong interference effects that lead to slow learning and poor generalization. If we know in advance that a set of training cases may be naturally divided into subsets that correspond to distinct subtasks, interference can be reduced by using a system composed of several different "expert" networks plus a gating network that decides which of the experts should be used for each training case.[1] Hampshire and Waibel (1989) have described a system of this kind that can be used when the division into subtasks is known prior to training, and Jacobs et al. (1990) have described a related system that *learns* how to allocate cases to experts. The idea behind such a system is that the gating network allocates a new case to one or a few experts, and, if the output is incorrect, the weight changes are localized to these experts (and the gating network).

---

[1]This idea was first presented by Jacobs and Hinton at the Connectionist Summer School in Pittsburgh in 1988.

So there is no interference with the weights of other experts that special-
ize in quite different cases. The experts are therefore local in the sense
that the weights in one expert are decoupled from the weights in other
experts. In addition they will often be local in the sense that each expert
will be allocated to only a small local region of the space of possible input
vectors.

Unfortunately, both Hampshire and Waibel and Jacobs *et al.* use an
error function that does not encourage localization. They assume that the
final output of the whole system is a linear combination of the outputs of
the local experts, with the gating network determining the proportion of
each local output in the linear combination. So the final error on case c
is

$$E^c = \|\mathbf{d}^c - \sum_i p_i^c \mathbf{o}_i^c\|^2 \tag{1.1}$$

where $\mathbf{o}_i^c$ is the output vector of expert i on case c, p_i^c is the proportional
contribution of expert i to the combined output vector, and $\mathbf{d}^c$ is the
desired output vector in case c.

This error measure compares the desired output with a blend of the
outputs of the local experts, so, to minimize the error, each local expert
must make its output cancel the residual error that is left by the combined
effects of all the other experts. When the weights in one expert change,
the residual error changes, and so the error derivatives for all the other
local experts change.[2] This strong coupling between the experts causes
them to cooperate nicely, but tends to lead to solutions in which many
experts are used for each case. It is possible to encourage competition by
adding penalty terms to the objective function to encourage solutions in
which only one expert is active (Jacobs *et al.* 1990), but a simpler remedy
is to redefine the error function so that the local experts are encouraged
to compete rather than cooperate.

Instead of linearly combining the outputs of the separate experts, we
imagine that the gating network makes a stochastic decision about which
single expert to use on each occasion (see Fig. 1). The error is then the
expected value of the squared difference between the desired and actual
output vectors

$$E^c = \langle \|\mathbf{d}^c - \mathbf{o}_i^c\|^2 \rangle = \sum_i p_i^c \|\mathbf{d}^c - \mathbf{o}_i^c\|^2 \tag{1.2}$$

Notice that in this new error function, each expert is required to pro-
duce the whole of the output vector rather than a residual. As a result,
the goal of a local expert on a given training case is not directly affected
by the weights within other local experts. There is still some indirect

---

[2]For Hampshire and Waibel, this problem does not arise because they do not learn the
task decomposition. They train each expert *separately* on its own preassigned subtask.
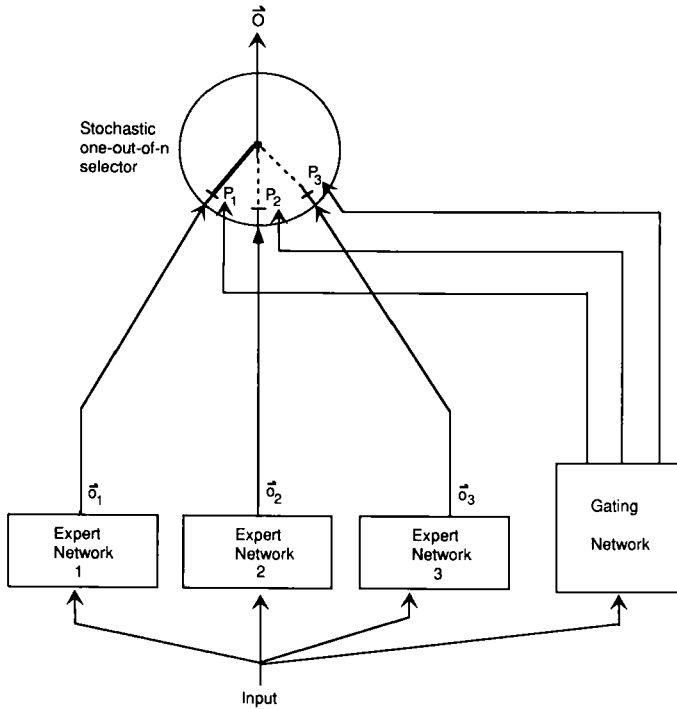
Figure 1: A system of expert and gating networks. Each expert is a feed-forward network and all experts receive the same input and have the same number of outputs. The gating network is also feedforward, and typically receives the same input as the expert networks. It has normalized outputs $p_j = \exp(x_j)/\sum_i \exp(x_i)$, where x_j is the total weighted input received by output unit j of the gating network. The selector acts like a multiple input, single output stochastic switch; the probability that the switch will select the output from expert j is p_j.

coupling because if some other expert changes its weights, it may cause the gating network to alter the responsibilities that get assigned to the experts, but at least these responsibility changes cannot alter the *sign* of the error that a local expert senses on a given training case. If both the gating network and the local experts are trained by gradient descent in this new error function, the system tends to devote a single expert to each training case. Whenever an expert gives less error than the weighted average of the errors of all the experts (using the outputs of the gating network to decide how to weight each expert's error) its responsibility for that case

will be increased, and whenever it does worse than the weighted average
its responsibility will be decreased.

The error function in equation 1.2 works in practice but in the simulations reported below we used a different error function which gives
better performance:

$$E^c = -\log \sum_i p_i^c e^{-\frac{1}{2}\|\mathbf{d}^c - \mathbf{o}_i^c\|^2} \tag{1.3}$$

The error defined in equation 1.3 is simply the negative log probability
of generating the desired output vector under the mixture of gaussians
model described at the end of the next section. To see why this error
function works better, it is helpful to compare the derivatives of the two
error functions with respect to the output of an expert. From equation 1.2
we get

$$\frac{\partial E^c}{\partial \mathbf{o}_i^c} = -2p_i^c(\mathbf{d}^c - \mathbf{o}_i^c) \tag{1.4}$$

while from equation 1.3 we get

$$\frac{\partial E^c}{\partial \mathbf{o}_i^c} = -\left[\frac{p_i^c e^{-\frac{1}{2}\|\mathbf{d}^c - \mathbf{o}_i^c\|^2}}{\sum_j p_j^c e^{-\frac{1}{2}\|\mathbf{d}^c - \mathbf{o}_j^c\|^2}} \right] (\mathbf{d}^c - \mathbf{o}_i^c) \tag{1.5}$$

In equation 1.4 the term p_i^c is used to weight the derivative for expert i.
In equation 1.5 we use a weighting term that takes into account how well
expert i does relative to other experts. This is a more useful measure of
the relevance of expert i to training case c, especially early in the training. Suppose, for example, that the gating network initially gives equal
weights to all experts and $\|\mathbf{d}^c - \mathbf{o}_i^c\| > 1$ for all the experts. Equation 1.4
will adapt the best-fitting expert the *slowest*, whereas equation 1.5 will
adapt it the fastest.

2 Making Competitive Learning Associative _____

It is natural to think that the "data" vectors on which a competitive network is trained play a role similar to the *input* vectors of an associative
network that maps input vectors to output vectors. This correspondence
is assumed in models that use competitive learning as a preprocessing
stage within an associative network (Moody and Darken 1989). A quite
different view is that the data vectors used in competitive learning correspond to the *output* vectors of an associative network. The competitive
network can then be viewed as an inputless stochastic generator of *output*
vectors and competitive learning can be viewed as a procedure for making the network generate output vectors with a distribution that matches
the distribution of the "data" vectors. The weight vector of each competitive hidden unit represents the mean of a multidimensional gaussian

distribution, and output vectors are generated by first picking a hidden unit and then picking an output vector from the gaussian distribution determined by the weight vector of the chosen hidden unit. The log probability of generating any particular output vector o^c is then

$$\log P^c = \log \sum_i p_i k e^{-\frac{1}{2}\|\boldsymbol{\mu}_i - o^c\|^2} \tag{2.1}$$

where i is an index over the hidden units, $\boldsymbol{\mu}_i$ is the "weight" vector of the hidden unit, k is a normalizing constant, and p_i is the probability of picking hidden unit i, so the p_i are constrained to sum to 1. In the statistics literature (McLachlan and Basford 1988), the p_i are called "mixing proportions."

"Soft" competitive learning modifies the weights (and also the variances and the mixing proportions) so as to increase the product of the probabilities (i.e., the likelihood) of generating the output vectors in the training set (Nowlan 1990a). "Hard" competitive learning is a simple approximation to soft competitive learning in which we ignore the possibility that a data vector could be generated by several different hidden units. Instead, we assume that it must be generated by the hidden unit with the closest weight vector, so only this weight vector needs to be modified to increase the probability of generating the data vector.

If we view a competitive network as generating output vectors, it is not immediately obvious what role input vectors could play. However, competitive learning can be generalized in much the same way as Barto (1985) generalized learning automata by adding an input vector and making the actions of the automaton be conditional on the input vector. We replace each hidden unit in a competitive network by an entire expert network whose output vector specifies the mean of a multidimensional gaussian distribution. So the means are now a function of the current input vector and are represented by activity levels rather than weights. In addition, we use a gating network which allows the mixing proportions of the experts to be determined by the input vector. This gives us a system of competing local experts with the error function defined in equation 1.3. We could also introduce a mechanism to allow the input vector to dynamically determine the covariance matrix for the distribution defined by each expert network, but we have not yet experimented with this possibility.

3 Application to Multispeaker Vowel Recognition

The mixture of experts model was evaluated on a speaker independent, four-class, vowel discrimination problem (Nowlan 1990b). The data consisted of the first and second formants of the vowels [i], [I], [a], and [A] (usually denoted [Λ]) from 75 speakers (males, females, and children) uttered in a **hVd** context (Peterson and Barney 1952). The data forms two
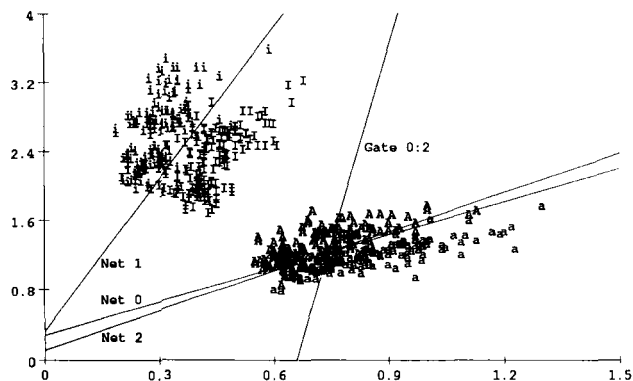
Figure 2: Data for vowel discrimination problem, and expert and gating network decision lines. The horizontal axis is the first formant value, and the vertical axis is the second formant value (the formant values have been linearly scaled by dividing by a factor of 1000). Each example is labeled with its corresponding vowel symbol. Vowels [i] and [I] form one overlapping pair of classes, vowels [a] and [A] form the other pair. The lines labeled Net 0, 1, and 2 represent the decision lines for 3 expert networks. On one side of these lines the output of the corresponding expert is less than 0.5, on the other side the output is greater than 0.5. Although the mixture in this case contained 4 experts, one of these experts made no significant contribution to the final mixture since its mixing proportion p_i was effectively 0 for all cases. The line labeled Gate 0:2 indicates the decision between expert 0 and expert 2 made by the gating network. To the left of this line $p_2 > p_0$, to the right of this line $p_0 > p_2$. The boundary between classes [a] and [A] is formed by the combination of the left part of Net 2's decision line and the right part of Net 0's decision line. Although the system tends to use as few experts as it can to solve a problem, it is also sensitive to specific problem features such as the slightly curved boundary between classes [a] and [A].

pairs of overlapping classes, and different experts learn to concentrate on one pair of classes or the other (Fig. 2).

We compared standard backpropagation networks containing a single hidden layer of 6 or 12 units with mixtures of 4 or 8 very simple experts. The architecture of each expert was restricted so it could form only a linear decision surface, which is defined as the set of input vectors for which the expert gives an output of exactly 0.5. All models were trained with data from the first 50 speakers and tested with data from the remaining 25 speakers. The small number of parameters for each expert allows excellent generalization performance (Table 1), and permits

System	Train % correct	Test % correct	Average number of epochs	SD
4 Experts	88	90	1124	23
8 Experts	88	90	1083	12
BP 6 Hid	88	90	2209	83
BP 12 Hid	88	90	2435	124

Table 1: Summary of Performance on Vowel Discrimination Task. Results are based on 25 simulations for each of the alternative models. The first column of the table indicates the system simulated. The second column gives the percent of training cases classified correctly by the final set of weights, while the third column indicates the percent of testing cases classified correctly. The last two columns contain the average number of epochs required to reach the error criterion, and the standard deviation of the distribution of convergence times. Although the squared error was used to decide when to stop training, the criterion for correct performance is based on a weighted average of the outputs of all the experts. Each expert assigns a probability distribution over the classes and these distributions are combined using proportions given by the gating network. The most probable class is then taken to be the response of the system. The identical performance of all the systems is due to the fact that, with this data set, the set of misclassified examples is not sensitive to small changes in the decision surfaces. Also, the test set is easier than the training set.

a graphic representation of the process of task decomposition (Figure 3). The number of hidden units in the backpropagation networks was chosen to give roughly equal numbers of parameters for the backpropagation networks and mixture models. All simulations were performed using a simple gradient descent algorithm with fixed step size ϵ. To simplify the comparisons, no momentum or other acceleration techniques were used. The value of ϵ for each system was chosen by performing a limited exploration of the convergence from the same initial conditions for a range of ϵ. Batch training was used with one weight update for each pass through the training set (*epoch*). Each system was trained until an average squared error of 0.08 over the training set was obtained.

The mixtures of experts reach the error criterion significantly faster than the backpropagation networks ($p \gg 0.999$), requiring only about half as many epochs on average (Table 1). The learning time for the mixture model also scales well as the number of experts is increased: The mixture of 8 experts has a small, but statistically significant ($p > 0.95$), advantage in the average number of epochs required to reach the error criterion. In contrast, the 12 hidden unit backpropagation network requires more epochs ($p > 0.95$) to reach the error criterion than the network with 6
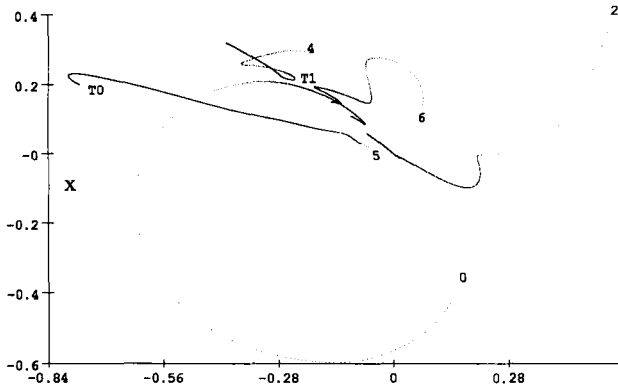
Figure 3: The trajectories of the decision lines of some experts during one simulation. The horizontal axis is the first formant value, and the vertical axis is the second formant value. Each trajectory is represented by a sequence of dots, one per epoch, each dot marking the intersection of the expert's decision line and the normal to that line passing through the origin. For clarity, only 5 of the 8 experts are shown and the number of the expert is shown at the start of the trajectory. The point labeled T0 indicates the optimal decision line for a single expert trained to discriminate [i] from [I]. Similarly, T1 represents the optimal decision line to discriminate [a] from [A]. The point labeled X is the decision line learned by a single expert trained with data from all 4 classes, and represents a type of *average* solution.

hidden units (Table 1). All statistical comparisons are based on a t test with 48 degrees of freedom and a pooled variance estimator.

Figure 3 shows how the decision lines of different experts move around as the system learns to allocate pieces of the task to different experts. The system begins in an unbiased state, with the gating network assigning equal mixing proportions to all experts in all cases. As a result, each expert tends to get errors from roughly equal numbers of cases in all 4 classes, and all experts head towards the point X, which represents the optimal decision line for an expert that must deal with all the cases. Once one or more experts begin to receive more error from cases in one class pair than the other, this symmetry is broken and the trajectories begin to diverge as different experts concentrate on one class pair or the other. In this simulation, expert 5 learns to concentrate on discriminating classes [i] and [I] so its decision line approaches the optimal line for this discrimination (T0). Experts 4 and 6 both concentrate on discriminating classes [a] and [A], so their trajectories approach the

optimal single line (T1) and then split to form a piecewise linear approx-
imation to the slightly curved optimal decision surface (see Fig. 2). Only
experts 4, 5, and 6 are active in the final mixture. This solution is typical
— in all simulations with mixtures of 4 or 8 experts all but 2 or 3 experts
had mixing proportions that were effectively 0 for all cases.

