A Study of Neural and Fuzzy Parameters for Explicit and Implicit Knowledge-based Systems

Daniel C. Neagu

Department of Computing, School of Informatics, University of Bradford, BD7 1DP Bradford, UK

Abstract. In this paper, a framework of a unified neural and neuro-fuzzy approach to integrate implicit and explicit knowledge in hybrid intelligent systems is presented. In the developed hybrid system, training data used for neural and neuro-fuzzy models represents implicit domain knowledge. On the other hand, the explicit domain knowledge is represented by fuzzy rules, directly mapped into equivalent connectionist structures. A formal model for a hybrid intelligent system implemented as neural, neuro-fuzzy and fuzzy modules is proposed. Furthermore, this paper explores the influences of the main identified parameters of the proposed model on the accuracy of the hybrid intelligent system in a predictive data mining application.

1 Introduction

In recent years, hybrid intelligent systems (HIS) have drawn an increasing research interest. This approach has been successfully applied in various areas, such as speech and natural language understanding [1], [2], robotics [2], medical diagnosis [3], fault diagnosis of industrial equipment [2], financial applications [4], bioinformatics [5], predictive toxicology [6]. A particular attention is paid to HIS incorporating connectionist structures, known as hybrid neural systems [2]. However, there is still a need to homogeneously describe the modular structures of such knowledge-based systems in order to propose further approaches for development of suitable solutions.

One of the delicate problems encountered to develop a good hybrid neural model for a real-world data-driven application is parameter tuning. One of the main constraints is determined by insufficiency of amount, distribution and quality of existing data, such that the model cannot meet the expectations at a particular development stage. Many different methods exist for adapting HIS from data, for instance Adaptive Network-based Fuzzy Inference Systems (ANFIS), nonlinear global search techniques (e.g. genetic algorithms) or adaptive on-line incremental or hybrid (supervised and unsupervised) learning algorithms [1]. The drawbacks of such methods are their relative dependency on the data quality and further inconsistent and unpredicted global performance. Moreover, the development of a good quality system in the case of a modular combination of individual models comes with increased difficulty tasks.

An original HIS approach based on implicit and explicit knowledge representation has been found suitable to develop better models, surpassing some of the above listed

Neagu D. (2004).
A Study of Neural and Fuzzy Parameters for Explicit and Implicit Knowledge-based Systems.
In *Proceedings of the First International Workshop on Artificial Neural Networks: Data Preparation Techniques and Application Development*, pages 49-59
DOI: 10.5220/0001149700490059

disadvantages. Some of its applications in Predictive Data Mining are discussed in [6]. Its modular architecture comes with the advantage of incorporating training data as connectionist structures and human expertise in form of fuzzy rules. The approach demonstrates better robustness because of the modular combinations [5] of various incorporated expert opinions. However, one of the encountered challenges is the significance of its parameters to the quality of the global model.

The next sections will be focused on the formalism proposed to describe the parameterized structure of HIS and the synergy derived from the use of its complementary components (Section 2). A formal description of HIS is proposed in Section 3, together with considerations on the universe of discourse and some issues on integration algorithms for the development of the global structure. Some implications and significance of parameters to the system will be further illustrated through a case study. The application, described in Section 4, covers the use of structural, learning and descriptive parameters of various knowledge models to tune an integrated system. A particular case study from predictive toxicology is presented, along with some preliminary experimental results on the influence of the main parameters of the proposed intelligent system based on the modular integration of implicit and explicit knowledge modules. In the last section, the advantages of using modular HIS to develop knowledge fusion models and list some potential further research directions are summarized.

2 Knowledge Representation

The last ten years have produced a tremendous amount of research on fuzzy logic and connectionist fields. The current directions of research explore high-level connectionism and hybrid intelligent systems [2], [7]. The two approaches can be used in a complementary way, HIS combining connectionist and symbolic features. In such systems, the learner can insert fuzzy rules into neural networks. Once the domain knowledge has a neural representation, training examples are used to refine initial knowledge or additional structures. Finally, it processes the output for given instances and, using specific methods [8]-[10], can extract symbolic information from trained networks, to explain and interpret the refined connectionist knowledge.

