Revue Roumaine de Chimie, **2008**, *53*(4), 283–290

ARTIFICIAL INTELLIGENCE METHODS APPLIED TO PREDICTION OF THE LIQUID CRYSTALLINE BEHAVIOR OF SOME FERROCENE DERIVATIVES

Cătălin LISA, Silvia CURTEANU,* Victor BULACOVSCHI and Daniela APREUTESEI

"Gh. Asachi" Technical University of Iaşi, Faculty of Chemical Engineering,
71A, Bd. D. Mangeron 700050, Iaşi, Roumania

Received October 18, 2007

The present paper studies the prediction possibilities of the liquid crystalline behavior for some symmetrically derivatives with two ferrocene units using the artificial intelligence methods – neural networks. The liquid crystalline property is correlated with chemical structure of the compounds, quantified by a series of molecular descriptors, which are estimated by mechanical molecular simulation. Good predictions are obtained in the validation phase, which proves the generalization capability of the neural model and its availability for liquid crystal property estimation.

INTRODUCTION

Generally, in the study of liquid crystals, a major objective is the possibility to control liquid crystalline properties of material via the molecular structure.[1] The obtaining of liquid crystals with a wide mesomorphic domain is possible by using a careful combination of structural elements such as: linear shape, rigidity, high value of length/width ratio. The structural factors affect the interaction nature between liquid crystals molecules and are very important for getting the adequate mesomorphic behavior. Even small changes of the shape and structure of the molecule could have a great influence over the type of mesophase and transition temperature domain.

The main advantage of using artificial neural network (ANN) to the prediction of properties of organic compounds is the fact that neural networks can simulate the nonlinear relationship between structural information and properties of compounds during the training process, and generalize the knowledge among homologous series without need for theoretical formula. The ability of neural networks is significant in the determination of quantitative structure-property relationships, because compounds with known properties can be used to train ANN, thus properties of other compounds that cannot be ascertained by experimentation can be, subsequently, determined.

In this paper we used an organic compounds database[2-4] (35 in all) which includes a wide variety of symmetrically derivatives with two ferrocene units.

This work has shown the excellent capability of an artificial neural network approach for the prediction of liquid crystalline behavior of some organic compounds. Simple architecture modular neural networks and simple methods of establishing the networks' structure are proposed for process modeling: feed-forward networks with single, two or three hidden layers. Good predictions are obtained in the validation phase, which proves the generalization capability of the neural model and its availability for liquid crystal property estimation.

METHODS

Liquid crystals with wide mesomorphic domain are obtained by careful combination of some important structural elements such as: linear shape, rigidity, and high values of the length/diameter ratio. In order to assure the desired rigidity but also the increase of the system's polarizability, the

* Corresponding author: silvia_curteanu@yahoo.com

mesogenic unit that is about to be attached to ferrocene structure is made of two aromatic rings, connected or not by connecting groups. The units used in this general structure and their combinations induce not only the type of liquid crystals mesophases, but also the physical properties of the compounds.

New symmetrically derivatives with two ferrocene units, with liquid crystal properties and various molecular designs have been synthesized and their mesomorphic properties analyzed. The synthesis of mesomorphic crystals properties have been reported previously.[2,3] Our database contains also ferrocene derivatives synthesized by Philippe Massiot.[4]