Acknowledgments

References

Barto, A. G. 1985. Learning by statistical cooperation of self-interested neuron-
    like computing elements. *Human Neurobiol.* 4, 229–256.
Hampshire, J., and Waibel, A. 1989. *The meta-pi network: Building distributed
    knowledge representations for robust pattern recognition.* Tech. Rep. CMU-CS-
    89-166, Carnegie Mellon University, Pittsburgh, PA.
Jacobs, R. A., and Jordan, M. I. 1991. Learning piecewise control strategies in a
    modular connectionist architecture, in preparation.
Jacobs, R. A., Jordan, M. I., and Barto, A. G. 1991. Task decomposition through
    competition in a modular connectionist architecture: The what and where
    vision tasks. *Cog. Sci.,* in press.
McLachlan, G. J., and Basford, K. E. 1988. *Mixture Models: Inference and Appli-
    cations to Clustering.* Marcel Dekker, New York.
Moody, J., and Darken, C. 1989. Fast learning in networks of locally-tuned
    processing units. *Neural Comp.* 1(2), 281–294.
Nowlan, S. J. 1990a. Maximum likelihood competitive learning. In *Advances
    in Neural Information Processing Systems 2,* D. S. Touretzky, ed., pp. 574–582.
    Morgan Kaufmann, San Mateo, CA.
Nowlan, S. J. 1990b. *Competing experts: An experimental investigation of associative
    mixture models.* Tech. Rep. CRG-TR-90-5, University of Toronto, Toronto,
    Canada.
Peterson, G. E., and Barney, H. L. 1952. Control methods used in a study of the
    vowels. *J. Acoust. Soc. Am.* 24, 175–184.

This article has been cited by:

1. Matthew M. Walsh, John R. Anderson. 2013. Electrophysiological Responses to Feedback during the Application of Abstract Rules. *Journal of Cognitive Neuroscience* **25**:11, 1986-2002. [Abstract] [Full Text] [PDF] [PDF Plus]

2. Bruno Damas, José Santos-Victor. 2013. Online Learning of Single- and Multivalued Functions with an Infinite Mixture of Linear Experts. *Neural Computation* **25**:11, 3044-3091. [Abstract] [Full Text] [PDF] [PDF Plus]

3. Qingguo Tang, Rohana J. Karunamuni. 2013. Minimum distance estimation in a finite mixture regression model. *Journal of Multivariate Analysis* **120**, 185-204. [CrossRef]

4. Filipe Rodrigues, Francisco Pereira, Bernardete Ribeiro. 2013. Learning from multiple annotators: Distinguishing good from random labelers. *Pattern Recognition Letters* **34**:12, 1428-1436. [CrossRef]

5. Saeed Reza Kheradpisheh, Fatemeh Behjati-Ardakani, Reza Ebrahimpour. 2013. Combining classifiers using nearest decision prototypes. *Applied Soft Computing* . [CrossRef]

6. Mansour Sheikhan, Amir Ali Sha'bani. 2013. PSO-optimized modular neural network trained by OWO-HWO algorithm for fault location in analog circuits. *Neural Computing and Applications* **23**:2, 519-530. [CrossRef]

7. Young-Seol Lee, Sung-Bae Cho. 2013. Activity recognition with android phone using mixture-of-experts co-trained with labeled and unlabeled data. *Neurocomputing* . [CrossRef]

8. Rahul Kala, Anupam Shukla, Ritu Tiwari. 2013. Breast Cancer Diagnosis Using Optimized Attribute Division in Modular Neural Networks. *Journal of Information Technology Research* **4**:1, 34-47. [CrossRef]

9. I. A. Kruglov, O. A. Mishulina. 2013. Neural network modeling of vector multivariable functions in ill-posed approximation problems. *Journal of Computer and Systems Sciences International* **52**:4, 503-518. [CrossRef]

10. Steven J. SimskeIntroduction and Overview 1-41. [CrossRef]

11. Subrajeet Mohapatra, Dipti Patra, Sanghamitra Satpathy. 2013. An ensemble classifier system for early diagnosis of acute lymphoblastic leukemia in blood microscopic images. *Neural Computing and Applications* . [CrossRef]

12. Bogusław CyganekObject Recognition 408-486. [CrossRef]

13. Heeyoul Choi, Seungjin Choi, Yoonsuck Choe. 2013. Parameter Learning for Alpha Integration. *Neural Computation* **25**:6, 1585-1604. [Abstract] [Full Text] [PDF] [PDF Plus]

14. Abbas Khalili, Shili Lin. 2013. Regularization in Finite Mixture of Regression Models with Diverging Number of Parameters. *Biometrics* **69**:2, 436-446. [CrossRef]

15. David J. Nott, Lucy Marshall, Mark Fielding, Shie-Yui Liong. 2013. Mixtures of experts for understanding model discrepancy in dynamic computer models. *Computational Statistics & Data Analysis* . [CrossRef]

16. J I Park, N Kim, M K Jeong, K S Shin. 2013. Multiphase support vector regression for function approximation with break-points. *Journal of the Operational Research Society* **64**:5, 775-785. [CrossRef]

17. Mehrdad Javadi, Seyed Ali Asghar Abbaszadeh Arani, Atena Sajedin, Reza Ebrahimpour. 2013. Classification of ECG arrhythmia by a modular neural network based on Mixture of Experts and Negatively Correlated Learning. *Biomedical Signal Processing and Control* **8**:3, 289-296. [CrossRef]

18. Davy Sannen, Hendrik Brussel. 2013. The active grading ensemble framework for learning visual quality inspection from multiple humans. *Pattern Analysis and Applications* **16**:2, 223-234. [CrossRef]

19. RASTISLAV J. R. STRUHARIK, LADISLAV A. NOVAK. 2013. HARDWARE IMPLEMENTATION OF DECISION TREE ENSEMBLES. *Journal of Circuits, Systems and Computers* 1350032. [CrossRef]

20. Michał Woźniak, Manuel Graña, Emilio Corchado. 2013. A survey of multiple classifier systems as hybrid systems. *Information Fusion* . [CrossRef]

21. Jeff A. Tracey, Jun Zhu, Erin Boydston, Lisa Lyren, Robert N. Fisher, Kevin R. Crooks. 2013. Mapping behavioral landscapes for animal movement: a finite mixture modeling approach. *Ecological Applications* **23**:3, 654-669. [CrossRef]

22. Rahul Kala, Anupam Shukla, Ritu TiwariBreast Cancer Diagnosis Using Optimized Attribute Division in Modular Neural Networks 34-47. [CrossRef]

23. P. Hall, Y. Xia, J.-H. Xue. 2013. Simple tiered classifiers. *Biometrika* . [CrossRef]

24. Tarek Helmy, S. M. Rahman, Muhammad Imtiaz Hossain, Abdulaziz Abdelraheem. 2013. Non-linear Heterogeneous Ensemble Model for Permeability Prediction of Oil Reservoirs. *Arabian Journal for Science and Engineering* . [CrossRef]

25. Mehmet Gönen, Ethem Alpaydın. 2013. Localized algorithms for multiple kernel learning. *Pattern Recognition* **46**:3, 795-807. [CrossRef]

26. Andriy Norets, Satoru Takahashi. 2013. On the surjectivity of the mapping between utilities and choice probabilities. *Quantitative Economics* **4**:1, 149-155. [CrossRef]

27. Konrad Jackowski. 2013. Fixed-size ensemble classifier system evolutionarily adapted to a recurring context with an unlimited pool of classifiers. *Pattern Analysis and Applications* . [CrossRef]

28. Haibo He, Yuan Cao, Yi Cao, Jinyu Wen. 2013. Ensemble learning for wind profile prediction with missing values. *Neural Computing and Applications* **22**:2, 287-294. [CrossRef]

29. Zhigang Liu, Wenfan Li, Wanlu Sun. 2013. A novel method of short-term load forecasting based on multiwavelet transform and multiple neural networks. *Neural Computing and Applications* **22**:2, 271-277. [CrossRef]

30. Yanan Fan, David J. Nott, Scott A. Sisson. 2013. Approximate Bayesian computation via regression density estimation. *Stat* **2**:1, 34-48. [CrossRef]

31. Siew Li Tan, David J. Nott. 2013. Variational approximation for mixtures of linear mixed models. *Journal of Computational and Graphical Statistics* 130123061318003. [CrossRef]

32. Reza Ebrahimpour, Seyed Ali Asghar Abbaszadeh Arani, Saeed Masoudnia. 2013. Improving combination method of NCL experts using gating network. *Neural Computing and Applications* **22**:1, 95-101. [CrossRef]

33. Rodrigo Arnaldo Scarpel. 2013. Forecasting air passengers at São Paulo International Airport using a mixture of local experts model. *Journal of Air Transport Management* **26**, 35-39. [CrossRef]

34. Adam Lammert, Louis Goldstein, Shrikanth Narayanan, Khalil Iskarous. 2013. Statistical methods for estimation of direct and differential kinematics of the vocal tract. *Speech Communication* **55**:1, 147-161. [CrossRef]

35. Sungmin Myoung. 2013. Modified Mixture of Experts for the Diagnosis of Perfusion Magnetic Resonance Imaging Measures in Locally Rectal Cancer Patients. *Healthcare Informatics Research* **19**:2, 130. [CrossRef]

36. Eri Itoh, Shinji Suzuki. 2013. Evaluation on novel architecture for harmonizing manual and automatic flight controls under atmospheric turbulence. *Aerospace Science and Technology* **24**:1, 241-254. [CrossRef]

37. Li-Jie Zhao, Tian-You Chai, De-Cheng Yuan. 2012. Selective ensemble extreme learning machine modeling of effluent quality in wastewater treatment plants. *International Journal of Automation and Computing* **9**:6, 627-633. [CrossRef]

38. Mattias Villani, Robert Kohn, David J. Nott. 2012. Generalized smooth finite mixtures. *Journal of Econometrics* **171**:2, 121-133. [CrossRef]

39. Michael K. Pitt, Ralph dos Santos Silva, Paolo Giordani, Robert Kohn. 2012. On some properties of Markov chain Monte Carlo simulation methods based on the particle filter. *Journal of Econometrics* **171**:2, 134-151. [CrossRef]

40. Patrick P.K. Chan, Daniel S. Yeung, Wing W.Y. Ng, Chih Min Lin, James N.K. Liu. 2012. Dynamic fusion method using Localized Generalization Error Model. *Information Sciences* **217**, 1-20. [CrossRef]

41. Eduardo F. Mendes, Wenxin Jiang. 2012. On Convergence Rates of Mixtures of Polynomial Experts. *Neural Computation* **24**:11, 3025-3051. [Abstract] [Full Text] [PDF] [PDF Plus]

42. Saeed Masoudnia, Reza Ebrahimpour, Seyed Ali Asghar Abbaszadeh Arani. 2012. Combining features of negative correlation learning with mixture of experts in proposed ensemble methods. *Applied Soft Computing* **12**:11, 3539-3551. [CrossRef]

43. Rodrigo C. Barros, Márcio P. Basgalupp, André C. P. L. F. Carvalho, Marcos G. Quiles. 2012. Clus-DTI: improving decision-tree classification with a clustering-based decision-tree induction algorithm. *Journal of the Brazilian Computer Society* **18**:4, 351-362. [CrossRef]

44. Tim Shallice, Richard P. Cooper. 2012. The Organisation of Mind: Response to commentators. *Cortex* **48**:10, 1383-1387. [CrossRef]

45. Jae Joon Ahn, Young Min Kim, Keunje Yoo, Joonhong Park, Kyong Joo Oh. 2012. Using GA-Ridge regression to select hydro-geological parameters influencing groundwater pollution vulnerability. *Environmental Monitoring and Assessment* **184**:11, 6637-6645. [CrossRef]

46. Marek Kurzynski, Michal Wozniak. 2012. Combining classifiers under probabilistic models: experimental comparative analysis of methods. *Expert Systems* **29**:4, 374-393. [CrossRef]

47. J. F. De Paz, M. Navarro, C. I. Pinzon, V. Julian, D. I. Tapia, J. Bajo. 2012. Mathematical model for a temporal-bounded classifier in security environments. *Logic Journal of IGPL* **20**:4, 712-721. [CrossRef]

48. Saeed Masoudnia, Reza Ebrahimpour, Seyed Ali Asghar Abbaszadeh Arani. 2012. Incorporation of a Regularization Term to Control Negative Correlation in Mixture of Experts. *Neural Processing Letters* **36**:1, 31-47. [CrossRef]

49. Hamid Shayegh Boroujeni, Nasrollah Moghadam Charkari. 2012. Robust moving shadow detection with hierarchical mixture of MLP experts. *Signal, Image and Video Processing* . [CrossRef]

50. Reza Ebrahimpour, Naser Sadeghnejad, Atena Sajedin, Nima Mohammadi. 2012. Electrocardiogram beat classification via coupled boosting by filtering and preloaded mixture of experts. *Neural Computing and Applications* . [CrossRef]

51. David J. Nott, Siew Li Tan, Mattias Villani, Robert Kohn. 2012. Regression Density Estimation With Variational Methods and Stochastic Approximation. *Journal of Computational and Graphical Statistics* **21**:3, 797-820. [CrossRef]

52. B. VermaNeural Network Based Classifier Ensembles 229-239. [CrossRef]

53. CHUANYU SUN, XIAO-LIN WU, KENT A. WEIGEL, GUILHERME J. M. ROSA, STEWART BAUCK, BRENT W. WOODWARD, ROBERT D. SCHNABEL, JEREMY F. TAYLOR, DANIEL GIANOLA. 2012. An ensemble-based approach to imputation of moderate-density genotypes for genomic selection with application to Angus cattle. *Genetics Research* **94**:03, 133-150. [CrossRef]

54. Andriy Norets, Justinas Pelenis. 2012. Bayesian modeling of joint and conditional distributions. *Journal of Econometrics* **168**:2, 332-346. [CrossRef]

55. Kristine Monteith, Tony Martinez. 2012. AGGREGATE CERTAINTY ESTIMATORS. *Computational Intelligence* no-no. [CrossRef]

56. Saeed Masoudnia, Reza Ebrahimpour. 2012. Mixture of experts: a literature survey. *Artificial Intelligence Review* . [CrossRef]

57. Pejman Tahmasebi, Ardeshir Hezarkhani. 2012. A fast and independent architecture of artificial neural network for permeability prediction. *Journal of Petroleum Science and Engineering* **86-87**, 118-126. [CrossRef]

58. Maya Wardeh, Frans Coenen, Trevor Bench Capon. 2012. PISA: A framework for multiagent classification using argumentation. *Data & Knowledge Engineering* **75**, 34-57. [CrossRef]

59. Reza Ebrahimpour, Naser Sadeghnejad, Seyed Ali Asghar Abbaszadeh Arani, Nima Mohammadi. 2012. Boost-wise pre-loaded mixture of experts for classification tasks. *Neural Computing and Applications* . [CrossRef]

60. MATTEO RE, GIORGIO VALENTINIEnsemble Methods **20124949**, . [CrossRef]

61. M. J. Frank, D. Badre. 2012. Mechanisms of Hierarchical Reinforcement Learning in Corticostriatal Circuits 1: Computational Analysis. *Cerebral Cortex* **22**:3, 509-526. [CrossRef]

62. Charles Kemp, Patrick Shafto, Joshua B. Tenenbaum. 2012. An integrated account of generalization across objects and features. *Cognitive Psychology* **64**:1-2, 35-73. [CrossRef]

63. Ricardo Ñanculef, Carlos Valle, Héctor Allende, Claudio Moraga. 2012. Training Regression Ensembles by Sequential Target Correction and Resampling. *Information Sciences* . [CrossRef]

64. Nima Hatami. 2012. Thinned-ECOC ensemble based on sequential code shrinking. *Expert Systems with Applications* **39**:1, 936-947. [CrossRef]

65. Xiuqin Bai, Weixin Yao, John E. Boyer. 2012. Robust fitting of mixture regression models. *Computational Statistics & Data Analysis* . [CrossRef]

66. Souhaib Ben Taieb, Gianluca Bontempi, Amir F. Atiya, Antti Sorjamaa. 2012. A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition. *Expert Systems with Applications* . [CrossRef]

67. Sungmin Myoung, Ji Hong Chang, Kijun Song. 2012. A Mixture of Experts Model for the Diagnosis of Liver Cirrhosis by Measuring the Liver Stiffness. *Healthcare Informatics Research* **18**:1, 29. [CrossRef]

68. P. Giordani, X. Mun, R. Kohn. 2012. Efficient Estimation of Covariance Matrices using Posterior Mode Multiple Shrinkage. *Journal of Financial Econometrics* **11**:1, 154-192. [CrossRef]

69. Scott Doyle, Michael D Feldman, Natalie Shih, John Tomaszewski, Anant Madabhushi. 2012. Cascaded discrimination of normal, abnormal, and confounder classes in histopathology: Gleason grading of prostate cancer. *BMC Bioinformatics* **13**:1, 282. [CrossRef]

70. Anil K. Ghosh, Fred Godtliebsen. 2011. On hybrid classification using model assisted posterior estimates. *Pattern Recognition* . [CrossRef]

71. Jong-Won Yoon, Sung-Ihk Yang, Sung-Bae Cho. 2011. Adaptive mixture-of-experts models for data glove interface with multiple users. *Expert Systems with Applications* . [CrossRef]

72. Andrew R. Webb, Keith D. CopseyReferences 591-636. [CrossRef]

73. Tomasz Woloszynski, Marek Kurzynski. 2011. A probabilistic model of classifier competence for dynamic ensemble selection. *Pattern Recognition* **44**:10-11, 2656-2668. [CrossRef]