The implicit knowledge is defined as connectionist representation of learning data. An explicit knowledge module has the role to adjust performances of implicit knowledge modules by using external information provided by experts, in form of Fuzzy Rule-based Systems. In our approach, connectionist integration of explicit and implicit knowledge appears a natural solution to develop homogeneous intelligent systems. Explicit and implicit rules are represented using MLP (Multi-Layer Perceptron) [11], neuro-fuzzy [12], fuzzy (FNN) or hybrid (HNN) neural nets [13]. Thus, fuzzy logic provides the inference mechanism under cognitive uncertainty, since neural nets offer advantages of learning, adaptation, fault-tolerance, parallelism and generalization.

The hybrid intelligent system considered in this paper is a multi-input single-output (MISO) neuro-fuzzy system (Fig. 1). The general goal is to model a combination of data and expert information to relate some inputs with the corresponding output value:

$$\Phi : \mathrm{D} \subseteq \boldsymbol{R}^{n} \rightarrow \boldsymbol{R}, \tag{1}$$

where $n \in \mathrm{N}$ is the number of the inputs from the universe of discourse U over the application domain. This leads to the following steps in a fuzzy neural computational process: (1) development of individual knowledge-based connectionist models, (2) modeling synaptic connections of individual models, which incorporate fuzziness into modules, and (3) adjusting the ensemble voting algorithm (Fig. 1).

The fundamental concepts and methods used in our approach [6][14][15] are based on the neuronal fuzzy model MAPI [7]. The MAPI neuron is used to implement the fuzzy neuro-symbolic processing units.

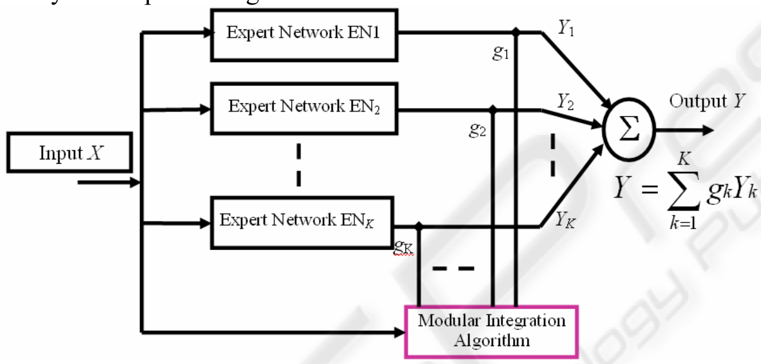


Fig. 1. The hierarchical architecture of the integrated intelligent system, based on modular combination of implicit and explicit knowledge modules.

2.1 The Neuronal Model

The artificial neuron MAPI (Matching, Aggregation, Projection, Inverse-matching neuron), proposed by Rocha [7] combines connectionist and fuzzy reasoning:

$$\text{MAPI}=\{\{X_p\}, Y, T, R, C, \theta, \{\alpha, f\}\}, \text{ where} \tag{2}$$

- $\{X_p\}$ is the family of pre-synaptic inputs over MAPI by all its n pre-synaptic axons;
- Y is the output code of MAPI;
- T is the family of transmitters used to exchange messages with other neurons;
- R is the family of receptors released by the pre-synaptic neurons;
- θ is the function used to aggregate the actual pre-synaptic activity;
- $\{\alpha, f\}$ is a family of thresholds and encoding functions defined as:

$$y = \begin{cases} w_l, if\ a_{MAPI} < \alpha_1 \\ w_u, if\ a_{MAPI} \geq \alpha_2 \\ f(\mathrm{a}_{\mathrm{MAPI}}), otherwise \end{cases} \tag{3}$$

- C is the set of controllers; each c_i actions over MAPI itself and other neurons.

The formal neuron exhibits capabilities of a multipurpose processing device, since it is able to handle different types of numerical calculations. This includes the processing capability of the classic neuron introduced by McCulloch and Pitts in 1943.

2.2 Explicit Knowledge Representation

According to the methodology presented in [13], Fuzzy Rule-Based Systems can be mapped into equivalent ANNs. We define the *explicit knowledge* as a knowledge base represented by neural networks, computationally identical to a given fuzzy rules set, and created by mapping a priori known fuzzy rules. The fuzzy rule set is described as a discrete fuzzy rule-based system (DFRBS [13]). Both, Mamdani and Sugeno zero and first order Fuzzy Inference Systems can be represented [7], [13] as EKMs. The intrinsic representation of explicit knowledge is based on MAPI fuzzy neurons [7]. Numerical weights corresponding to connections between neurons are computed using either Combine Rules First Method [7], [13] or Fire Each Rule Method [13].