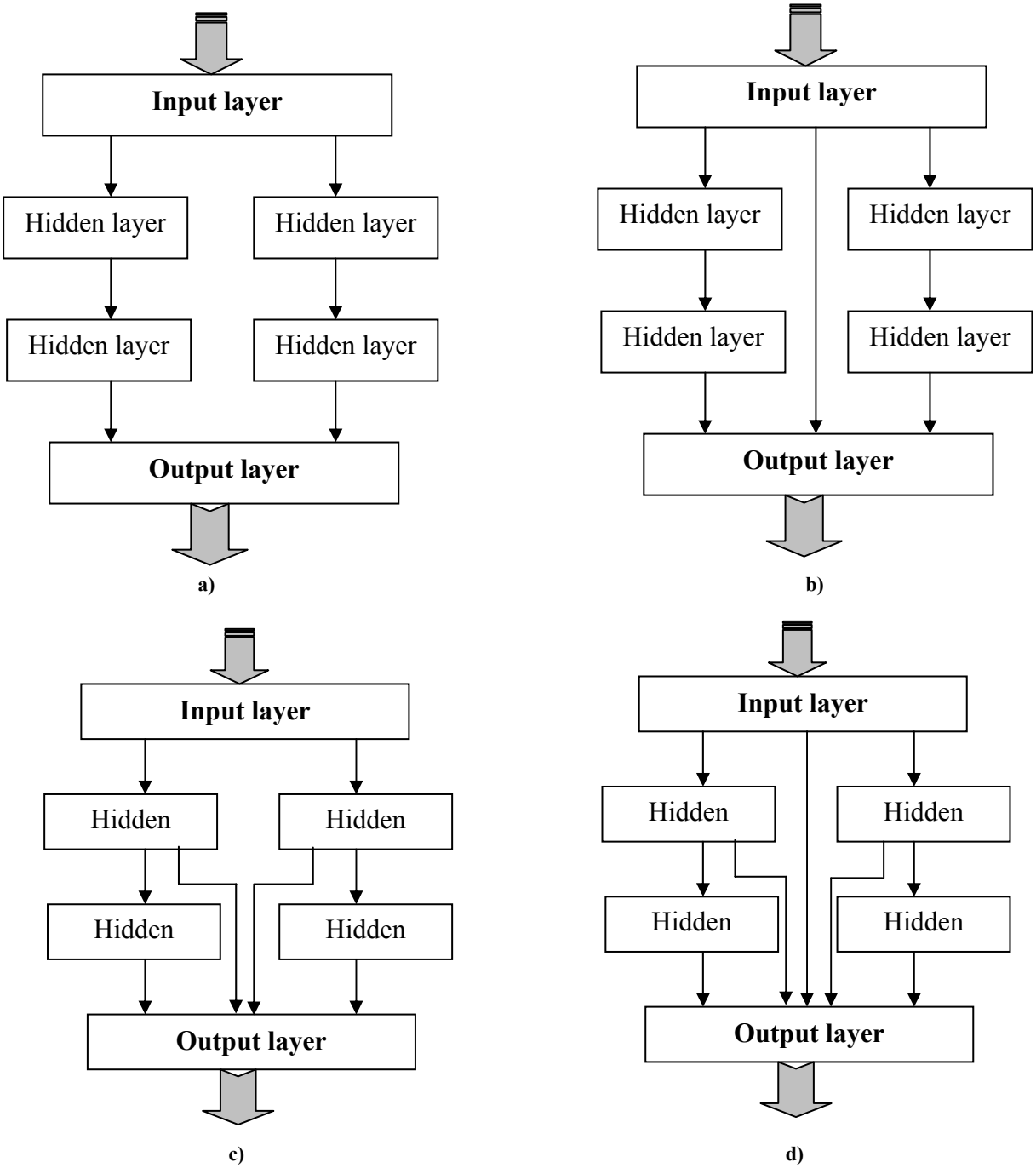


Fig. 1 – Modular Neural Network architectures based on the standard backpropagation algorihm:
a) Two hidden slabs with possibility of two activation functions; b) Two hidden slabs with possibility of three activation functions and a jump connection; c) Two hidden slabs with possibility of four activation functions and a jump connection between hidden layer and output layer; d) Two hidden slabs with possibility of five activation functions and a jump connection between hidden layer and output layer.

Artificial neural network modeling has been applied in numerous chemical applications.[5-8] The mathematical adaptability of ANN recommends them as a powerful tool for pattern classification and building predictive models. A particular advantage of ANNs is their inherent ability to incorporate nonlinear dependencies between the dependent and independent variables without using an explicit mathematical function.

An artificial neural network is a biologically inspired computer program designed to simulate the way in which the human brain processes information. ANNs are composed of a number of processing elements or units (neurons). Each processing element has weighted inputs, transfer function, and one output. Processing elements are connected with coefficients (weights) and are organized in a layered topology as follows: (i) the input layer, (ii) the output layer, and (iii) the hidden layers between them. The number of layers and the number of units in each layer determine the functional complexity of the ANN.

In this work, modular multilayer feed forward neural networks, trained with a backpropagation-learning algorithm, were used to predict the liquid crystals properties as function of some characteristics. Reasons for the use of this kind of neural network are the simplicity of its theory, ease of programming and good results and because this neural network is a universal function in the sense that if topology of the network is allowed to vary freely it can take the shape of any broken curve.[9]

The NeuroSolution 4.0 program was used for all neural network operations. Unless otherwise stated, the four modular neural networks with a single, two, three or four hidden layers were generated. Modular feedforward networks are a special class of multilayer perceptrons (MLPs). These networks process their input using several parallel MLPs, and then recombine the results (Fig. 1). These networks process their input using several parallel multilayer perceptrons, and then recombine the results. This action tends to create some structure within the topology, which will foster specialization of function in each sub-module. Using modular networks, one needs a smaller number of training examples for the same size network (*i.e.*, the same number of input variables). The hidden layers act like feature detectors. Different activation functions applied to the hidden layer detect different features as pattern processes through a network. Combining the results leads to better prediction. When each slab has different activation functions it offers two, three, four and five ways of viewing data. The output layer receives multiple different views of data features as detected in the hidden slabs plus the original inputs.