74. Ahmad Banakar, Mohammad Fazle Azeem. 2011. Local recurrent sigmoidal–wavelet neurons in feed-forward neural network for forecasting of dynamic systems: Theory. *Applied Soft Computing* . [CrossRef]

75. Cristian I. Pinzón, Juan F. De Paz, Martí Navarro, Javier Bajo, Vicente Julián, Juan. M. Corchado. 2011. Real-time CBR-agent with a mixture of experts in the reuse stage to classify and detect DoS attacks. *Applied Soft Computing* **11**:7, 4384-4398. [CrossRef]

76. SULTAN UDDIN AHMED, MD. SHAHJAHAN, KAZUYUKI MURASE. 2011. A LEMPEL-ZIV COMPLEXITY-BASED NEURAL NETWORK PRUNING ALGORITHM. *International Journal of Neural Systems* **21**:05, 427-441. [CrossRef]

77. Yunwu Xiong, Rony Wallach, Alex Furman. 2011. Modeling multidimensional flow in wettable and water-repellent soils using artificial neural networks. *Journal of Hydrology* . [CrossRef]

78. Yan Yang, Jinwen Ma. 2011. Asymptotic Convergence Properties of the EM Algorithm for Mixture of Experts. *Neural Computation* **23**:8, 2140-2168. [Abstract] [Full Text] [PDF] [PDF Plus] [Supplementary Content]

79. Iffat A. Gheyas, Leslie S. Smith. 2011. A novel neural network ensemble architecture for time series forecasting. *Neurocomputing* . [CrossRef]

80. Shuhaida Ismail, Ani Shabri, Ruhaidah Samsudin. 2011. A hybrid model of self-organizing maps (SOM) and least square support vector machine (LSSVM) for time-series forecasting. *Expert Systems with Applications* **38**:8, 10574-10578. [CrossRef]

81. Lei XuLearning Algorithms for RBF Functions and Subspace Based Functions 1034-1065. [CrossRef]

82. Lei Xu, Shun-ichi AmariCombining Classifiers and Learning Mixture-of-Experts 243-252. [CrossRef]

83. Pramod P. Nair. 2011. A multigradient algorithm using a mixture of experts architecture for land cover classification of multisensor images. *International Journal of Remote Sensing* 1-9. [CrossRef]

84. Bruno Baruque, Santiago Porras, Emilio Corchado. 2011. Hybrid Classification Ensemble Using Topology-preserving Clustering. *New Generation Computing* **29**:3, 329-344. [CrossRef]

85. Ivette Luna, Rosangela Ballini. 2011. Top-down strategies based on adaptive fuzzy rule-based systems for daily time series forecasting. *International Journal of Forecasting* **27**:3, 708-724. [CrossRef]

86. Sotirios P. Chatzis, Yiannis Demiris. 2011. The copula echo state network. *Pattern Recognition* . [CrossRef]

87. Reza Ebrahimpour, Hossein Nikoo, Saeed Masoudnia, Mohammad Reza Yousefi, Mohammad Sajjad Ghaemi. 2011. Mixture of MLP-experts for trend forecasting of time series: A case study of the Tehran stock exchange. *International Journal of Forecasting* **27**:3, 804-816. [CrossRef]

88. Mehdi Salkhordeh Haghighi, Abedin Vahedian, Hadi Sadoghi Yazdi. 2011. Extended decision template presentation for combining classifiers. *Expert Systems with Applications* **38**:7, 8414-8418. [CrossRef]

89. Dong-Hun Seo, Won-Don Lee. 2011. A New Ensemble System using Dynamic Weighting Method. *The Journal of the Korean Institute of Information and Communication Engineering* **15**:6, 1213-1220. [CrossRef]

90. Haibo HeEnsemble Learning 108-139. [CrossRef]

91. David Martínez-Rego, Oscar Fontenla-Romero, Amparo Alonso-Betanzos. 2011. Efficiency of local models ensembles for time series prediction. *Expert Systems with Applications* **38**:6, 6884-6894. [CrossRef]

92. Shiliang Sun, Xin Xu. 2011. Variational Inference for Infinite Mixtures of Gaussian Processes With Applications to Traffic Flow Prediction. *IEEE Transactions on Intelligent Transportation Systems* **12**:2, 466-475. [CrossRef]

93. Martina Sattlecker, Rebecca Baker, Nick Stone, Conrad Bessant. 2011. Support vector machine ensembles for breast cancer type prediction from mid-FTIR micro-calcification spectra. *Chemometrics and Intelligent Laboratory Systems* . [CrossRef]

94. David P. Williams. 2011. Label Alteration to Improve Underwater Mine Classification. *IEEE Geoscience and Remote Sensing Letters* **8**:3, 488-492. [CrossRef]

95. Mahdi Tabassian, Reza Ghaderi, Reza Ebrahimpour. 2011. Knitted fabric defect classification for uncertain labels based on Dempster–Shafer theory of evidence. *Expert Systems with Applications* **38**:5, 5259-5267. [CrossRef]

96. Isobel Claire Gormley, Thomas Brendan MurphyMixture of Experts Modelling with Social Science Applications 101-121. [CrossRef]

97. Feng Li, Mattias Villani, Robert KohnModelling Conditional Densities Using Finite Smooth Mixtures 123-144. [CrossRef]

98. Duy Nguyen-Tuong, Jan Peters. 2011. Model learning for robot control: a survey. *Cognitive Processing* . [CrossRef]

99. Xu Zhao, Yun Fu, Yuncai Liu. 2011. Human Motion Tracking by Temporal-Spatial Local Gaussian Process Experts. *IEEE Transactions on Image Processing* **20**:4, 1141-1151. [CrossRef]

100. Chen Wu. 2011. Deriving collective intelligence from reviews on the social Web using a supervised learning approach. *Expert Systems with Applications* . [CrossRef]

101. Parisa Rashidi, Diane J. Cook. 2011. Activity knowledge transfer in smart environments. *Pervasive and Mobile Computing* . [CrossRef]

102. Rainer Schlittgen. 2011. A weighted least-squares approach to clusterwise regression. *AStA Advances in Statistical Analysis* . [CrossRef]

103. Lei Xu, Yanda Li. 2011. Machine learning and intelligence science: Sino-foreign interchange workshop IScIDE2010 (A). *Frontiers of Electrical and Electronic Engineering in China* **6**:1, 1-5. [CrossRef]

104. Y. Sun, M.B. Brown, M. Prapopoulou, N. Davey, R.G. Adams, G.P. Moss. 2011. The application of stochastic machine learning methods in the prediction of skin penetration. *Applied Soft Computing* **11**:2, 2367-2375. [CrossRef]

105. Davy Sannen, Hendrik Van Brussel. 2011. A Multilevel Information Fusion Approach for Visual Quality Inspection. *Information Fusion* . [CrossRef]

106. Abdelhamid Bouchachia. 2011. Incremental learning with multi-level adaptation. *Neurocomputing* . [CrossRef]

107. Sally Wood. 2011. Bayesian Mixtures of Autoregressive Models. *Journal of Computational and Graphical Statistics* 1-22. [CrossRef]

108. M. Olteanu, J. Rynkiewicz. 2011. Asymptotic properties of mixture-of-experts models. *Neurocomputing* . [CrossRef]

109. A. Wefky, F. Espinosa, A. Prieto, J.J. Garcia, C. Barrios. 2011. Comparison of neural classifiers for vehicles gear estimation. *Applied Soft Computing* . [CrossRef]

110. Phaedon-Stelios Koutsourelakis, Elias Bilionis. 2011. Scalable Bayesian Reduced-Order Models for Simulating High-Dimensional Multiscale Dynamical Systems. *Multiscale Modeling & Simulation* **9**:1, 449-485. [CrossRef]

111. Georgios Kalantzis, Luis A. Vasquez-Quino, Travis Zalman, Guillem Pratx, Yu Lei. 2011. Toward IMRT 2D dose modeling using artificial neural networks: A feasibility study. *Medical Physics* **38**:10, 5807. [CrossRef]

112. Esma Kilic, Ethem Alpaydin. 2011. Learning the areas of expertise of classifiers in an ensemble. *Procedia Computer Science* **3**, 74-82. [CrossRef]

113. Ahmad Banakar. 2011. Lyapunov Stability Analysis of Gradient Descent-Learning Algorithm in Network Training. *ISRN Applied Mathematics* **2011**, 1-12. [CrossRef]

114. R. A. Mat Noor, Z. Ahmad, M. Mat Don, M. H. Uzir. 2010. Modelling and control of different types of polymerization processes using neural networks technique: A review. *The Canadian Journal of Chemical Engineering* **88**:6, 1065-1084. [CrossRef]

115. Zhouyu Fu, A Robles-Kelly, Jun Zhou. 2010. Mixing Linear SVMs for Nonlinear Classification. *IEEE Transactions on Neural Networks* **21**:12, 1963-1975. [CrossRef]

116. Abbas Khalili. 2010. New estimation and feature selection methods in mixture-of-experts models. *Canadian Journal of Statistics* **38**:4, 519-539. [CrossRef]

117. Salah Bouktif, Faheem Ahmed, Issa Khalil, Giuliano Antoniol. 2010. A novel composite model approach to improve software quality prediction. *Information and Software Technology* **52**:12, 1298-1311. [CrossRef]

118. Chrisantha Fernando, Richard Goldstein, Eörs Szathmáry. 2010. The Neuronal Replicator Hypothesis. *Neural Computation* **22**:11, 2809-2857. [Abstract] [Full Text] [PDF] [PDF Plus] [Supplementary Content]

119. Luís M. Silva, J. Marques de Sá, Luís A. Alexandre. 2010. The MEE Principle in Data Classification: A Perceptron-Based Analysis. *Neural Computation* **22**:10, 2698-2728. [Abstract] [Full Text] [PDF] [PDF Plus]

120. D.S. Young, D.R. Hunter. 2010. Mixtures of regressions with predictor-dependent mixing proportions. *Computational Statistics & Data Analysis* **54**:10, 2253-2266. [CrossRef]

121. José Luis Aznarte, José Manuel Benitez. 2010. Equivalences Between Neural-Autoregressive Time Series Models and Fuzzy Systems. *IEEE Transactions on Neural Networks* **21**:9, 1434-1444. [CrossRef]

122. Francis Colas, Julien Diard, Pierre Bessière. 2010. Common Bayesian Models for Common Cognitive Issues. *Acta Biotheoretica* **58**:2-3, 191-216. [CrossRef]

123. Ladan Shams, Ulrik R. Beierholm. 2010. Causal inference in perception. *Trends in Cognitive Sciences* **14**:9, 425-432. [CrossRef]

124. Alexandre Carvalho, Georgios Skoulakis. 2010. Time Series Mixtures of Generalized t Experts: ML Estimation and an Application to Stock Return Density Forecasting. *Econometric Reviews* **29**:5, 642-687. [CrossRef]

125. Michael C. Lee, Lilla Boroczky, Kivilcim Sungur-Stasik, Aaron D. Cann, Alain C. Borczuk, Steven M. Kawut, Charles A. Powell. 2010. Computer-aided diagnosis of pulmonary nodules using a two-step approach for feature selection and classifier ensemble construction. *Artificial Intelligence in Medicine* **50**:1, 43-53. [CrossRef]

126. Yasutake Takahashi, Yoshihiro Tamura, Minoru Asada, Mario Negrello. 2010. Emulation and behavior understanding through shared values. *Robotics and Autonomous Systems* **58**:7, 855-865. [CrossRef]

127. Min Liu, Mingyu Dong, Cheng Wu. 2010. A New ANFIS for Parameter Prediction With Numeric and Categorical Inputs. *IEEE Transactions on Automation Science and Engineering* **7**:3, 645-653. [CrossRef]

128. Jun Namikawa, Jun Tani. 2010. Learning to imitate stochastic time series in a compositional way by chaos. *Neural Networks* **23**:5, 625-638. [CrossRef]

129. Seyoung Kim, Padhraic Smyth, Hal Stern. 2010. A Bayesian Mixture Approach to Modeling Spatial Activation Patterns in Multisite fMRI Data. *IEEE Transactions on Medical Imaging* **29**:6, 1260-1274. [CrossRef]

130. Salim Chitroub. 2010. Classifier combination and score level fusion: concepts and practical aspects. *International Journal of Image and Data Fusion* **1**:2, 113-135. [CrossRef]

131. Elif Derya Übeyli, Konuralp Ilbay, Gul Ilbay, Deniz Sahin, Gur Akansel. 2010. Differentiation of Two Subtypes of Adult Hydrocephalus by Mixture of Experts. *Journal of Medical Systems* **34**:3, 281-290. [CrossRef]

132. K. A. Le Cao, E. Meugnier, G. J. McLachlan. 2010. Integrative mixture of experts to combine clinical factors and gene markers. *Bioinformatics* **26**:9, 1192-1198. [CrossRef]

133. Jean-Pierre Stockis, Jürgen Franke, Joseph Tadjuidje Kamgaing. 2010. On geometric ergodicity of CHARME models. *Journal of Time Series Analysis* **31**:3, 141-152. [CrossRef]

134. I. Zaier, C. Shu, T.B.M.J. Ouarda, O. Seidou, F. Chebana. 2010. Estimation of ice thickness on lakes using artificial neural network ensembles. *Journal of Hydrology* **383**:3-4, 330-340. [CrossRef]

135. J. McLean Sloughter. 2010. Probabilistic Wind Speed Forecasting Using Ensembles and Bayesian Model Averaging. *Journal of the American Statistical Association* **105**:489, 25-35. [CrossRef]

136. József Fiser, Pietro Berkes, Gergő Orbán, Máté Lengyel. 2010. Statistically optimal perception and learning: from behavior to neural representations. *Trends in Cognitive Sciences* **14**:3, 119-130. [CrossRef]

137. Chulsang Yoo, Jooyoung Park. 2010. A mixture-density-network based approach for finding rating curves: Facing multi-modality and unbalanced data distribution. *KSCE Journal of Civil Engineering* **14**:2, 243-250. [CrossRef]

138. Shile Zhang, Bin Li, Xiangyang Xue. 2010. Semi-automatic dynamic auxiliary-tag-aided image annotation#. *Pattern Recognition* **43**:2, 470-477. [CrossRef]

139. Rajib Nayak, James Gomes. 2010. Generalized hybrid control synthesis for affine systems using sequential adaptive networks. *Journal of Chemical Technology & Biotechnology* **85**:1, 59-76. [CrossRef]

140. Isaac Martín Diego, Alberto Muñoz, Javier M. Moguerza. 2010. Methods for the combination of kernel matrices within a support vector framework. *Machine Learning* **78**:1-2, 137-174. [CrossRef]

141. Petr Kadlec, Bogdan Gabrys. 2009. Architecture for development of adaptive on-line prediction models. *Memetic Computing* **1**:4, 241-269. [CrossRef]

142. Mattias Villani, Robert Kohn, Paolo Giordani. 2009. Regression density estimation using smooth adaptive Gaussian mixtures. *Journal of Econometrics* **153**:2, 155-173. [CrossRef]

143. Reza Ebrahimpour, Ehsanollah Kabir, Mohammad Reza Yousefi. 2009. Improving mixture of experts for view-independent face recognition using teacher-directed learning. *Machine Vision and Applications* . [CrossRef]

144. Kyuwan Choi, Hideaki Hirose, Yoshio Sakurai, Toshio Iijima, Yasuharu Koike. 2009. Prediction of arm trajectory from the neural activities of the primary motor cortex with modular connectionist architecture. *Neural Networks* **22**:9, 1214-1223. [CrossRef]

145. J. Bolton, P. Gader. 2009. Random Set Framework for Context-Based Classification With Hyperspectral Imagery. *IEEE Transactions on Geoscience and Remote Sensing* **47**:11, 3810-3821. [CrossRef]

146. RASTISLAV J. R. STRUHARIK, LADISLAV A. NOVAK. 2009. EVOLVING DECISION TREES IN HARDWARE. *Journal of Circuits, Systems and Computers* **18**:06, 1033-1060. [CrossRef]

147. Eri Itoh, Shinji Suzuki. 2009. Architecture for Harmonizing Manual and Automatic Flight Controls. *Journal of Aerospace Computing, Information, and Communication* **6**:10, 553-569. [CrossRef]

148. Christian Plagemann, Sebastian Mischke, Sam Prentice, Kristian Kersting, Nicholas Roy, Wolfram Burgard. 2009. A Bayesian regression approach to terrain mapping and an application to legged robot locomotion. *Journal of Field Robotics* **26**:10, 789-811. [CrossRef]

149. Clodoaldo A. M. Lima, André L. V. Coelho, Fernando J. Zuben. 2009. Pattern classification with mixtures of weighted least-squares support vector machine experts. *Neural Computing and Applications* **18**:7, 843-860. [CrossRef]

150. Hsiu-Ting Yu. 2009. S. FRÜHWIRTH-SCHNATTER (2006) Finite Mixture and Markov Switching Models. *Psychometrika* **74**:3, 559-560. [CrossRef]

151. Paolo Soda, Giulio Iannello, Mario Vento. 2009. A multiple expert system for classifying fluorescent intensity in antinuclear autoantibodies analysis. *Pattern Analysis and Applications* **12**:3, 215-226. [CrossRef]