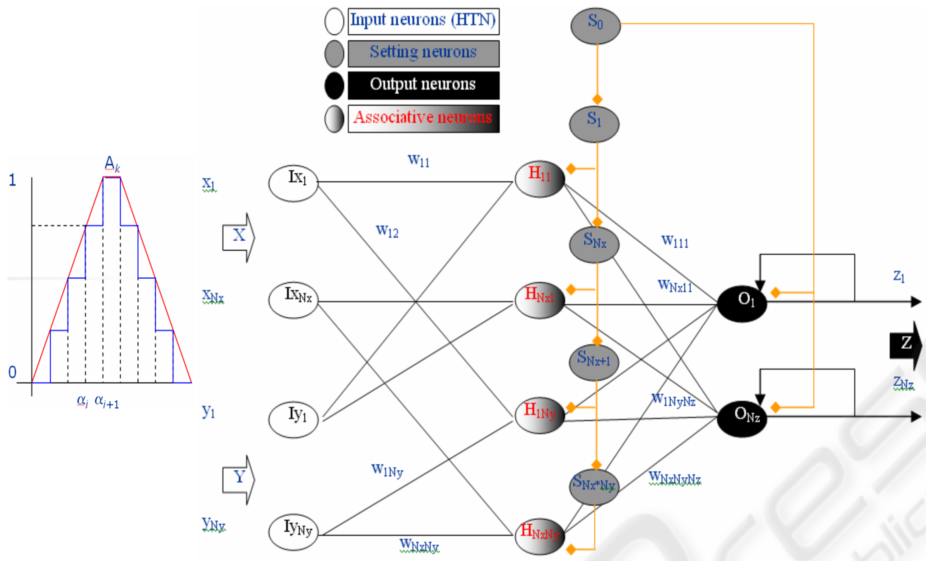
The neural reasoning engine is accorded to multiple premises fuzzy rules using fuzzy connectives. Considering the extended version of Modus Ponens [16]:

$$\begin{array}{c} \text{IF } X_1 \text{ is } A_1 \wedge \ldots \wedge X_p \text{ is } A_p \text{ then } Y \text{ is } B \\ \underline{(X_1 \text{ is } A'_1) \wedge \ldots \wedge (X_n \text{ is } A'_n)} \\ Y \text{ is } B' \end{array} \qquad (4)$$

where system inputs X_i, i=1,2,...,n, and output Y are linguistic variables. Thus, for example, let be considered a single rule with two antecedents:

$$\text{IF } X_1 \text{ is } A_1 \text{ AND } X_2 \text{ is } A_2 \text{ THEN } Y \text{ is } B \qquad (5)$$

where A_1, A_2, B are fuzzy sets having associated matching functions μ_{A1}, μ_{A2}, μ_B.



(a) (b)

Fig. 2. Connectionist Discrete Fuzzy Rule-based System: (a) discrete fuzzy set; (b) MAPI ANN equivalent with a rule with two premises (Combine Rules First Method)

Let the membership function $\mu_{A1}(\xi)$ be described by a vector X_1 of size m_1, so that:

$$x_{1i}=\mu_{A1}(\xi),\ \text{if}\ \alpha_i<\xi\leq\alpha_{i+1},\ i=1,2,...,\ m_1\text{-}1 \tag{6}$$

Introducing the discrete form (Fig. 2a) of fuzzy set $A_1=[x_{11}\ \dots\ x_{1m1}]$ in relation (5):

$$R{:}A_1\times A_2\times B\rightarrow[0,1],\ \mu_R(x_1,x_2,y)=(\mu_{A1}(\xi)\wedge\mu_{A2}(\psi))\Gamma\mu_B(\upsilon) \tag{7}$$

defines the discrete form of a fuzzy implication according to (4), where ^ and Γ are fuzzy connectives and implication operators. An equivalent structure using MAPI neurons to implement an explicit multi-premise rule [2] is shown in Fig. 2b.

2.3 Implicit Knowledge Representation

The *implicit knowledge* represents data collections acquired by learning procedures in connectionist structures. IKM structures have two representations: Multilayer Perceptrons as Crisp Neural Networks (IKM_CNN) or neuro-fuzzy nets (IKM_FNN) [14].