EXPERIMENTAL

The combination of different structural units in a molecule gives rise to physical properties which are very important when designing new liquid crystals. For practical use, the materials should not only have the molecular structure suitable for inducing liquid crystal properties, but also an appropriate combination of physical properties for that application. The factors influencing the molecular unit are varied and include core units, connecting groups, terminal groups, lateral groups and lengths of flexible chains. All these structural factors affect the nature of interactions between liquid crystalline molecules and are very important for obtaining the adequate mesomorphic behavior. The organic compounds used in this paper have similar structures with small structural changes that allow a systematical analysis of the factors that influence liquid crystals properties and determination of some parameters that will be used in prediction with neural networks. We used an organic compounds database (32 compounds), which includes a wide variety of symmetrically derivatives with two ferrocene units. The general structures of the analyzed compounds are presented in Fig. 2.

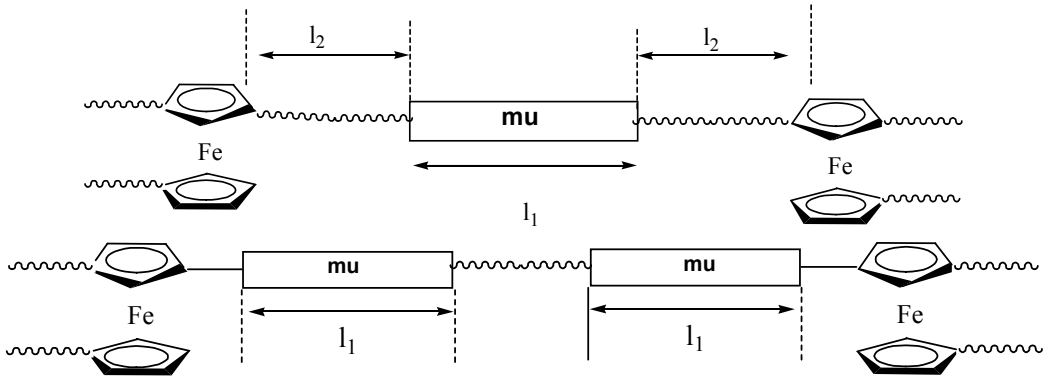


Fig. 2 – The general structures of the ferrocene derivatives.

The establishment of the numerical inputs for neural models (*molecular descriptors*) is a critical and difficult problem. This is due to the fact that the molecular descriptors must represent the molecular structural features related to the properties of interest as distinctly as possible. The prediction accuracy of the neural networks depends heavily on the amount of correction between the molecular descriptors and the structural features. We used as topological and geometrical molecular descriptors: molecular weight (**M**), polarizability (**P**), total length ($\mathbf{L_{tot}}$), and ratio volume/surface (**V/A**). The molecular descriptors were estimated by mechanical molecular simulation using Hyperchem program. The output variable for the model is represented by the liquid crystalline behaviour of the compound.

The molecular modeling studies (molecular mechanics) were carried out using the HyperChem 7.5 software package. Submitting a structure to a calculation can be expensive in terms of human time and effort. HyperChem lets us build and display molecules easily. Since HyperChem contains a graphical interface, we can monitor the construction of molecules. Using the Drawing tool, we can draw a two-dimensional (2D) representation of a molecule, and then use the Model Builder to generate a three-dimensional (3D) structure. The Model Builder adds implicit hydrogen atoms to the molecule at our request. We can also manipulate individual bonds, bond geometries, angles, torsions, and atomic charges during model building or after model building. HyperChem contains a database of amino and nucleic acid residues so we can quickly build compounds containing these subunits.

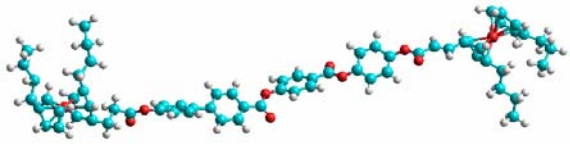
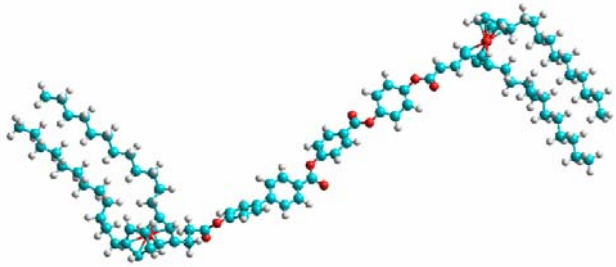
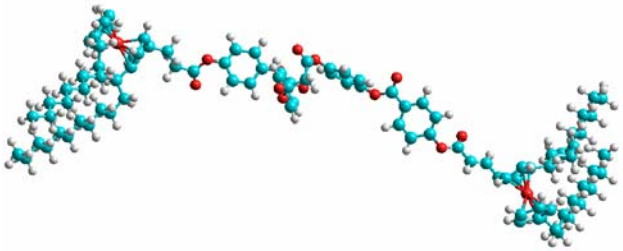
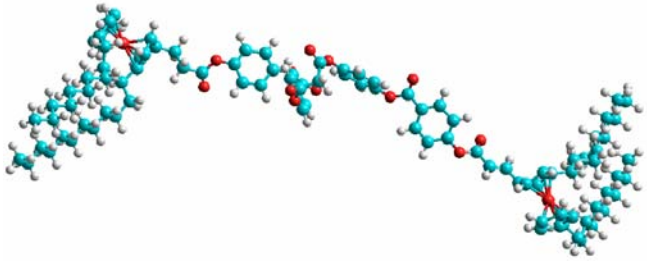
To calculate the molecular descriptors[10] we need to generate a well-defined structure. A calculation often requires a structure that exhibits a minimum on a potential energy surface. HyperChem contains several geometry optimizers to do this (Polak-Ribière, Steepest Descendent or Fletcher-Reeves (Conjugate gradient), optimizer RMS (Gradient) of 0.1 kcal/mol.Å). We can then calculate single point properties of a molecule or use the optimized structure as a starting point for subsequent calculations, such as molecular dynamics simulations.