152. Enso Ikonen, Kaddour Najim. 2009. Multiple Model-Based Control Using Finite Controlled Markov Chains. *Cognitive Computation* **1**:3, 234-243. [CrossRef]

153. Elif Derya Übeyli. 2009. Modified mixture of experts employing eigenvector methods and Lyapunov exponents for analysis of electroencephalogram signals. *Expert Systems* **26**:4, 339-354. [CrossRef]

154. Elif Derya Übeyli. 2009. Modified Mixture of Experts for Diabetes Diagnosis. *Journal of Medical Systems* **33**:4, 299-305. [CrossRef]

155. Andrea L. Gebhart, Richard N. Aslin, Elissa L. Newport. 2009. Changing Structures in Midstream: Learning Along the Statistical Garden Path. *Cognitive Science* **33**:6, 1087-1116. [CrossRef]

156. Adam Tashman, Robert Frey. 2009. Modeling risk in arbitrage strategies using finite mixtures. *Quantitative Finance* **9**:5, 495-503. [CrossRef]

157. Giovanni Pezzulo, Gianguglielmo Calvi. 2009. Computational explorations of perceptual symbol systems theory. *New Ideas in Psychology* . [CrossRef]

158. E.D. Ubeyli. 2009. Eigenvector Methods for Automated Detection of Electrocardiographic Changes in Partial Epileptic Patients. *IEEE Transactions on Information Technology in Biomedicine* **13**:4, 478-485. [CrossRef]

159. Zainal Ahmad, Rabiatul ʾAdawiah Mat Noor, Jie Zhang. 2009. Multiple neural networks modeling techniques in process control: a review. *Asia-Pacific Journal of Chemical Engineering* **4**:4, 403-419. [CrossRef]

160. J. Carreau, Y. Bengio. 2009. A Hybrid Pareto Mixture for Conditional Asymmetric Fat-Tailed Distributions. *IEEE Transactions on Neural Networks* **20**:7, 1087-1101. [CrossRef]

161. Hiroshi Imamizu, Mitsuo Kawato. 2009. Brain mechanisms for predictive control by switching internal models: implications for higher-order cognitive functions. *Psychological Research Psychologische Forschung* **73**:4, 527-544. [CrossRef]

162. Sheng-Hsun Hsu, JJ Po-An Hsieh, Ting-Chih Chih, Kuei-Chu Hsu. 2009. A two-stage architecture for stock price forecasting by integrating self-organizing map and support vector regression. *Expert Systems with Applications* **36**:4, 7947-7951. [CrossRef]

163. Aydın Ulaş, Murat Semerci, Olcay Taner Yıldız, Ethem Alpaydın. 2009. Incremental construction of classifier and discriminant ensembles. *Information Sciences* **179**:9, 1298-1318. [CrossRef]

164. Yaoyao Zhu, Xiaolei Huang, Wei Wang, Daniel Lopresti, Rodney Long, Sameer Antani, Zhiyun Xue, George Thoma. 2009. Balancing the Role of Priors in Multi-Observer Segmentation Evaluation. *Journal of Signal Processing Systems* **55**:1-3, 185-207. [CrossRef]

165. J. Molina-Vilaplana, J. L. Contreras-Vidal, M. T. Herrero-Ezquerro, J. Lopez-Coronado. 2009. A model for altered neural network dynamics related to prehension movements in Parkinson disease. *Biological Cybernetics* **100**:4, 271-287. [CrossRef]

166. Manuel Carcenac. 2009. A modular neural network for super-resolution of human faces. *Applied Intelligence* **30**:2, 168-186. [CrossRef]

167. Elif Derya Übeyli. 2009. Features for analysis of electrocardiographic changes in partial epileptic patients. *Expert Systems with Applications* **36**:3, 6780-6789. [CrossRef]

168. N. Gradojevic, R. Gencay, D. Kukolj. 2009. Option Pricing With Modular Neural Networks. *IEEE Transactions on Neural Networks* **20**:4, 626-637. [CrossRef]

169. P. Rojanavasu, Hai Huong Dam, H.A. Abbass, C. Lokan, O. Pinngern. 2009. A Self-Organized, Distributed, and Adaptive Rule-Based Induction System. *IEEE Transactions on Neural Networks* **20**:3, 446-459. [CrossRef]

170. Kai A. Krueger, Peter Dayan. 2009. Flexible shaping: How learning in small steps helps. *Cognition* **110**:3, 380-394. [CrossRef]

171. Ori Rosen, David S. Stoffer, Sally Wood. 2009. Local Spectral Analysis via a Bayesian Mixture of Smoothing Splines. *Journal of the American Statistical Association* **104**:485, 249-262. [CrossRef]

172. S FROLOV, A BAPTISTA, Y ZHANG, C SEATON. 2009. Estimation of ecologically significant circulation features of the Columbia River estuary and plume using a reduced-dimension Kalman filter. *Continental Shelf Research* **29**:2, 456-466. [CrossRef]

173. Michael J. Procopio, Jane Mulligan, Greg Grudic. 2009. Learning terrain segmentation with classifier ensembles for autonomous robot navigation in unstructured environments. *Journal of Field Robotics* **26**:2, 145-175. [CrossRef]

174. Serkan Tapkın, Ozdemir Akyılmaz. 2009. A new approach to neural trip distribution models: NETDIM. *Transportation Planning and Technology* **32**:1, 93-114. [CrossRef]

175. Mike Oaksford, Nick Chater. 2009. The uncertain reasoner: Bayes, logic, and rationality. *Behavioral and Brain Sciences* **32**:01, 105-120. [CrossRef]

176. M.D. Muhlbaier, A. Topalis, R. Polikar. 2009. Learn^{++} .NC: Combining Ensemble of Classifiers With Dynamically Weighted Consult-and-Vote for Efficient Incremental Learning of New Classes. *IEEE Transactions on Neural Networks* **20**:1, 152-168. [CrossRef]

177. Sergios Theodoridis, Konstantinos KoutroumbasNonlinear Classifiers 151-260. [CrossRef]

178. S MABU, K SHIMADA, K HIRASAWA, J HU. 2009. Study of multi-branch structure of Universal Learning Networks. *Applied Soft Computing* **9**:1, 393-403. [CrossRef]

179. Xin Li, Yunfei Zheng. 2009. Patch-Based Video Processing: A Variational Bayesian Approach. *IEEE Transactions on Circuits and Systems for Video Technology* **19**:1, 27-40. [CrossRef]

180. J NAMIKAWA, J TANI. 2008. A model for learning to segment temporal sequences, utilizing a mixture of RNN experts together with adaptive variance. *Neural Networks* **21**:10, 1466-1475. [CrossRef]

181. COLIN FYFE, WESAM BARBAKH, WEI CHUAN OOI, HANSEOK KO. 2008. TOPOLOGICAL MAPPINGS OF VIDEO AND AUDIO DATA. *International Journal of Neural Systems* **18**:06, 481-489. [CrossRef]

182. Fernanda L. Minku, Teresa B. Ludermir. 2008. Clustering and co-evolution to construct neural network ensembles: An experimental study. *Neural Networks* **21**:9, 1363-1379. [CrossRef]

183. Luís M. Silva, J. Marques de Sá, Luís A. Alexandre. 2008. Data classification with multilayer perceptrons using a generalized error function. *Neural Networks* **21**:9, 1302-1310. [CrossRef]

184. H. Cevikalp, R. Polikar. 2008. Local Classifier Weighting by Quadratic Programming. *IEEE Transactions on Neural Networks* **19**:10, 1832-1838. [CrossRef]

185. Pietro Zito, Haibo Chen, Margaret C. Bell. 2008. Predicting Real-Time Roadside CO and $\hbox{NO}_{2}$ Concentrations Using Neural Networks. *IEEE Transactions on Intelligent Transportation Systems* **9**:3, 514-522. [CrossRef]

186. M.A. Wiering, H. van Hasselt. 2008. Ensemble Algorithms in Reinforcement Learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* **38**:4, 930-936. [CrossRef]

187. KOSTAS FRAGOS, SPIROS PANETSOS. 2008. DISAMBIGUATION OF GREEK POLYSEMOUS WORDS USING HIERARCHICAL PROBABILISTIC NETWORKS AND A CHI-SQUARE FEATURE SELECTION STRATEGY. *International Journal on Artificial Intelligence Tools* **17**:04, 687-701. [CrossRef]

188. Reza Ebrahimpour, Ehsanollah Kabir, Mohammad Reza Yousefi. 2008. Teacher-directed learning in view-independent face recognition with mixture of experts using overlapping eigenspaces. *Computer Vision and Image Understanding* **111**:2, 195-206. [CrossRef]

189. Dongbing Gu. 2008. Distributed EM Algorithm for Gaussian Mixtures in Sensor Networks. *IEEE Transactions on Neural Networks* **19**:7, 1154-1166. [CrossRef]

190. M.M. Islam, Xin Yao, S.M. Shahriar Nirjon, M.A. Islam, K. Murase. 2008. Bagging and Boosting Negatively Correlated Neural Networks. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* **38**:3, 771-784. [CrossRef]

191. Sally A Wood, Robert Kohn, Remy Cottet, Wenxin Jiang, Martin Tanner. 2008. Locally Adaptive Nonparametric Binary Regression. *Journal of Computational and Graphical Statistics* **17**:2, 352-372. [CrossRef]

192. Giovanni Pezzulo. 2008. Coordinating with the Future: The Anticipatory Nature of Representation. *Minds and Machines* **18**:2, 179-225. [CrossRef]

193. Elif Derya Übeyli. 2008. Implementing wavelet transform/mixture of experts network for analysis of electrocardiogram beats. *Expert Systems* **25**:2, 150-162. [CrossRef]

194. E UBEYLI. 2008. Wavelet/mixture of experts network structure for EEG signals classification. *Expert Systems with Applications* **34**:3, 1954-1962. [CrossRef]

195. Leo Galway, Darryl Charles, Michaela Black. 2008. Machine learning in digital games: a survey. *Artificial Intelligence Review* **29**:2, 123-161. [CrossRef]

196. J. Peres, R. Oliveira, S. Feyo de Azevedo. 2008. Bioprocess hybrid parametric/nonparametric modelling based on the concept of mixture of experts. *Biochemical Engineering Journal* **39**:1, 190-206. [CrossRef]

197. Jun Tani, Ryu Nishimoto, Jun Namikawa, Masato Ito. 2008. Codevelopmental Learning Between Human and Humanoid Robot Using a Dynamic Neural-

Network Model. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* **38**:1, 43-59. [CrossRef]

198. Berk Gokberk, Helin Dutagaci, Ayd¿n Ulas, Lale Akarun, B#lent Sankur. 2008. Representation Plurality and Fusion for 3-D Face Recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* **38**:1, 155-173. [CrossRef]

199. Shalom Darmanjian, Jose C. Principe. 2008. Boosted and Linked Mixtures of HMMs for Brain-Machine Interfaces. *EURASIP Journal on Advances in Signal Processing* **2008**, 1-13. [CrossRef]

200. R POLIKAR, A TOPALIS, D PARIKH, D GREEN, J FRYMIARE, J KOUNIOS, C CLARK. 2008. An ensemble based data fusion approach for early diagnosis of Alzheimer's disease. *Information Fusion* **9**:1, 83-95. [CrossRef]

201. T TANIGUCHI, T SAWARAGI. 2008. Incremental acquisition of multiple nonlinear forward models based on differentiation process of schema model. *Neural Networks* **21**:1, 13-27. [CrossRef]

202. Yong Liu,, Xin Yao,. 2008. Nature Inspired Neural Network Ensemble Learning. *Journal of Intelligent Systems* **17**:supplement, 5. [CrossRef]

203. Minh Ha Nguyen, H.A. Abbass, R.I. McKay. 2008. Analysis of CCME: Coevolutionary Dynamics, Automatic Problem Decomposition, and Regularization. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* **38**:1, 100-109. [CrossRef]

204. Sébastien Hélie, Gyslain Giguère, Denis Cousineau, Robert Proulx. 2007. Using knowledge partitioning to investigate the psychological plausibility of mixtures of experts. *Artificial Intelligence Review* **25**:1-2, 119-138. [CrossRef]

205. Spyros N. Raptis, Elpida S. Tzafestas. 2007. Robot inverse kinematics via neural and neurofuzzy networks: architectural and computational aspects for improved performance. *Journal of Information and Optimization Sciences* **28**:6, 905-933. [CrossRef]

206. Sheng-Uei Guan, Chunyu Bao, TseNgee Neo. 2007. Reduced Pattern Training Based on Task Decomposition Using Pattern Distributor. *IEEE Transactions on Neural Networks* **18**:6, 1738-1749. [CrossRef]

207. Manuel Carcenac. 2007. A modular neural network applied to image transformation and mental images. *Neural Computing and Applications* . [CrossRef]

208. Mark Eastwood, Bogdan Gabrys. 2007. The Dynamics of Negative Correlation Learning. *The Journal of VLSI Signal Processing Systems for Signal, Image, and Video Technology* **49**:2, 251-263. [CrossRef]

209. Shun-ichi Amari. 2007. Integration of Stochastic Models by Minimizing α-Divergence. *Neural Computation* **19**:10, 2780-2796. [Abstract] [PDF] [PDF Plus]

210. J SU. 2007. Camera calibration based on receptive fields. *Pattern Recognition* **40**:10, 2837-2845. [CrossRef]

211. Mitsuo Kawato, Daniel WolpertInternal Models for Motor Control 291-307. [CrossRef]

212. M. Aziz Muslim, Masumi Ishikawa, Tetsuo Furukawa. 2007. Task segmentation in a mobile robot by mnSOM: a new approach to training expert modules. *Neural Computing and Applications* **16**:6, 571-580. [CrossRef]

213. Zhuoxin Sun, Ori Rosen, Allan R. Sampson. 2007. Multivariate Bernoulli Mixture Models with Application to Postmortem Tissue Studies in Schizophrenia. *Biometrics* **63**:3, 901-909. [CrossRef]

214. L XU. 2007. A unified perspective and new results on RHT computing, mixture based learning, and multi-learner based problem solving#. *Pattern Recognition* **40**:8, 2129-2153. [CrossRef]

215. Reza Ebrahimpour, Ehsanollah Kabir, Mohammad Reza Yousefi. 2007. Face Detection Using Mixture of MLP Experts. *Neural Processing Letters* **26**:1, 69-82. [CrossRef]

216. Hiroshi Imamizu, Norikazu Sugimoto, Rieko Osu, Kiyoka Tsutsui, Kouichi Sugiyama, Yasuhiro Wada, Mitsuo Kawato. 2007. Explicit contextual information selectively contributes to predictive switching of internal models. *Experimental Brain Research* **181**:3, 395-408. [CrossRef]

217. Mark D. Skowronski, John G. Harris. 2007. Noise-Robust Automatic Speech Recognition Using a Predictive Echo State Network. *IEEE Transactions on Audio, Speech and Language Processing* **15**:5, 1724-1730. [CrossRef]

218. D. M. Gavrila, S. Munder. 2007. Multi-cue Pedestrian Detection and Tracking from a Moving Vehicle. *International Journal of Computer Vision* **73**:1, 41-59. [CrossRef]

219. C LIMA, A COELHO, F VONZUBEN. 2007. Hybridizing mixtures of experts with support vector machines: Investigation into nonlinear dynamic systems identification. *Information Sciences* **177**:10, 2049-2074. [CrossRef]

220. Abdelhamid Bouchachia. 2007. Learning with partly labeled data. *Neural Computing and Applications* **16**:3, 267-293. [CrossRef]

221. Giovanna Jona Lasinio, Fabio Divino, Annibale Biggeri. 2007. Environmental risk assessment in the Tuscany region: a proposal. *Environmetrics* **18**:3, 315-332. [CrossRef]

222. Yung-Keun Kwon, Byung-Ro Moon. 2007. A Hybrid Neurogenetic Approach for Stock Forecasting. *IEEE Transactions on Neural Networks* **18**:3, 851-864. [CrossRef]

223. R VANDERMERWE, T LEEN, Z LU, S FROLOV, A BAPTISTA. 2007. Fast neural network surrogates for very high dimensional physics-based models in computational oceanography. *Neural Networks* **20**:4, 462-478. [CrossRef]

224. Yasutoshi Nomura, Hitoshi Furuta, Michiyuki Hirokane. 2007. An Integrated Fuzzy Control System for Structural Vibration. *Computer-Aided Civil and Infrastructure Engineering* **22**:4, 306-316. [CrossRef]

225. M SKOWRONSKI, J HARRIS. 2007. Automatic speech recognition using a predictive echo state network classifier. *Neural Networks* **20**:3, 414-423. [CrossRef]

226. Ludmila Kuncheva, Juan Rodriguez. 2007. Classifier Ensembles with a Random Linear Oracle. *IEEE Transactions on Knowledge and Data Engineering* **19**:4, 500-508. [CrossRef]

227. Devi Parikh, Robi Polikar. 2007. An Ensemble-Based Incremental Learning Approach to Data Fusion. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)* **37**:2, 437-450. [CrossRef]

228. Colin Fyfe. 2007. Two topographic maps for data visualisation. *Data Mining and Knowledge Discovery* **14**:2, 207-224. [CrossRef]

229. David J. Miller, Siddharth Pal. 2007. Transductive Methods for the Distributed Ensemble Classification Problem. *Neural Computation* **19**:3, 856-884. [Abstract] [PDF] [PDF Plus]