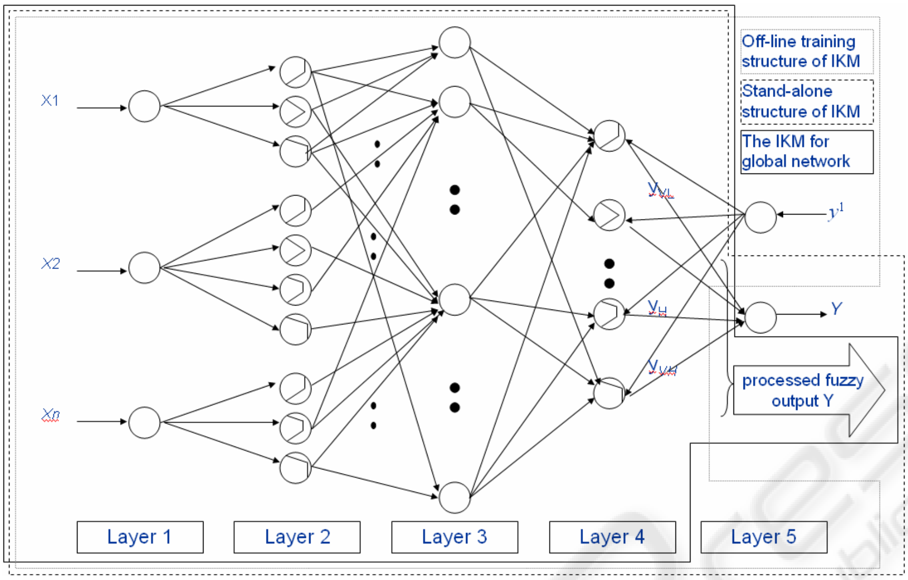


Fig. 3. Implicit Knowledge Module based on neural fuzzy processing IKM_FNN.

An IKM_CNN is a Multi-Layer Perceptron [11] whose typical equation for the weight changes by various learning algorithms is described by:

$$\Delta W_p = -\eta \nabla E(W_p) + \alpha \Delta W_{p-1} \tag{8}$$

in which ΔW_p represents the updates to the weight vector, $E(W_p)$ is the error function at the p-th iteration, η is the learning rate and α is the momentum term. The learning rate determines the speed the network moves along the error surface following its gradient. The momentum term smoothes out fluctuations in the error-weight space.

An IKM_FNN is a multi-layered neural structure based on an input layer to perform membership degrees of the current values, a fully connected three-layered MLP and a defuzzification layer (Fig. 3). MAPI input nodes implement membership functions for each linguistic input. The objective is to learn fuzzy associations between inputs and output: IKM_FNNs implement models dependent on learning and structural parameters, and on fuzzification algorithm (according to equations 7 and 8).

3 Implicit and Explicit Knowledge-based Intelligent System

Let's consider a MISO HIS with n inputs. Let also be considered $U = \prod_{i=1}^{n+1} D_i$ the universe of discourse over the application domain as the Cartesian product of sets D_i, i=1..n+1, having the input variables $X_i \in D_i$, i=1..n, and the output $Y \in D_{n+1}$.

A HIS as an integrated model of the problem Φ based on implicit and explicit knowledge modules is a good approximation of Φ as defined by:

$$HIS = \left\{ \underset{j=1..m}{M_j} / \forall \varepsilon > 0, \exists X \in \prod_{i=1}^{n} D_i, \forall Y = \Phi(X) : \left\| M_j(X) - Y \right\| < \varepsilon \right\} \tag{9}$$

where the knowledge modules are functional models $M_j : \prod_{i=1}^{n} D_{ij} \rightarrow D_{n+1,j}$.

The modules M_j are, in our approach, either implicit or explicit knowledge models $M_j \in \{M_{IKM\_CNN}, M_{IKM\_FNN}, M_{EKM\_Mamdani}, M_{EKM\_Sugeno}\}$. For any of these models, based on the connectionist homogeneous implementation of any M_j model, we can propose, following (3), (7) and (8), a formal parameter-based description of HIS:

$$M_j = \langle \Theta, \Lambda, \Omega \rangle \tag{10}$$

where Θ is the set of *topological parameters* (i.e. number of layers, number of neurons on each layer, connection matrices) of individual models and also of general structure (type and number of individual models and gating networks), Λ is the set of *learning parameters* (learning rate, momentum term, any early stopping attribute for implicit knowledge modules, but NIL for explicit knowledge modules) and Ω is the set of *description parameters* (defining for any fuzzy model number and type of fuzzy sets, and parameters of membership functions associated to linguistic variables).

Three distinctive cases to develop further integrated models can be identified:

Case 1: $D_j = \prod_{i=1}^{n} D_{ij}$ for all *j*=1..*m*. The model is a modular architecture [2], [15] combining experts on the whole input domain.

Case 2: $\bigcap_{j=1}^{m} \prod_{i=1}^{n} D_{ij} = 0$ and $D_i \cap D_j = 0$, for *j,k* = 1,..,*m*. The HIS model is a collection of *m* expert models on disjunctive input domains; the system is a top-down integrated decomposition model, by dividing the initial problem in separate less-complex sub-problems.

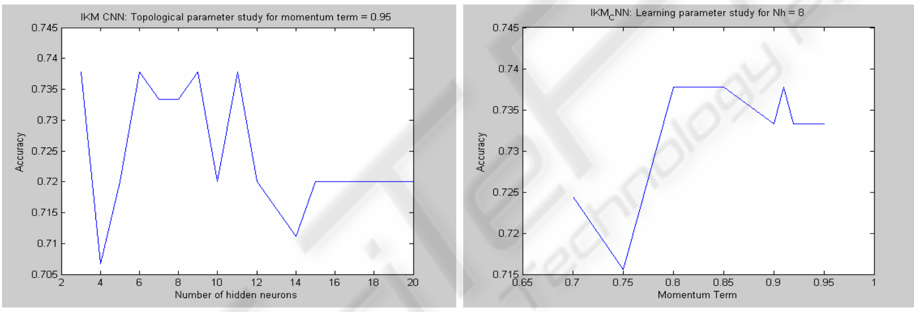
Case 3: $\bigcap_{j=1}^{m} \prod_{i=1}^{n} D_{ij} \neq 0$. The models are built on overlapping sub-domains and further algorithms to refine the problem as cases 1 or 2 are required [14], [15].

So far, few different strategies to combine IKM and EKM in a global HIS have been proposed [15]: Fire Each Module (FEM), Unsupervised-trained Gating Network (UGN), Supervised-trained Gating Network (SGN), majority voting etc. FEM is an adapted Fire Each Rule method [13] for modular networks, in two versions: statistical combination of crisp outputs (FEMS) or fuzzy inference of linguistic outputs (FEMF). The second strategy proposes competitive aggregation of EKMs and IKMs, while the SGN uses a supervised trained layer to process the overall output of modules.

4 The Influence of HIS Parameters: a Case Study

The case study considers the influence of HIS parameters to satisfy conditions for Case 1 (see Section 3). According to formulas (3), (8), (9), the influence of momentum term to IKM_CNN, IKM_FNN (trained with gradient descent adaptive learning rate backpropagation), the influence of membership functions to IKM_FNN and EKM_Sugeno and also the influence of learning and description parameters to the global models, developed using FEMS, FEMF, SGN integration algorithms, are considered. The main objective of this case study is to define the main parameters of the HIS model and to describe their importance in terms of prediction accuracy.

The case study is based on Predictive Toxicology data: the 2D ciliate (Tetrahymena pyriformis) population growth impairment (IGC50) values from TETRATOX database [17]. For the sake of simplicity, just two input chemical descriptors were finally chosen. The whole set of available patterns has been divided in two independent sets, for training and testing (70/30). For the accuracy measure, the absolute error of the predicted cases for the whole data set is used. The system consists on implicit knowledge modules (IKM_CNN, IKM_FNN) and explicit knowledge modules (EKM_Sugeno implementing a Quantitative SAR [6]).



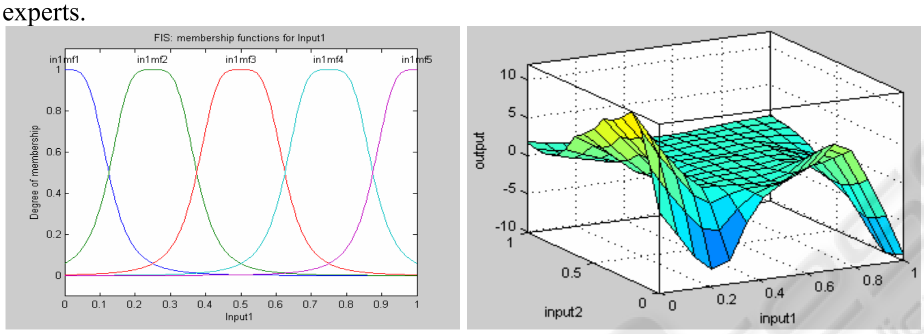
(a) number of hidden neurons (b) momentum term

Fig. 4. Tuning IKM_CNN: (a) topological parameters; (b) learning parameters.