230. Mark S. Handcock, Adrian E. Raftery, Jeremy M. Tantrum. 2007. Model-based clustering for social networks. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **170**:2, 301-354. [CrossRef]

231. Anelia Angelova, Larry Matthies, Daniel Helmick, Pietro Perona. 2007. Learning and prediction of slip from visual information. *Journal of Field Robotics* **24**:3, 205-231. [CrossRef]

232. Elif Derya Übeyli. 2007. Comparison of different classification algorithms in clinical decision-making. *Expert Systems* **24**:1, 17-31. [CrossRef]

233. Xia Hong, Sheng Chen, Chris J. Harris. 2007. A Kernel-Based Two-Class Classifier for Imbalanced Data Sets. *IEEE Transactions on Neural Networks* **18**:1, 28-41. [CrossRef]

234. P HARTONO, S HASHIMOTO. 2007. Learning from imperfect data. *Applied Soft Computing* **7**:1, 353-363. [CrossRef]

235. Justin C. Sanchez, José C. Principe. 2007. Brain–Machine Interface Engineering. *Synthesis Lectures on Biomedical Engineering* **2**:1, 1-234. [CrossRef]

236. Sheng-Uei Guan,, Kai Wang,. 2007. Hierarchical Incremental Class Learning with Output Parallelism. *Journal of Intelligent Systems* **16**:2, 167. [CrossRef]

237. Nicolas Amaury, Emilio Benfenati, Severin Bumbaru, Antonio Chana, Marian Craciun, Jacques R. Chrétien, Giuseppina Gini, Gongde Guo, Frank Lemke, Viorel Minzu, Johann-Adolf Müller, Daniel Neagu, Marco Pintore, Silviu Augustin Stroia, Paul TrundleHybrid systems 149-183. [CrossRef]

238. Junfeng Pan, Qiang Yang, Yiming Yang, Lei Li, Frances Li, George Li. 2007. Cost-Sensitive-Data Preprocessing for Mining Customer Relationship Management Databases. *IEEE Intelligent Systems* **22**:1, 46-51. [CrossRef]

239. Kumiko Nishi, Ichiro Takeuchi. 2007. Casualty Insurance Pure Premium Estimation Using Two-Stage Regression Tree. *Transactions of the Japanese Society for Artificial Intelligence* **22**, 183-190. [CrossRef]

240. Matthew C. Casey, Khurshid Ahmad. 2006. A competitive neural model of small number detection. *Neural Networks* **19**:10, 1475-1489. [CrossRef]

241. Sheng-Uei Guan, Chunyu Bao, Ru-Tian Sun. 2006. Hierarchical Incremental Class Learning with Reduced Pattern Training. *Neural Processing Letters* **24**:2, 163-177. [CrossRef]

242. I. Guler, E.D. Ubeyli. 2006. Automated Diagnostic Systems With Diverse and Composite Features for Doppler Ultrasound Signals. *IEEE Transactions on Biomedical Engineering* **53**:10, 1934-1942. [CrossRef]

243. Y OHIGASHI, T OMORI. 2006. Modeling of autonomous problem solving process by dynamic construction of task models in multiple tasks environment. *Neural Networks* **19**:8, 1169-1180. [CrossRef]

244. Kazuyuki Samejima, Ken'Ichi Katagiri, Kenji Doya, Mitsuo Kawato. 2006. Multiple model-based reinforcement learning for nonlinear control. *Electronics and Communications in Japan (Part III: Fundamental Electronic Science)* **89**:9, 54-69. [CrossRef]

245. J. Verbeek. 2006. Learning nonlinear image manifolds by global alignment of local linear models. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **28**:8, 1236-1250. [CrossRef]

246. A SHARMA, K PALIWAL, G ONWUBOLU. 2006. Class-dependent PCA, MDC and LDA: A combined classifier for pattern classification. *Pattern Recognition* **39**:7, 1215-1229. [CrossRef]

247. G BARRETO, L SOUZA. 2006. Adaptive filtering with the self-organizing map: A performance comparison. *Neural Networks* **19**:6-7, 785-798. [CrossRef]

248. I AUTIO. 2006. Using natural class hierarchies in multi-class visual classification. *Pattern Recognition* **39**:7, 1290-1299. [CrossRef]

249. G.-B. Huang, L. Chen, C.-K. Siew. 2006. Universal Approximation Using Incremental Constructive Feedforward Networks With Random Hidden Nodes. *IEEE Transactions on Neural Networks* **17**:4, 879-892. [CrossRef]

250. W WANG, P GELDER, J VRIJLING, J MA. 2006. Forecasting daily streamflow using hybrid ANN models. *Journal of Hydrology* **324**:1-4, 383-399. [CrossRef]

251. S-P Kim, J C Sanchez, Y N Rao, D Erdogmus, J M Carmena, M A Lebedev, M A L Nicolelis, J C Principe. 2006. A comparison of optimal MIMO linear and nonlinear models for brain–machine interfaces. *Journal of Neural Engineering* **3**:2, 145-161. [CrossRef]

252. V.R. Khare, X. Yao, B. Sendhoff. 2006. Multi-network evolutionary systems and automatic decomposition of complex problems. *International Journal of General Systems* **35**:3, 259-274. [CrossRef]

253. Michael Rimer, Tony Martinez. 2006. Classification-based objective functions. *Machine Learning* **63**:2, 183-205. [CrossRef]

254. Robi PolikarPattern Recognition . [CrossRef]

255. Mingyang Xu, Michael W. Golay. 2006. Data-guided model combination by decomposition and aggregation. *Machine Learning* **63**:1, 43-67. [CrossRef]

256. J. Zhang, Q. Jin, Y. Xu. 2006. Inferential Estimation of Polymer Melt Index Using Sequentially Trained Bootstrap Aggregated Neural Networks. *Chemical Engineering & Technology* **29**:4, 442-448. [CrossRef]

257. Erhan Oztop, Mitsuo Kawato, Michael Arbib. 2006. Mirror neurons and imitation: A computationally guided review. *Neural Networks* **19**:3, 254-271. [CrossRef]

258. D LOYOLAR. 2006. Applications of neural network methods to the processing of earth observation satellite data. *Neural Networks* **19**:2, 168-177. [CrossRef]

259. D SOLOMATINE, M SIEK. 2006. Modular learning models in forecasting natural phenomena. *Neural Networks* **19**:2, 215-224. [CrossRef]

260. J. Cho, J.C. Principe, D. Erdogmus, M.A. Motter. 2006. Modeling and Inverse Controller Design for an Unmanned Aerial Vehicle Based on the Self-Organizing Map. *IEEE Transactions on Neural Networks* **17**:2, 445-460. [CrossRef]

261. J ZHANG. 2006. Improved on-line process fault diagnosis through information fusion in multiple neural networks. *Computers & Chemical Engineering* **30**:3, 558-571. [CrossRef]

262. Marcus FreanConnectionist Architectures: Optimization . [CrossRef]

263. Yang Ge, Wenxin Jiang. 2006. On Consistency of Bayesian Inference with Mixtures of Logistic Regression. *Neural Computation* **18**:1, 224-243. [Abstract] [PDF] [PDF Plus]

264. Xin Yao, Yong Xu. 2006. Recent Advances in Evolutionary Computation. *Journal of Computer Science and Technology* **21**:1, 1-18. [CrossRef]

265. Alexandre X. Carvalho, Martin A. Tanner. 2006. Modeling nonlinearities with mixtures-of-experts of time series models. *International Journal of Mathematics and Mathematical Sciences* **2006**, 1-23. [CrossRef]

266. Giuliano Armano, Alessandro Orro, Eloisa Vargiu. 2006. MASSP3: A System for Predicting Protein Secondary Structure. *EURASIP Journal on Advances in Signal Processing* **2006**, 1-10. [CrossRef]

267. J. Peres, F. Freitas, MAM Reis, S. Feyo de Azevedo, R. OliveiraHybrid modular mechanistic/ANN modelling of a wastewater phosphorus removal process **21**, 1717-1722. [CrossRef]

268. Tadahiro TANIGUCHI, Tetsuo SAWARAGI. 2006. Construction of Behavioral Concepts through Social Interactions based on Reward Design: Schema-Based Incremental Reinforcement Learning. *Journal of Japan Society for Fuzzy Theory and Intelligent Informatics* **18**:4, 629-640. [CrossRef]

269. Steven Walczak, Madhavan Parthasarathy. 2006. Modeling online service discontinuation with nonparametric agents. *Information Systems and e-Business Management* **4**:1, 49-70. [CrossRef]

270. Sergios Theodoridis, Konstantinos KoutroumbasNonlinear Classifiers 121-211. [CrossRef]

271. Biswanath BhattacharyaReferences . [CrossRef]

272. Z AHMAD, J ZHANG. 2005. Combination of multiple neural networks using data fusion techniques for enhanced nonlinear process modelling. *Computers & Chemical Engineering* **30**:2, 295-308. [CrossRef]

273. S SCHAAL, N SCHWEIGHOFER. 2005. Computational motor control in humans and robots. *Current Opinion in Neurobiology* **15**:6, 675-682. [CrossRef]

274. J. Arenas-Garcia, V. Gomez-Verdejo, A.R. Figueiras-Vidal. 2005. New Algorithms for Improved Adaptive Convex Combination of LMS Transversal Filters. *IEEE Transactions on Instrumentation and Measurement* **54**:6, 2239-2249. [CrossRef]

275. G.-B. Huang, K.Z. Mao, C.-K. Siew, D.-S. Huang. 2005. Fast Modular Network Implementation for Support Vector Machines. *IEEE Transactions on Neural Networks* **16**:6, 1651-1663. [CrossRef]

276. Geoffrey HintonArtificial Intelligence: Neural Networks . [CrossRef]

277. Elif Derya Übeyli. 2005. A Mixture of Experts Network Structure for Breast Cancer Diagnosis. *Journal of Medical Systems* **29**:5, 569-579. [CrossRef]

278. Peter D Neilson, Megan D Neilson. 2005. An overview of adaptive model theory: solving the problems of redundancy, resources, and nonlinear interactions in human movement control. *Journal of Neural Engineering* **2**:3, S279-S312. [CrossRef]

279. G. Armano, L. Milanesi, A. Orro. 2005. Multiple Alignment Through Protein Secondary-Structure Information. *IEEE Transactions on Nanobioscience* **4**:3, 207-211. [CrossRef]

280. DAVID WINDRIDGE, JOSEF KITTLER. 2005. PERFORMANCE MEASURES OF THE TOMOGRAPHIC CLASSIFIER FUSION METHODOLOGY. *International Journal of Pattern Recognition and Artificial Intelligence* **19**:06, 731-753. [CrossRef]

281. K. Chen. 2005. On the Use of Different Speech Representations for Speaker Modeling. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)* **35**:3, 301-314. [CrossRef]

282. J.I. Arribas, J. Cid-Sueiro. 2005. A Model Selection Algorithm for a Posteriori Probability Estimation With Neural Networks. *IEEE Transactions on Neural Networks* **16**:4, 799-809. [CrossRef]

283. I. Mora-Jimenez, J. Cid-Sueiro. 2005. A Universal Learning Rule That Minimizes Well-Formed Cost Functions. *IEEE Transactions on Neural Networks* **16**:4, 810-820. [CrossRef]

284. Tugba Taskaya-Temizel, Matthew C. Casey. 2005. A comparative study of autoregressive neural network hybrids. *Neural Networks* **18**:5-6, 781-789. [CrossRef]

285. John A. Drakopoulos, Ahmad Abdulkader. 2005. Training neural networks with heterogeneous data. *Neural Networks* **18**:5-6, 595-601. [CrossRef]

286. Takumi Ichimura, Shinichi Oeda, Machi Suka, Katsumi Yoshida. 2005. A learning method of immune multi-agent neural networks. *Neural Computing and Applications* **14**:2, 132-148. [CrossRef]

287. Masahiko Morita, Kouhei Matsuzawa, Shigemitsu Morokami. 2005. A model of context-dependent association using selective desensitization of nonmonotonic neural elements. *Systems and Computers in Japan* **36**:7, 73-83. [CrossRef]

288. Abedalrazq Khalil, Mohammad N. Almasri, Mac McKee, Jagath J. Kaluarachchi. 2005. Applicability of statistical learning algorithms in groundwater quality modeling. *Water Resources Research* **41**:5, n/a-n/a. [CrossRef]

289. BRENT FERGUSON, RANADHIR GHOSH, JOHN YEARWOOD. 2005. MODULAR NEURAL NETWORK DESIGN FOR THE PROBLEM OF ALPHABETIC CHARACTER RECOGNITION. *International Journal of Pattern Recognition and Artificial Intelligence* **19**:02, 249-269. [CrossRef]

290. Alexandre X. Carvalho, Martin A. Tanner. 2005. Modeling nonlinear time series with local mixtures of generalized linear models. *Canadian Journal of Statistics* **33**:1, 97-113. [CrossRef]

291. G ARMANO, M MARCHESI, A MURRU. 2005. A hybrid genetic-neural architecture for stock indexes forecasting. *Information Sciences* **170**:1, 3-33. [CrossRef]

292. A HAGEN. 2005. Recent advances in the multi-stream HMM/ANN hybrid approach to noise robust ASR. *Computer Speech & Language* **19**:1, 3-30. [CrossRef]

293. Seiji Ishihara, Harukazu Igarashi. 2005. A Task Decomposition Algorithm Based on the Distribution of Input Pattern Vectors for Classification Problems. *IEEJ Transactions on Electronics, Information and Systems* **125**:7, 1043-1048. [CrossRef]

294. A.X. Carvalho, M.A. Tanner. 2005. Mixtures-of-Experts of Autoregressive Time Series: Asymptotic Normality and Model Specification. *IEEE Transactions on Neural Networks* **16**:1, 39-56. [CrossRef]

295. David WindridgeMorphologically Debiased Classifier Fusion: A Tomography-Theoretic Approach **134**, 181-266. [CrossRef]

296. J. Lan, J. Cho, D. Erdogmus, J.C. Principe, M.A. Motter, J. Xu. 2005. Local Linear PID Controllers for Nonlinear Control. *Control and Intelligent Systems* **33**:1. . [CrossRef]

297. R PAINE, J TANI. 2004. Motor primitive and sequence self-organization in a hierarchical recurrent neural network. *Neural Networks* **17**:8-9, 1291-1309. [CrossRef]

298. C.K. Loo, M. Rajeswari, M.V.C. Rao. 2004. Novel Direct and Self-Regulating Approaches to Determine Optimum Growing Multi-Experts Network Structure. *IEEE Transactions on Neural Networks* **15**:6, 1378-1395. [CrossRef]

299. Roelof K. Brouwer. 2004. A hybrid neural network for input that is both categorical and quantitative. *International Journal of Intelligent Systems* **19**:10, 979-1001. [CrossRef]

300. A.S. Cofino, J.M. Gutierrez, M.L. Ivanissevich. 2004. Evolving modular networks with genetic algorithms: application to nonlinear time series. *Expert Systems* **21**:4, 208-216. [CrossRef]

301. SHIMON COHEN, NATHAN INTRATOR. 2004. ON DIFFERENT MODEL SELECTION CRITERIA IN A FORWARD AND BACKWARD REGRESSION HYBRID NETWORK. *International Journal of Pattern Recognition and Artificial Intelligence* **18**:05, 847-865. [CrossRef]

302. CHRISTOPH KÖNIG, GIUSEPPINA GINI, MARIAN CRACIUN, EMILIO BENFENATI. 2004. MULTICLASS CLASSIFIER FROM A COMBINATION OF LOCAL EXPERTS: TOWARD DISTRIBUTED COMPUTATION FOR REAL-PROBLEM CLASSIFIERS. *International Journal of Pattern Recognition and Artificial Intelligence* **18**:05, 801-817. [CrossRef]

303. Jayanta Basak, Ravi Kothari. 2004. A Classification Paradigm for Distributed Vertically Partitioned Data. *Neural Computation* **16**:7, 1525-1544. [Abstract] [PDF] [PDF Plus]

304. F.A. Mussa-Ivaldi, S.A. Solla. 2004. Neural Primitives for Motion Control. *IEEE Journal of Oceanic Engineering* **29**:3, 640-650. [CrossRef]

305. L. Xu. 2004. Advances on BYY Harmony Learning: Information Theoretic Perspective, Generalized Projection Geometry, and Independent Factor Autodetermination. *IEEE Transactions on Neural Networks* **15**:4, 885-902. [CrossRef]

306. S.K. Warfield, K.H. Zou, W.M. Wells. 2004. Simultaneous Truth and Performance Level Estimation (STAPLE): An Algorithm for the Validation of Image Segmentation. *IEEE Transactions on Medical Imaging* **23**:7, 903-921. [CrossRef]

307. M.-W. Chang, C.-J. Lin, R.C.-H. Weng. 2004. Analysis of Switching Dynamics With Competing Support Vector Machines. *IEEE Transactions on Neural Networks* **15**:3, 720-727. [CrossRef]

308. J Brown. 2004. How laminar frontal cortex and basal ganglia circuits interact to control planned and reactive saccades. *Neural Networks* **17**:4, 471-510. [CrossRef]