IKM_CNNs were generated for various values of topological parameters (number of hidden neurons of the connectionist structure, Fig. 4a) and learning parameter (momentum term, Fig. 4b) and the best expert has been chosen IKM_CNN with 8 hidden neurons and momentum term of 0.85.

A further study on parameters description related to linguistic variables considered membership functions generated by ANFIS (Fig. 5a). Two approaches to fuzzify the variables were considered: Gaussian Bell membership functions for a Sugeno order 1 fuzzy system (Fig. 5b), and a balanced fuzzy split of domain intervals. For various combinations of membership functions (triangular, trapezoidal, Gaussian, Bell) the best results were of ANFIS generated Gaussian Bell 3-3-5 fuzzification procedure.

The global system, based on the best generated individual experts and one explicit QSAR [17], has been applied to the test data. Comparative results are depicted in Fig. 6. The results show that a tuned HIS model comes with better predictive abilities than traditional approaches (QSAR). The identified parameters, following the formal description (10) of the proposed HIS have a critical impact on the coverage and accuracy of the developed modular expert models. Following the results, any training algorithm proposes satisfactory performances as individual models, but modular combinations based on further parameters tuning will definitively increase the global predictions accuracy. However, the case study is based on a stepwise approach of parameters tuning, where the final models were built on already improved individual experts.



(a) number and shape of linguistic values (b) ANFIS fuzzy rules analysis

Fig. 5. Tuning description parameters for IKM_FNN and EKM.

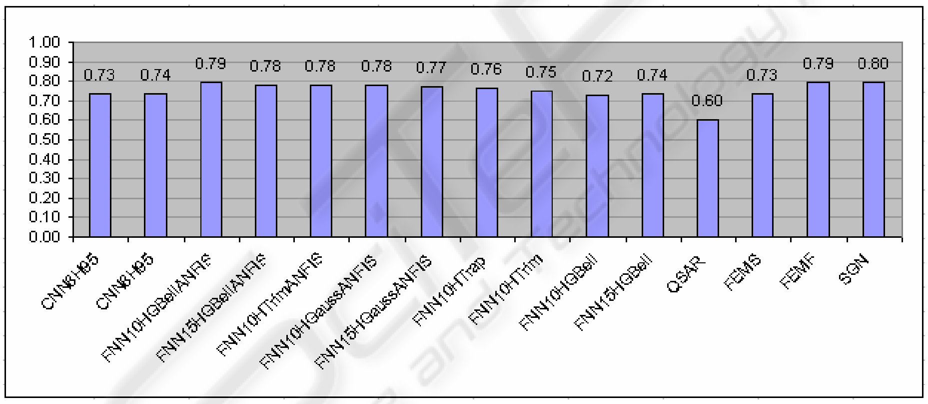


Fig. 6. Performances of best generated experts and global algorithms for HIS development.

5 Conclusions and Future Work

This paper briefly explains how different modular combinations of connectionist and fuzzy inference systems could be formulated using a parameter-based data driven functional approach and then investigates whether they can provide an improved level of performance, sufficiently good and robust to provide reliable models for predictive data mining. Experiment results reveal that all considered tuned parameters and combining paradigms can alter the developed hybrid models to show better performances to represent predictive toxicology data accurately.

The main problem regarding HIS development is the difficulty of delivering an optimized structure, due to existence of limitations in knowledge elicitation. Implicit and explicit knowledge models were analyzed in order to propose a formal description of HIS based on neural, neuro-fuzzy and fuzzy modules. The proposed model exhibits effective solutions for evaluation of available systems against representative samples, to choose the best combination of the available methods. The advantages of developing HIS models to combine implicit and explicit knowledge structures are identified. The implications and significance of individual and collective parameters tuning to the global system have been illustrated through a case study from predictive toxicology. Classes of individual parameters and their importance were also reviewed.

Comparison of various models developed on the predictive toxicology data suggests that, rather than using just randomly chosen connectionist though trained models, the use of modular combinations of tuned fuzzy experts significantly improves the performance of the hybrid system. However, the data quality and preprocessing training data is also quite important for the success of the tuned hybrid intelligent systems.

A reliable algorithm to optimally tune these parameters into the framework of the global method of combining the modules in HIS is critical to the quality of further predictions and the maintainability of the systems. Future work will be carried out to analyze new possibilities of parameters tuning for different expert domain models, mainly to consider the disjunctive experts collaboration in hybrid intelligent system area.

Acknowledgments

The author is grateful to T.W. Schultz, M.T.D. Cronin, A. Aptula and T.I. Netzeva for providing toxicity data and models used in this study.

This work is partially funded by the EU FP5 Quality of Life project DEMETRA under the contract QLK5-CT-2002-00691 [18].

References

1. Kasabov, N.; Kozma, R.: Hybrid intelligent adaptive systems: a framework and a case study on speech recognition. International Journal of Intelligent Systems. 13/6 (1998) 455-466
2. Wermter, S., Sun, R.: Hybrid Neural Systems. Springer Verlag: Heidelberg (2000)
3. Brasil, L.M., de Azevedo, F.M., Bareto, J.M.: A hybrid expert architecture for medical diagnosis. Proceedings of the 3th International Conference on Artificial Neural networks and Genetic Algorithms ICANNGA, Norwich, England (1997) 176-180
4. Takagi, H.: Cooperative system of neural networks and fuzzy logic and its application to consumer products. Ind Apps. Fuzzy Cont&Intell Syst. Van Nostrand Reinhold: NY (1994)
5. Neagu, D., Palade, V.: A Neuro-Fuzzy Approach for Functional Genomics Data Interpretation and Analysis. Journal of Neural Computing and Applications. 12/3-4, Springer-Verlag (2003) 153-159

6. Neagu, D. Gini, G.: Neuro-Fuzzy Knowledge Integration applied in Toxicity Prediction. In Jain, R., Abraham, A., Faucher, C., van der Zwaag, B.J. (eds.): Innovations In Knowledge Engineering. Advanced Knowledge International, Adelaide – Australia (2003) 311-342
7. da Rocha, AF: Neural Nets: A Theory for Brains and Machines: Springer-Verlag (1992)
8. Omlin, C.W., Giles, C.L.: Extraction of Rules from Discrete-Time Recurrent Neural Networks. Neural Networks 9 (1996) 41-52
9. Benitez, JM, Castro, JL, Requena, I: Are ANNs Black Boxes? IEEE Transactions Neural Networks 8/5 (1997) 1157-1164
10. Neagu, D., Palade, V.: An interactive fuzzy operator used in rule extraction from neural networks. Neural Networks World Journal 10/4 (2000) 675-684
11. Rumelhart, D., McClelland, J.: Parallel Distributed Processing: MIT Press (1988)
12. Nauck, D., Kruse, R.: NEFCLASS-X: A Neuro-Fuzzy Tool to Build Readable Fuzzy Classifiers. BT Tech. J. 16/3 (1998) 180-192
13. Buckley J.J. Hayashi Y.: Neural nets for fuzzy systems, Fuzzy Sets and Systems 71 (1995) 265-276
14. Neagu, D., Avouris, N. M., Kalapanidas, E., Palade, V.: Neurosymbolic Integration in a Knowledge-based System for Air Quality Prediction. Applied Intelligence. 17/2 (2002) 141–169
15. Neagu, D., Palade, V.: Modular neuro-fuzzy networks: an overview of explicit and implicit knowledge integration. Proceedings of Int'l FLAIRS-02 Conference, Pensacola, Florida: AAAI Press (2002) 277-281
16. Zadeh, L.A.: The role of fuzzy logic in the management of uncertainty in expert systems. Fuzzy Sets and Systems 11/3 (1983) 199-227
17. Schultz, TW, Sinks, GD, Cronin, MTD: Identification of mechanisms of toxic action of phenols to Tetrahymena pyriformis from molecular descriptors. QSARs in Env.Sc.7 (1997) 329-342
18. DEMETRA-Development of Environmental Modules for Evaluation of Toxicity of pesticide Residues in Agriculture: http://www.demetra-tox.net/