309. S.-K. Ng, G.J. McLachlan. 2004. Using the EM Algorithm to Train Neural Networks: Misconceptions and a New Algorithm for Multiclass Classification. *IEEE Transactions on Neural Networks* **15**:3, 738-749. [CrossRef]

310. M.A. Moussa. 2004. Combining Expert Neural Networks Using Reinforcement Feedback for Learning Primitive Grasping Behavior. *IEEE Transactions on Neural Networks* **15**:3, 629-638. [CrossRef]

311. WEI HUANG, K. K. LAI, Y. NAKAMORI, SHOUYANG WANG. 2004. FORECASTING FOREIGN EXCHANGE RATES WITH ARTIFICIAL NEURAL NETWORKS: A REVIEW. *International Journal of Information Technology & Decision Making* **03**:01, 145-165. [CrossRef]

312. J. Su, J. Wang, Y. Xi. 2004. Incremental Learning With Balanced Update on Receptive Fields for Multi-Sensor Data Fusion. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)* **34**:1, 659-665. [CrossRef]

313. Hyun-Chul Kim, Daijin Kim, Sung Yang Bang, Sang-Youn Lee. 2004. Face recognition using the second-order mixture-of-eigenfaces method. *Pattern Recognition* **37**:2, 337-349. [CrossRef]

314. Andrew Estabrooks, Taeho Jo, Nathalie Japkowicz. 2004. A Multiple Resampling Method for Learning from Imbalanced Data Sets. *Computational Intelligence* **20**:1, 18-36. [CrossRef]

315. Sheng-Uei Guan,, TseNgee Neo,, Chunyu Bao,. 2004. Task Decomposition Using Pattern Distributor. *Journal of Intelligent Systems* **13**:2, 123. [CrossRef]

316. J. Peres, R. Oliveira, L.S. Serafim, P. Lemos, M.A. Reis, S. Feyo de AzevedoHybrid Modelling of a PHA Production Process Using Modular Neural Networks **18**, 733-738. [CrossRef]

317. Michael L. Kalish, Stephan Lewandowsky, John K. Kruschke. 2004. Population of Linear Experts: Knowledge Partitioning and Function Learning. *Psychological Review* **111**:4, 1072-1099. [CrossRef]

318. Lee-Xieng Yang, Stephan Lewandowsky. 2004. Knowledge Partitioning in Categorization: Constraints on Exemplar Models. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **30**:5, 1045-1064. [CrossRef]

319. Sheng-Uei Guan,, Jun Liu,, Yinan Qi,. 2004. An Incremental Approach to Contribution-Based Feature Selection. *Journal of Intelligent Systems* **13**:1, 15. [CrossRef]

320. Sheng-Uei Guan,, Jun Liu,. 2004. Incremental Neural Network Training with an Increasing Input Dimension. *Journal of Intelligent Systems* **13**:1, 43. [CrossRef]

321. Kai Huang, Robert F. Murphy. 2004. From quantitative microscopy to automated image understanding. *Journal of Biomedical Optics* **9**:5, 893. [CrossRef]

322. Lewis Bott, Evan Heit. 2004. Nonmonotonic Extrapolation in Function Learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **30**:1, 38-50. [CrossRef]

323. Ori Rosen, Ayala Cohen. 2003. Analysis of growth curves via mixtures. *Statistics in Medicine* **22**:23, 3641-3654. [CrossRef]

324. L Chu Kiong. 2003. Nonlinear dynamic system identification and control via constructivism inspired neural network. *Applied Soft Computing* **3**:3, 237-257. [CrossRef]

325. C.M. Bachmann, M.H. Bettenhausen, R.A. Fusina, T.F. Donato, A.L. Russ, J.W. Burke, G.M. Lamela, W.J. Rhea, B.R. Truitt, J.H. Porter. 2003. A credit assignment approach to fusing classifiers of multiseason hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing* **41**:11, 2488-2499. [CrossRef]

326. D.J. Miller, J. Browning. 2003. A mixture model and EM-based algorithm for class discovery, robust classification, and outlier rejection in mixed labeled/unlabeled data sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **25**:11, 1468-1483. [CrossRef]

327. L Chu Kiong. 2003. Extrapolation detection and novelty-based node insertion for sequential growing multi-experts network. *Applied Soft Computing* **3**:2, 159-175. [CrossRef]

328. K Samejima. 2003. Inter-module credit assignment in modular reinforcement learning. *Neural Networks* **16**:7, 985-994. [CrossRef]

329. S Mounce. 2003. Sensor-fusion of hydraulic data for burst detection and location in a treated water distribution system. *Information Fusion* **4**:3, 217-229. [CrossRef]

330. Jun Tani, M. Ito. 2003. Self-organization of behavioral primitives as multiple attractor dynamics: A robot experiment. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* **33**:4, 481-488. [CrossRef]

331. R Brouwer. 2003. Training a feed-forward network with incomplete data due to missing input variables. *Applied Soft Computing* **3**:1, 23-36. [CrossRef]

332. Md.M. Islam, Xin Yao, K. Murase. 2003. A constructive algorithm for training cooperative neural network ensembles. *IEEE Transactions on Neural Networks* **14**:4, 820-834. [CrossRef]

333. H KIM. 2003. An efficient model order selection for PCA mixture model. *Pattern Recognition Letters* **24**:9-10, 1385-1393. [CrossRef]

334. Chee Peng Lim, R.F. Harrison. 2003. Online pattern classification with multiple neural network systems: an experimental study. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)* **33**:2, 235-247. [CrossRef]

335. RONAN COLLOBERT, YOSHUA BENGIO, SAMY BENGIO. 2003. SCALING LARGE LEARNING PROBLEMS WITH HARD PARALLEL MIXTURES. *International Journal of Pattern Recognition and Artificial Intelligence* **17**:03, 349-365. [CrossRef]

336. H Kim. 2003. Extensions of LDA by PCA mixture model and class-wise features. *Pattern Recognition* **36**:5, 1095-1105. [CrossRef]

337. N. Garcia-Pedrajas, C. Hervas-Martinez, J. Munoz-Perez. 2003. COVNET: a cooperative coevolutionary model for evolving artificial neural networks. *IEEE Transactions on Neural Networks* **14**:3, 575-596. [CrossRef]

338. Liu Yong, Zou Xiu-fen. 2003. Analysis of negative correlation learning. *Wuhan University Journal of Natural Sciences* **8**:1, 165-175. [CrossRef]

339. Liu Yong, Zou Xiu-fen. 2003. From designing a single neural network to designing neural network ensembles. *Wuhan University Journal of Natural Sciences* **8**:1, 155-164. [CrossRef]

340. D. Windridge, J. Kittler. 2003. A morphologically optimal strategy for classifier combinaton: multiple expert fusion as a tomographic process. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **25**:3, 343-353. [CrossRef]

341. Shimon Edelman, Nathan Intrator. 2003. Towards structural systematicity in distributed, statically bound visual representations. *Cognitive Science* **27**:1, 73-109. [CrossRef]

342. J Tani. 2003. Learning to generate articulated behavior through the bottom-up and the top-down interaction processes. *Neural Networks* **16**:1, 11-23. [CrossRef]

343. S Cohen. 2002. Automatic model selection in a hybrid perceptron/radial network. *Information Fusion* **3**:4, 259-266. [CrossRef]

344. M.A.L. Thathachar, P.S. Sastry. 2002. Varieties of learning automata: an overview. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)* **32**:6, 711-722. [CrossRef]

345. M Su. 2002. A hybrid learning system for image deblurring. *Pattern Recognition* **35**:12, 2881-2894. [CrossRef]

346. Andrew D. Back, Tianping Chen. 2002. Universal Approximation of Multiple Nonlinear Operators by Neural Networks. *Neural Computation* **14**:11, 2561-2566. [Abstract] [PDF] [PDF Plus]

347. A. Kehagias, V. Petridis. 2002. Predictive modular neural networks for unsupervised segmentation of switching time series: the data allocation problem. *IEEE Transactions on Neural Networks* **13**:6, 1432-1449. [CrossRef]

348. L XU. 2002. BYY harmony learning, structural RPCL, and topological self-organizing on mixture models1. *Neural Networks* **15**:8-9, 1125-1151. [CrossRef]

349. J. SVENSSON. 2002. AN MLP TRAINING ALGORITHM TAKING INTO ACCOUNT KNOWN ERRORS ON INPUTS AND OUTPUTS. *International Journal of Neural Systems* **12**:05, 369-379. [CrossRef]

350. Michalis K. Titsias, Aristidis Likas. 2002. Mixture of Experts Classification Using a Hierarchical Mixture Model. *Neural Computation* **14**:9, 2221-2244. [Abstract] [PDF] [PDF Plus]

351. R Brouwer. 2002. A feed-forward network for input that is both categorical and quantitative. *Neural Networks* **15**:7, 881-890. [CrossRef]

352. GIULIANO ARMANO, ANDREA MURRU, FABIO ROLI. 2002. STOCK MARKET PREDICTION BY A MIXTURE OF GENETIC-NEURAL EXPERTS. *International Journal of Pattern Recognition and Artificial Intelligence* **16**:05, 501-526. [CrossRef]

353. Lizhong Wu, S.L. Oviatt, P.R. Cohen. 2002. From members to teams to committee-a robust approach to gestural and multimodal recognition. *IEEE Transactions on Neural Networks* **13**:4, 972-982. [CrossRef]

354. H ElKadi. 2002. Energy-based fatigue life prediction of fiberglass/epoxy composites using modular neural networks. *Composite Structures* **57**:1-4, 85-89. [CrossRef]

355. E. Mizutani, K. Nishio. 2002. Multi-illuminant color reproduction for electronic cameras via CANFIS neuro-fuzzy modular network device characterization. *IEEE Transactions on Neural Networks* **13**:4, 1009-1022. [CrossRef]

356. Peter DayanReinforcement Learning . [CrossRef]

357. PITOYO HARTONO, SHUJI HASHIMOTO. 2002. EXTRACTING THE PRINCIPAL BEHAVIOR OF A PROBABILISTIC SUPERVISOR THROUGH NEURAL NETWORKS ENSEMBLE. *International Journal of Neural Systems* **12**:03n04, 291-301. [CrossRef]

358. Purushottam Papatla, Mariam (Fatemeh) Zahedi, Marijana Zekic-Susac. 2002. Leveraging the Strengths of Choice Models and Neural Networks: A Multiproduct Comparative Analysis. *Decision Sciences* **33**:3, 433-461. [CrossRef]

359. A.M. Martinez. 2002. Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**:6, 748-763. [CrossRef]

360. Ronan Collobert, Samy Bengio, Yoshua Bengio. 2002. A Parallel Mixture of SVMs for Very Large Scale Problems. *Neural Computation* **14**:5, 1105-1114. [Abstract] [PDF] [PDF Plus]

361. Z Zhou. 2002. Ensembling neural networks: Many could be better than all. *Artificial Intelligence* **137**:1-2, 239-263. [CrossRef]

362. Sheng-Uei Guan, Shanchun Li. 2002. Parallel growing and training of neural networks using output parallelism. *IEEE Transactions on Neural Networks* **13**:3, 542-550. [CrossRef]

363. J Dong. 2002. Local learning framework for handwritten character recognition. *Engineering Applications of Artificial Intelligence* **15**:2, 151-159. [CrossRef]

364. L.I. Kuncheva. 2002. Switching between selection and fusion in combining classifiers: an experiment. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)* **32**:2, 146-156. [CrossRef]

365. E. Moreau, C. Mallet, S. Thiria, B. Mabboux, F. Badran, C. Klapisz. 2002. Atmospheric Liquid Water Retrieval Using a Gated Experts Neural Network. *Journal of Atmospheric and Oceanic Technology* **19**:4, 457-467. [CrossRef]

366. C.C. Chibelushi, F. Deravi, J.S.D. Mason. 2002. A review of speech-based bimodal recognition. *IEEE Transactions on Multimedia* **4**:1, 23-37. [CrossRef]

367. Michael A. Erickson, John K. Kruschke. 2002. Rule-based extrapolation in perceptual categorization. *Psychonomic Bulletin & Review* **9**:1, 160-168. [CrossRef]

368. G. Mayraz, G.E. Hinton. 2002. Recognizing handwritten digits using hierarchical products of experts. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**:2, 189-197. [CrossRef]

369. Sheng-Uei Guan,, Jun Liu,. 2002. Incremental Ordered Neural Network Training. *Journal of Intelligent Systems* **12**:3, 137. [CrossRef]

370. David Mitchell, Robert Pavur. 2002. Using modular neural networks for business decisions. *Management Decision* **40**:1, 58-63. [CrossRef]

371. Sham Kakade, Peter Dayan. 2002. Acquisition and extinction in autoshaping. *Psychological Review* **109**:3, 533-544. [CrossRef]

372. Sheng-Uei Guan,, Peng Li,. 2002. Feature Selection for Modular Neural Network Classifiers. *Journal of Intelligent Systems* **12**:3, 173. [CrossRef]

373. Sun-Gi Hong, Sang-Keon Oh, Min-Soeng Kim, Ju-Jang Lee. 2002. Evolving mixture of experts for nonlinear time series modelling and prediction. *Electronics Letters* **38**:1, 34. [CrossRef]

374. Stephan Lewandowsky, Michael Kalish, S. K. Ngang. 2002. Simplified learning in complex situations: Knowledge partitioning in function learning. *Journal of Experimental Psychology: General* **131**:2, 163-193. [CrossRef]

375. Sheng-Uei Guan,, Peng Li,. 2002. A Hierarchical Incremental Learning Approach to Task Decomposition. *Journal of Intelligent Systems* **12**:3, 201. [CrossRef]

376. Giulio Cifarelli. 2001. Introduction. *The European Journal of Finance* **7**:4, 286-288. [CrossRef]

377. Mehdi Azzouzi, Ian T. Nabney. 2001. Dynamic local models for segmentation and prediction of financial time series. *The European Journal of Finance* **7**:4, 289-311. [CrossRef]

378. Christian Broberger, Tomas Hökfelt. 2001. Hypothalamic and vagal neuropeptide circuitries regulating food intake. *Physiology & Behavior* **74**:4-5, 669-682. [CrossRef]

379. R. Polikar, L. Upda, S.S. Upda, V. Honavar. 2001. Learn++: an incremental learning algorithm for supervised neural networks. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)* **31**:4, 497-508. [CrossRef]

380. Tao Hong, M.T.C. Fang. 2001. Detection and classification of partial discharge using a feature decomposition-based modular neural network. *IEEE Transactions on Instrumentation and Measurement* **50**:5, 1349-1354. [CrossRef]

381. T. S. Hu, K. C. Lam, S. T. Ng. 2001. River flow time series prediction with a range-dependent neural network. *Hydrological Sciences Journal* **46**:5, 729-745. [CrossRef]

382. L See. 2001. Multi-model data fusion for hydrological forecasting. *Computers & Geosciences* **27**:8, 987-994. [CrossRef]

383. L Kuncheva. 2001. Using measures of similarity and inclusion for multiple classifier fusion by decision templates. *Fuzzy Sets and Systems* **122**:3, 401-407. [CrossRef]

384. Jochen Triesch, Christoph von der Malsburg. 2001. Democratic Integration: Self-Organized Integration of Adaptive Cues. *Neural Computation* **13**:9, 2049-2074. [Abstract] [PDF] [PDF Plus]

385. Tze Leung Lai, Samuel Po-Shing Wong. 2001. Stochastic Neural Networks With Applications to Nonlinear Time Series. *Journal of the American Statistical Association* **96**:455, 968-981. [CrossRef]

386. A.A. Ilumoka. 2001. Efficient and accurate crosstalk prediction via neural net-based topological decomposition of 3-D interconnect. *IEEE Transactions on Advanced Packaging* **24**:3, 268-276. [CrossRef]

387. David C. Becalick, Timothy J. Coats. 2001. Comparison of Artificial Intelligence Techniques with UKTRISS for Estimating Probability of Survival after Trauma. *The Journal of Trauma: Injury, Infection and Critical Care* **51**:1, 123-133. [CrossRef]

388. Jefferson T. Davis, Athanasios Episcopos, Sannaka Wettimuny. 2001. Predicting direction shifts on Canadian-US exchange rates with artificial neural networks. *International Journal of Intelligent Systems in Accounting, Finance & Management* **10**:2, 83-96. [CrossRef]

389. Yuhong Yang. 2001. Adaptive Regression by Mixing. *Journal of the American Statistical Association* **96**:454, 574-588. [CrossRef]

390. J PERES, R OLIVEIRA, S FEYODEAZEVEDO. 2001. Knowledge based modular networks for process modelling and control. *Computers & Chemical Engineering* **25**:4-6, 783-791. [CrossRef]

391. Edmondo Trentin. 2001. Networks with trainable amplitude of activation functions. *Neural Networks* **14**:4-5, 471-493. [CrossRef]

392. J Svensson, M von Hellermann, R W T König. 2001. *Plasma Physics and Controlled Fusion* **43**:4, 389-403. [CrossRef]

393. Hsin-Chia Fu, Yen-Po Lee, Cheng-Chin Chiang, Hsiao-Tien Pao. 2001. Divide-and-conquer learning and modular perceptron networks. *IEEE Transactions on Neural Networks* **12**:2, 250-263. [CrossRef]

394. A Fiordaliso. 2001. Autostructuration of fuzzy systems by rules sensitivity analysis. *Fuzzy Sets and Systems* **118**:2, 281-296. [CrossRef]

395. YOUNÈS BENNANI, FABRÌCE BOSSAERT. 2001. MODULAR CONNECTIONIST MODELLING AND CLASSIFICATION APPROACHES FOR LOCAL DIAGNOSIS IN TELECOMMUNICATION TRAFFIC MANAGEMENT. *International Journal of Computational Intelligence and Applications* **01**:01, 53-70. [CrossRef]

396. H. Yin, N.M. Allinson. 2001. Self-organizing mixture networks for probability density estimation. *IEEE Transactions on Neural Networks* **12**:2, 405-411. [CrossRef]

397. A Fiordaliso. 2001. A constrained Takagi–Sugeno fuzzy system that allows for better interpretation and analysis. *Fuzzy Sets and Systems* **118**:2, 307-318. [CrossRef]

398. Lei Xu. 2001. Best Harmony, Unified RPCL and Automated Model Selection for Unsupervised and Supervised Learning on Gaussian Mixtures, Three-Layer Nets and ME-RBF-SVM Models. *International Journal of Neural Systems* **11**:01, 43-69. [CrossRef]

399. L Kuncheva. 2001. Decision templates for multiple classifier fusion: an experimental comparison. *Pattern Recognition* **34**:2, 299-314. [CrossRef]

400. B Zhang. 2001. A nonlinear neural network model of mixture of local principal component analysis: application to handwritten digits recognition. *Pattern Recognition* **34**:2, 203-214. [CrossRef]

401. ZOUBIN GHAHRAMANI. 2001. AN INTRODUCTION TO HIDDEN MARKOV MODELS AND BAYESIAN NETWORKS. *International Journal of Pattern Recognition and Artificial Intelligence* **15**:01, 9-42. [CrossRef]

402. R. Feraund, O.J. Bernier, J.-E. Viallet, M. Collobert. 2001. A fast and accurate face detector based on neural networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **23**:1, 42-53. [CrossRef]

403. C.J. Harris, X. Hong. 2001. Neurofuzzy mixture of experts network parallel learning and model construction algorithms. *IEE Proceedings - Control Theory and Applications* **148**:6, 456. [CrossRef]

404. S Sternberg. 2001. Separate modifiability, mental modules, and the use of pure and composite measures to reveal them. *Acta Psychologica* **106**:1-2, 147-246. [CrossRef]

405. X. Dai. 2001. CMA-based nonlinear blind equaliser modelled by a two-layer feedforward neural network. *IEE Proceedings - Communications* **148**:4, 243. [CrossRef]

406. Y. Bengio, V.-P. Lauzon, R. Ducharme. 2001. Experiments on the application of IOHMMs to model financial returns series. *IEEE Transactions on Neural Networks* **12**:1, 113-123. [CrossRef]

407. Naonori Ueda. 2001. *Transactions of the Japanese Society for Artificial Intelligence* **16**, 299-308. [CrossRef]

408. Qiang Gan, C.J. Harris. 2001. A hybrid learning scheme combining EM and MASMOD algorithms for fuzzy local linearization modeling. *IEEE Transactions on Neural Networks* **12**:1, 43-53. [CrossRef]

409. Martin Kreutz, Maik Anschütz, Thorsten Grünendick, Andreas Rick, Stefan Gehlen, Klaus Hoffmann. 2001. AUTOMATED DIAGNOSIS OF SKIN CANCER USING DIGITAL IMAGE PROCESSING AND MIXTURE-OF-EXPERTS. *Biomedizinische Technik/Biomedical Engineering* **46**:s1, 376-377. [CrossRef]

410. S LIEHR, K PAWELZIK. 2000. A trading strategy with variable investment from minimizing risk to profit ratio. *Physica A: Statistical Mechanics and its Applications* **287**:3-4, 524-538. [CrossRef]

411. Dirk Husmeier. 2000. The Bayesian Evidence Scheme for Regularizing Probability-Density Estimating Neural Networks. *Neural Computation* **12**:11, 2685-2717. [Abstract] [PDF] [PDF Plus]

412. L.C. Jain, L.I. Kuncheva. 2000. Designing classifier fusion systems by genetic algorithms. *IEEE Transactions on Evolutionary Computation* **4**:4, 327-336. [CrossRef]

413. T. Higuchi, Xin Yao, Yong Liu. 2000. Evolutionary ensembles with negative correlation learning. *IEEE Transactions on Evolutionary Computation* **4**:4, 380-387. [CrossRef]

414. M.A. Carreira-Perpinan. 2000. Mode-finding for mixtures of Gaussian distributions. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**:11, 1318-1323. [CrossRef]

415. D WEST. 2000. Neural network credit scoring models. *Computers & Operations Research* **27**:11-12, 1131-1152. [CrossRef]

416. E Adi-Japha. 2000. Regulation of division of labour between cognitive systems controlling action. *Cognition* **76**:1, 1-11. [CrossRef]

417. S. Gutta, J.R.J. Huang, P. Jonathon, H. Wechsler. 2000. Mixture of experts for classification of gender, ethnic origin, and pose of human faces. *IEEE Transactions on Neural Networks* **11**:4, 948-960. [CrossRef]

418. Andreas S. Weigend, Shanming Shi. 2000. Predicting daily probability distributions of S&P500 returns. *Journal of Forecasting* **19**:4, 375-392. [CrossRef]

419. Wenxin Jiang. 2000. The VC Dimension for Mixtures of Binary Classifiers. *Neural Computation* **12**:6, 1293-1301. [Abstract] [PDF] [PDF Plus]

420. Ichiro Takeuchi, Takeshi Furuhashi. 2000. A description of dynamic behavior of sensory/motor systems with fuzzy symbolic dynamic systems. *Artificial Life and Robotics* **4**:2, 84-88. [CrossRef]

421. Wenxin Jiang, M.A. Tanner. 2000. On the asymptotic normality of hierarchical mixtures-of-experts for generalized linear models. *IEEE Transactions on Information Theory* **46**:3, 1005-1013. [CrossRef]

422. Fu-Lai Chung, Ji-Cheng Duan. 2000. On multistage fuzzy neural network modeling. *IEEE Transactions on Fuzzy Systems* **8**:2, 125-142. [CrossRef]

423. LARS ASKER, MATS DANIELSON, LOVE EKENBERG. 2000. COMMITTEES OF LEARNING AGENTS. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **08**:02, 187-202. [CrossRef]

424. A. Khotanzad, H. Elragal, T.-L. Lu. 2000. Combination of artificial neural-network forecasters for prediction of natural gas consumption. *IEEE Transactions on Neural Networks* **11**:2, 464-473. [CrossRef]

425. A. Karniel, G.F. Inbar. 2000. Human motor control: learning to control a time-varying, nonlinear, many-to-one system. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)* **30**:1, 1-11. [CrossRef]

426. Shiro Ikeda. 2000. Acceleration of the EM algorithm. *Systems and Computers in Japan* **31**:2, 10-18. [CrossRef]

427. Chapter 8 Design issues — Neural networks **1**, 89-102. [CrossRef]

428. Qun Zhao, Jose C. Principe, Victor L. Brennan, Dongxin Xu, Zheng Wang. 2000. Synthetic aperture radar automatic target recognition with three strategies of learning and representation. *Optical Engineering* **39**:5, 1230. [CrossRef]

429. Azriel Rosenfeld, Harry Wechsler. 2000. Pattern recognition: Historical perspective and future directions. *International Journal of Imaging Systems and Technology* **11**:2, 101-116. [CrossRef]

430. Evan Heit, Lewis BottKnowledge selection in category learning **39**, 163-199. [CrossRef]

431. D West. 2000. Improving diagnostic accuracy using a hierarchical neural network to model decision subtasks. *International Journal of Medical Informatics* **57**:1, 41-55. [CrossRef]

432. A.K. Jain, P.W. Duin, Jianchang Mao. 2000. Statistical pattern recognition: a review. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**:1, 4-37. [CrossRef]

433. G. Deng, H. Ye, L.W. Cahill. 2000. Adaptive combination of linear predictors for lossless image compression. *IEE Proceedings - Science, Measurement and Technology* **147**:6, 414. [CrossRef]

434. J. L. Castro, M. Delgado, C. J. Mantas. 2000. MORSE: A general model to represent structured knowledge. *International Journal of Intelligent Systems* **15**:1, 27-43. [CrossRef]

435. J. Peres, R. Oliveira, S. Feyo de AzevedoKnowledge based modular networks for process modelling and control **8**, 247-252. [CrossRef]

436. Paul J. WerbosNeurocontrollers . [CrossRef]

437. Yair BartalDivide-and-Conquer Methods . [CrossRef]

438. A.N. Srivastava, R. Su, A.S. Weigend. 1999. Data mining for features using scale-sensitive gated experts. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **21**:12, 1268-1279. [CrossRef]

439. Ferdinando A Mussa-Ivaldi. 1999. Modular features of motor control and learning. *Current Opinion in Neurobiology* **9**:6, 713-717. [CrossRef]

440. Mitsuo Kawato. 1999. Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology* **9**:6, 718-727. [CrossRef]

441. Chin-Teng Lin, I-Fang Chung. 1999. A reinforcement neuro-fuzzy combiner for multiobjective control. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)* **29**:6, 726-744. [CrossRef]

442. KE CHEN, HUISHENG CHI. 1999. A MODULAR NEURAL NETWORK ARCHITECTURE FOR PATTERN CLASSIFICATION BASED ON DIFFERENT FEATURE SETS. *International Journal of Neural Systems* **09**:06, 563-581. [CrossRef]

443. K Chen. 1999. Improved learning algorithms for mixture of experts in multiclass classification. *Neural Networks* **12**:9, 1229-1252. [CrossRef]

444. W Jiang. 1999. On the identifiability of mixtures-of-experts. *Neural Networks* **12**:9, 1253-1258. [CrossRef]

445. Bao-Liang Lu, H. Kita, Y. Nishikawa. 1999. Inverting feedforward neural networks using linear and nonlinear programming. *IEEE Transactions on Neural Networks* **10**:6, 1271-1290. [CrossRef]

446. J Tani. 1999. Learning to perceive the world as articulated: an approach for hierarchical learning in sensory-motor systems. *Neural Networks* **12**:7-8, 1131-1141. [CrossRef]

447. M Dailey. 1999. Organization of face and object recognition in modular neural network models. *Neural Networks* **12**:7-8, 1053-1074. [CrossRef]

448. H Ando. 1999. Unsupervised visual learning of three-dimensional objects using a modular network architecture. *Neural Networks* **12**:7-8, 1037-1051. [CrossRef]

449. H Gross. 1999. Generative character of perception: a neural architecture for sensorimotor anticipation. *Neural Networks* **12**:7-8, 1101-1129. [CrossRef]

450. J Hansen. 1999. Combining predictors: comparison of five meta machine learning methods. *Information Sciences* **119**:1-2, 91-105. [CrossRef]

451. D.J. Miller, Lian Yan. 1999. Critic-driven ensemble classification. *IEEE Transactions on Signal Processing* **47**:10, 2833-2844. [CrossRef]

452. Bao-Liang Lu, M. Ito. 1999. Task decomposition and module combination based on class relations: a modular neural network for pattern classification. *IEEE Transactions on Neural Networks* **10**:5, 1244-1256. [CrossRef]

453. B. Apolloni, I. Zoppis. 1999. Sub-symbolically managing pieces of symbolical functions for sorting. *IEEE Transactions on Neural Networks* **10**:5, 1099-1122. [CrossRef]

454. A.D. Wilson, A.F. Bobick. 1999. Parametric hidden Markov models for gesture recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **21**:9, 884-900. [CrossRef]

455. Sun-Yuan Kung, J. Taur, Shang-Hung Lin. 1999. Synergistic modeling and applications of hierarchical fuzzy neural networks. *Proceedings of the IEEE* **87**:9, 1550-1574. [CrossRef]

456. I-Cheng Yeh. 1999. Modeling chaotic two-dimensional mapping with fuzzy-neuron networks. *Fuzzy Sets and Systems* **105**:3, 421-427. [CrossRef]

457. Wenxin Jiang, Martin A. Tanner. 1999. On the Approximation Rate of Hierarchical Mixtures-of-Experts for Generalized Linear Models. *Neural Computation* **11**:5, 1183-1198. [Abstract] [PDF] [PDF Plus]

458. J Zhang. 1999. Inferential estimation of polymer quality using bootstrap aggregated neural networks. *Neural Networks* **12**:6, 927-938. [CrossRef]

459. Y Yeun. 1999. Function approximations by coupling neural networks and genetic programming trees with oblique decision trees. *Artificial Intelligence in Engineering* **13**:3, 223-239. [CrossRef]

460. S.A. Rizvi, L.-C. Wang, N.M. Nasrabadi. 1999. Rate-constrained modular predictive residual vector quantization of digital images. *IEEE Signal Processing Letters* **6**:6, 135-137. [CrossRef]

461. R Sun. 1999. Multi-agent reinforcement learning: weighting and partitioning. *Neural Networks* **12**:4-5, 727-753. [CrossRef]

462. Ori Rosen, Martin Tanner. 1999. Mixtures of proportional hazards regression models. *Statistics in Medicine* **18**:9, 1119-1131. [CrossRef]

463. I.A. Taha, J. Ghosh. 1999. Symbolic interpretation of artificial neural networks. *IEEE Transactions on Knowledge and Data Engineering* **11**:3, 448-463. [CrossRef]

464. J Alexander. 1999. Template-based procedures for neural network interpretation. *Neural Networks* **12**:3, 479-498. [CrossRef]

465. 1999. Developing robust non-linear models through bootstrap aggregated neural networks. *Neurocomputing* **25**:1-3, 93-113. [CrossRef]

466. GASSER AUDA, MOHAMED KAMEL. 1999. MODULAR NEURAL NETWORKS: A SURVEY. *International Journal of Neural Systems* **09**:02, 129-151. [CrossRef]

467. A. Sarajedini, R. Hecht-Nielsen, P.M. Chau. 1999. Conditional probability density function estimation with sigmoidal neural networks. *IEEE Transactions on Neural Networks* **10**:2, 231-238. [CrossRef]

468. Ran Avnimelech, Nathan Intrator. 1999. Boosting Regression Estimators. *Neural Computation* **11**:2, 499-520. [Abstract] [PDF] [PDF Plus]

469. Suzanna Becker. 1999. Implicit Learning in 3D Object Recognition: The Importance of Temporal Context. *Neural Computation* **11**:2, 347-374. [Abstract] [PDF] [PDF Plus]

470. C. Di Natale, E. Proietti, R. Diamanti, A. D'Amico. 1999. Modeling of APCVD-doped silicon dioxide deposition process by a modular neural network. *IEEE Transactions on Semiconductor Manufacturing* **12**:1, 109-115. [CrossRef]

471. Matthew N. Dailey, Garrison W. CottrellChapter 10 Prosopagnosia in modular neural network models **121**, 165-184. [CrossRef]

472. Robert A. Jacobs. 1999. Computational studies of the development of functionally specialized neural modules. *Trends in Cognitive Sciences* **3**:1, 31-38. [CrossRef]

473. Kenneth J. Kurtz, Dedre Gentner, Virginia GunnReasoning 145-200. [CrossRef]

474. V. Ramamurti, J. Ghosh. 1999. Structurally adaptive modular networks for nonstationary environments. *IEEE Transactions on Neural Networks* **10**:1, 152-160. [CrossRef]

475. M A Wani, D T Pham. 1999. Efficient control chart pattern recognition through synergistic and distributed artificial neural networks. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture* **213**:2, 157-169. [CrossRef]

476. P Schyns. 1999. Dr. Angry and Mr. Smile: when categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition* **69**:3, 243-265. [CrossRef]

477. S Khan. 1998. Statistical and neural classifiers to detect traffic operational problems on urban arterials. *Transportation Research Part C: Emerging Technologies* **6**:5-6, 291-314. [CrossRef]

478. Tom Ziemke. 1998. Adaptive Behavior in Autonomous Agents. *Presence: Teleoperators and Virtual Environments* **7**:6, 564-587. [Abstract] [PDF] [PDF Plus]

479. Rainer Dietrich, Manfred Opper. 1998. *Journal of Physics A: Mathematical and General* **31**:46, 9131-9147. [CrossRef]

480. I-Cheng Yeh. 1998. Modeling Concrete Strength with Augment-Neuron Networks. *Journal of Materials in Civil Engineering* **10**:4, 263-268. [CrossRef]

481. R O'Reilly. 1998. Six principles for biologically based computational models of cortical cognition. *Trends in Cognitive Sciences* **2**:11, 455-462. [CrossRef]

482. C.L. Fancourt, J.C. Principe. 1998. Competitive principal component analysis for locally stationary time series. *IEEE Transactions on Signal Processing* **46**:11, 3068-3081. [CrossRef]

483. R. Sun, T. Peterson. 1998. Autonomous learning of sequential tasks: experiments and analyses. *IEEE Transactions on Neural Networks* **9**:6, 1217-1234. [CrossRef]

484. K. Rose. 1998. Deterministic annealing for clustering, compression, classification, regression, and related optimization problems. *Proceedings of the IEEE* **86**:11, 2210-2239. [CrossRef]

485. J. Yen, Liang Wang, C.W. Gillespie. 1998. Improving the interpretability of TSK fuzzy models by combining global learning and local learning. *IEEE Transactions on Fuzzy Systems* **6**:4, 530-537. [CrossRef]

486. D Wolpert. 1998. Multiple paired forward and inverse models for motor control. *Neural Networks* **11**:7-8, 1317-1329. [CrossRef]

487. J Buessler. 1998. Visually guided movements: learning with modular neural maps in robotics. *Neural Networks* **11**:7-8, 1395-1415. [CrossRef]

488. S.K.T. Kriebel, W. Brauer, W. Eifler. 1998. Coastal upwelling prediction with a mixture of neural networks. *IEEE Transactions on Geoscience and Remote Sensing* **36**:5, 1508-1518. [CrossRef]

489. Chi-Hang Lam, F. Shin. 1998. Formation and dynamics of modules in a dual-tasking multilayer feed-forward neural network. *Physical Review E* **58**:3, 3673-3677. [CrossRef]

490. C. Ornes, J. Sklansky. 1998. A visual neural classifier. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)* **28**:4, 620-625. [CrossRef]

491. A Mirhosseini. 1998. Human Face Image Recognition: An Evidence Aggregation Approach. *Computer Vision and Image Understanding* **71**:2, 213-230. [CrossRef]

492. James A. Reggia, Sharon Goodall, Yuri Shkuro. 1998. Computational Studies of Lateralization of Phoneme Sequence Generation. *Neural Computation* **10**:5, 1277-1297. [Abstract] [PDF] [PDF Plus]

493. Sun-Yuan Kung, Jenq-Neng Hwang. 1998. Neural networks for intelligent multimedia processing. *Proceedings of the IEEE* **86**:6, 1244-1272. [CrossRef]

494. Y. Shimshoni, N. Intrator. 1998. Classification of seismic signals by integrating ensembles of neural networks. *IEEE Transactions on Signal Processing* **46**:5, 1194-1201. [CrossRef]

495. Ke Chen. 1998. A connectionist method for pattern classification with diverse features. *Pattern Recognition Letters* **19**:7, 545-558. [CrossRef]

496. A.J. Zeevi, R. Meir, V. Maiorov. 1998. Error bounds for functional approximation and estimation using mixtures of experts. *IEEE Transactions on Information Theory* **44**:3, 1010-1025. [CrossRef]

497. P Szymanski. 1998. Hybrid interior point training of modular neural networks. *Neural Networks* **11**:2, 215-234. [CrossRef]

498. R. Rae, H.J. Ritter. 1998. Recognition of human head orientation based on artificial neural networks. *IEEE Transactions on Neural Networks* **9**:2, 257-265. [CrossRef]

499. M.M. Poulton, R.A. Birken. 1998. Estimating one-dimensional models from frequency-domain electromagnetic data using modular neural networks. *IEEE Transactions on Geoscience and Remote Sensing* **36**:2, 547-555. [CrossRef]

500. David J. Miller, Hasan S. Uyar. 1998. Combined Learning and Use for a Mixture Model Equivalent to the RBF Classifier. *Neural Computation* **10**:2, 281-293. [Abstract] [PDF] [PDF Plus]

501. E Bax. 1998. Validation of average error rate over classifiers. *Pattern Recognition Letters* **19**:2, 127-132. [CrossRef]

502. Kishan Mehrotra, Chilukuri K. MohanModular neural networks **3**, 147-181. [CrossRef]

503. Ian T Nabney, Christopher M BishopModeling Wind Direction from Satellite Scatterometer Data 295-301. [CrossRef]

504. Itiel E. Dror, Christopher S. SchreinerChapter 4 Neural networks and perception **126**, 77-85. [CrossRef]

505. 1998. Analysis of powder factors for tunnel blasting using neural networks. *Fragblast* **2**:4, 433-448. [CrossRef]

506. Monica Bianchini, Paolo Frasconi, Marco Gori, Marco MagginiOptimal learning in artificial neural networks: A theoretical view **2**, 1-51. [CrossRef]

507. Colin CampbellConstructive learning techniques for designing neural network systems **3**, 91-145. [CrossRef]

508. Yoram Singer. 1997. Adaptive Mixtures of Probabilistic Transducers. *Neural Computation* **9**:8, 1711-1733. [Abstract] [PDF] [PDF Plus]

509. Athanasios Kehagias, Vassilios Petridis. 1997. Time-Series Segmentation Using Predictive Modular Neural Networks. *Neural Computation* **9**:8, 1691-1709. [Abstract] [PDF] [PDF Plus]

510. A.L. McIlraith, H.C. Card. 1997. Birdsong recognition using backpropagation and multivariate statistics. *IEEE Transactions on Signal Processing* **45**:11, 2740-2748. [CrossRef]

511. A.V. Rao, D. Miller, K. Rose, A. Gersho. 1997. Mixture of experts regression modeling by deterministic annealing. *IEEE Transactions on Signal Processing* **45**:11, 2811-2820. [CrossRef]

512. N INTRATOR, S EDELMAN. 1997. Competitive learning in biological and artificial neural computation. *Trends in Cognitive Sciences* **1**:7, 268-272. [CrossRef]

513. C Williams. 1997. Instantiating Deformable Models with a Neural Net. *Computer Vision and Image Understanding* **68**:1, 120-126. [CrossRef]

514. J Newman. 1997. A Neural Global Workspace Model for Conscious Attention. *Neural Networks* **10**:7, 1195-1206. [CrossRef]

515. H Bersini. 1997. Now comes the time to defuzzify neuro-fuzzy models. *Fuzzy Sets and Systems* **90**:2, 161-169. [CrossRef]

516. Robert A. Jacobs. 1997. Nature, nurture, and the development of functional specializations: A computational approach. *Psychonomic Bulletin & Review* **4**:3, 299-309. [CrossRef]

517. Daniel M. Wolpert. 1997. Computational approaches to motor control. *Trends in Cognitive Sciences* **1**:6, 209-216. [CrossRef]

518. Yu Hen Hu, S. Palreddy, W.J. Tompkins. 1997. A patient-adaptable ECG beat classifier using a mixture of experts approach. *IEEE Transactions on Biomedical Engineering* **44**:9, 891-900. [CrossRef]

519. V. Petridis, A. Kehagias. 1997. Predictive modular fuzzy systems for time-series classification. *IEEE Transactions on Fuzzy Systems* **5**:3, 381-397. [CrossRef]

520. Kukjin Kang, Jong-Hoon Oh, Chulan Kwon. 1997. Learning by a population of perceptrons. *Physical Review E* **55**:3, 3257-3261. [CrossRef]

521. R Jacobs. 1997. A Bayesian Approach to Model Selection in Hierarchical Mixtures-of-Experts Architectures. *Neural Networks* **10**:2, 231-241. [CrossRef]

522. Robert A. Jacobs. 1997. Bias/Variance Analyses of Mixtures-of-Experts Architectures. *Neural Computation* **9**:2, 369-383. [Abstract] [PDF] [PDF Plus]

523. J.A. Benediktsson, J.R. Sveinsson, O.K. Ersoy, P.H. Swain. 1997. Parallel consensual neural networks. *IEEE Transactions on Neural Networks* **8**:1, 54-64. [CrossRef]

524. Ath. Kehagias, Vas. Petridis. 1997. Predictive Modular Neural Networks for Time Series Classification. *Neural Networks* **10**:1, 31-49. [CrossRef]

525. Sung-Bae Cho. 1997. Neural-network classifiers for recognizing totally unconstrained handwritten numerals. *IEEE Transactions on Neural Networks* **8**:1, 43-53. [CrossRef]

526. Shimon Edelman, Nathan IntratorLearning as Extraction of Low-Dimensional Representations **36**, 353-380. [CrossRef]

527. Anders Krogh, Peter Sollich. 1997. Statistical mechanics of ensemble learning. *Physical Review E* **55**:1, 811-825. [CrossRef]

528. Elizabeth Bates, Donna Thal, Doris Trauner, Judi Fenson, Dorothy Aram, Julie Eisele, Ruth Nass. 1997. From first words to grammar in children with focal brain injury. *Developmental Neuropsychology* **13**:3, 275-343. [CrossRef]

529. Chuanyi Ji, Sheng Ma. 1997. Combinations of weak classifiers. *IEEE Transactions on Neural Networks* **8**:1, 32-42. [CrossRef]

530. Assaf J. Zeevi, Ronny Meir. 1997. Density Estimation Through Convex Combinations of Densities: Approximation and Estimation Bounds. *Neural Networks* **10**:1, 99-109. [CrossRef]

531. V Desai. 1996. A comparison of neural networks and linear scoring models in the credit union environment. *European Journal of Operational Research* **95**:1, 24-37. [CrossRef]

532. R.C. Miall, D.M. Wolpert. 1996. Forward Models for Physiological Motor Control. *Neural Networks* **9**:8, 1265-1279. [CrossRef]

533. Pierre Baldi, Yves Chauvin. 1996. Hybrid Modeling, HMM/NN Architectures, and Protein Applications. *Neural Computation* **8**:7, 1541-1565. [Abstract] [PDF] [PDF Plus]

534. Y. Bengio, P. Frasconi. 1996. Input-output HMMs for sequence processing. *IEEE Transactions on Neural Networks* **7**:5, 1231-1249. [CrossRef]

535. Fengchun Peng, Robert A. Jacobs, Martin A. Tanner. 1996. Bayesian Inference in Mixtures-of-Experts and Hierarchical Mixtures-of-Experts Models with an Application to Speech Recognition. *Journal of the American Statistical Association* **91**:435, 953-960. [CrossRef]

536. David A. Cohn. 1996. Neural Network Exploration Using Optimal Experiment Design. *Neural Networks* **9**:6, 1071-1083. [CrossRef]

537. Christopher M. Bishop, Ian T. Nabney. 1996. Modeling Conditional Probability Distributions for Periodic Variables. *Neural Computation* **8**:5, 1123-1133. [Abstract] [PDF] [PDF Plus]

538. A. Khotanzad, J.J.-H. Liou. 1996. Recognition and pose estimation of unoccluded three-dimensional objects from a two-dimensional perspective view by banks of neural networks. *IEEE Transactions on Neural Networks* **7**:4, 897-906. [CrossRef]

539. J.del.R. Millan. 1996. Rapid, safe, and incremental learning of navigation strategies. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)* **26**:3, 408-420. [CrossRef]

540. Peter M. Williams. 1996. Using Neural Networks to Model Conditional Multivariate Densities. *Neural Computation* **8**:4, 843-854. [Abstract] [PDF] [PDF Plus]

541. D. Sarkar. 1996. Randomness in generalization ability: a source to improve it. *IEEE Transactions on Neural Networks* **7**:3, 676-685. [CrossRef]

542. E. Alpaydin, M.I. Jordan. 1996. Local linear perceptrons for classification. *IEEE Transactions on Neural Networks* **7**:3, 788-794. [CrossRef]

543. Klaus Pawelzik, Jens Kohlmorgen, Klaus-Robert Müller. 1996. Annealed Competition of Experts for a Segmentation and Classification of Switching Dynamics. *Neural Computation* **8**:2, 340-356. [Abstract] [PDF] [PDF Plus]

544. Vassilios Petridis, Athanasios Kehagias. 1996. A Recurrent Network Implementation of Time Series Classification. *Neural Computation* **8**:2, 357-372. [Abstract] [PDF] [PDF Plus]

545. David Heath, Simon Kasif, Steven SalzbergChapter 18 Committees of decision trees **113**, 305-317. [CrossRef]

546. V. Petridis, A. Kehagias. 1996. Modular neural networks for MAP classification of time series and the partition algorithm. *IEEE Transactions on Neural Networks* **7**:1, 73-86. [CrossRef]

547. Masahiko Shizawa. 1996. Multivalued regularization network-a theory of multilayer networks for learning many-to-h mappings. *Electronics and Communications in Japan (Part III: Fundamental Electronic Science)* **79**:9, 98-113. [CrossRef]

548. Sung-Bae Cho. 1995. Fuzzy aggregation of modular neural networks with ordered weighted averaging operators. *International Journal of Approximate Reasoning* **13**:4, 359-375. [CrossRef]

549. C Malsburg. 1995. Binding in models of perception and brain function. *Current Opinion in Neurobiology* **5**:4, 520-526. [CrossRef]

550. Younès Bennani. 1995. A Modular and Hybrid Connectionist System for Speaker Identification. *Neural Computation* **7**:4, 791-798. [Abstract] [PDF] [PDF Plus]

551. W Duch. 1995. Feature space mapping as a universal adaptive system. *Computer Physics Communications* **87**:3, 341-371. [CrossRef]

552. Peter Dayan, Richard S. Zemel. 1995. Competition and Multiple Cause Models. *Neural Computation* **7**:3, 565-579. [Abstract] [PDF] [PDF Plus]

553. G. Deco, D. Obradovic. 1995. Decorrelated Hebbian Learning for Clustering and Function Approximation. *Neural Computation* **7**:2, 338-348. [Abstract] [PDF] [PDF Plus]

554. M SPITZER. 1995. A neurocomputational approach to delusions. *Comprehensive Psychiatry* **36**:2, 83-105. [CrossRef]

555. Tor A. Johansen, Bjarne A. Foss. 1995. Identification of non-linear system structure and parameters using regime decomposition. *Automatica* **31**:2, 321-326. [CrossRef]

556. N Borghese. 1995. Generation of Temporal Sequences Using Local Dynamic Programming. *Neural Networks* **8**:1, 39-54. [CrossRef]

557. S Amari. 1995. Information geometry of the EM and em algorithms for neural networks. *Neural Networks* **8**:9, 1379-1408. [CrossRef]

558. Tor A. Johansen, Bjarne A. Foss. 1995. Semi-empirical modeling of non-linear dynamic systems through identification of operating regimes and local models. *Modeling, Identification and Control: A Norwegian Research Bulletin* **16**:4, 213-232. [CrossRef]

559. M Jordan. 1995. Convergence results for the EM approach to mixtures of experts architectures. *Neural Networks* **8**:9, 1409-1431. [CrossRef]

560. Steven J. Nowlan, Terrence J. Sejnowski. 1994. Filter selection model for motion segmentation and velocity integration. *Journal of the Optical Society of America A* **11**:12, 3177. [CrossRef]

561. S Sathiya Keerthi, B Ravindran. 1994. A tutorial survey of reinforcement learning. *Sadhana* **19**:6, 851-889. [CrossRef]

562. R. S. Shadafan, M. Niranjan. 1994. A Dynamic Neural Network Architecture by Sequential Partitioning of the Input Space. *Neural Computation* **6**:6, 1202-1222. [Abstract] [PDF] [PDF Plus]

563. Robert A. Jacobs, Stephen M. Kosslyn. 1994. Encoding Shape and Spatial Relations: The Role of Receptive Field Size in Coordinating Complementary Representations. *Cognitive Science* **18**:3, 361-386. [CrossRef]

564. Alan L. Yuille, Paul Stolorz, Joachim Utans. 1994. Statistical Physics, Mixtures of Distributions, and the EM Algorithm. *Neural Computation* **6**:2, 334-340. [Abstract] [PDF] [PDF Plus]

565. R BATTITI, A COLLA. 1994. Democracy in neural nets: Voting schemes for classification. *Neural Networks* **7**:4, 691-707. [CrossRef]

566. Garrison Cottrell, Kim Plunkett. 1994. Acquiring the Mapping from Meaning to Sounds. *Connection Science* **6**:4, 379-412. [CrossRef]

567. E.L. Hanzevack, T.W. Long, J.F. Menousek. 1994. NEURAL NETS SPEED FLOW CALCULATION. *Chemical Engineering Communications* **129**:1, 41-51. [CrossRef]

568. ANNETTE KARMILOFF-SMITH, ANDY CLARK. 1993. What's Special About the Development of the Human Mind/Brain?. *Mind & Language* **8**:4, 569-581. [CrossRef]

569. WILLIAM BECHTEL. 1993. The Path Beyond First-Order Connectionism. *Mind & Language* **8**:4, 531-539. [CrossRef]

570. William Bechtel. 1993. Currents in connectionism. *Minds and Machines* **3**:2, 125-153. [CrossRef]

571. J McClelland. 1993. Computational approaches to cognition: top-down approaches. *Current Opinion in Neurobiology* **3**:2, 209-216. [CrossRef]

572. Suzanna Becker, Geoffrey E. Hinton. 1993. Learning Mixture Models of Spatial Coherence. *Neural Computation* **5**:2, 267-277. [Abstract] [PDF] [PDF Plus]

573. Lei Xu. 1993. Least mean square error reconstruction principle for self-organizing neural-nets. *Neural Networks* **6**:5, 627-648. [CrossRef]

574. Robert A. Jacobs,, Michael I. Jordan. 1992. Computational Consequences of a Bias toward Short Connections. *Journal of Cognitive Neuroscience* **4**:4, 323-336. [Abstract] [PDF] [PDF Plus]

575. Robert A. Jacobs, Michael I. Jordan, Andrew G. Barto. 1991. Task Decomposition Through Competition in a Modular Connectionist Architecture: The What and Where Vision Tasks. *Cognitive Science* **15**:2, 219-250. [CrossRef]

576. Terence D. Sanger. 1991. A Tree-Structured Algorithm for Reducing Computation in Networks with Separable Basis Functions. *Neural Computation* **3**:1, 67-78. [Abstract] [PDF] [PDF Plus]