The Molecular Mechanics (MM+) force field was applied for preliminary structure optimization and study of the conformational behavior of each organic compound. Molecular mechanics has been shown to produce more realistic geometry values for the majority of organic molecules owing to its highly parameterized nature.[11]

The experimental results of HyperChem calculations for the several organic compounds analyzed are show in Table 1.

Table 1

Values of 3D and molecular descriptors

Sample cod (S)	**Three-dimensional (3D) structures for several organic compounds**	**M (a.m.u.)**	**L_{tot} (Å)**	**P**	**V/A**
7		1129	45.3	114.2	2.1
8		1578	52.3	172.9	2.0
9		1708	49.8	183.5	2.0
12		1680	50.2	179.9	2.0

RESULTS AND DISCUSSION

The neural network modeling implies the following stages: collecting the training data by experiments, making up the training and testing data sets, developing the neural network topology, training and, finally, establishing the performance of the neural network model by comparing the network prediction to unseen (validation) data. The success in obtaining a reliable and a robust neural network depends strongly on the choice of the process variable involved, the available data set and its domain used for training, as well as the training method.

The parameters considered as inputs of the neural model were: sample code (**S**), molecular weight (**M**), polarizability (**P**), total length ($\mathbf{L_{tot}}$), and ratio volume/surface (**V/A**) evaluated using molecular modeling simulation. Concerning the liquid crystal behavior (LC), we have coded with "1" the possibility to generate a mesophase and with "0" the crystalline or amorphous phases. This is the symbolic output variable for neural model.

Firstly, the experimental data were split into training and validation sets, with **25** and **7** experimental points, respectively. We developed and trained many modular neural networks and then we selected the best one that balances the size and the performance. A number of different modular neural network architectures (Fig. 1) with **4** inputs, a single, two or three hidden layers of **4**…**80** hidden neurons and **1** output were used. A layered feed-forward neural network was chosen because it may be the simplest type which is able to model systems with different degrees of complexity and nonlinearity. The best network topology was determined based upon the mean squared errors (MSE) on the training data. So, the training is considered terminated at the point where network error (MSE) becomes sufficiently small (<0.02).

Table 2 contains different feed forward topologies tested with selected training data and the main performances for these networks: MSE (Mean Squared Error), r (correlation) and E_p (percent error).

Good predictions are obtained with every modular neural models on training data because they have similar performances. Table 3 presents an example for the 2MLP(4:50:1) (modular neural network – **b**). Two columns named "LC net" appear in table 3: the first contains the real predictions of the network and the second, "LC net rounded", the rounded result corresponding to the experimental data.

One of the most important features of a neural network is the generalization capability. A network is said to generalize well when the input-output relationship, found by the network, is correct for input/output patterns of validation data, which were never used in training the network (unseen data). The good predictions of the networks on validation data are exemplified in Tables 4 and 5. The probability of a correct answer was 57.14 % for 2MLP(4:60:1) (modular neural network – **a**) (that means 4 predictions are correct from all 7 validation data) and 71.42 % for 2MLP(4:60:1) (modular neural network – **d**) (5 correct predictions). Wrong predictions are marked in grey in Tables 4 and 5.

The reason of choosing the neural networks with one hidden layer and 50 and 60 hidden neurons was the simplicity of the topology and the good performances in the training and validation phases. There are other networks in Table 2 with good performances, the best being a more complex neural network, 2MLP(4:60:60:60:1), with MSE = 0.0165, r = 0.990 and E_p = 2.01. A comparison between the predictions of the neural networks with one, two or three intermediate layers (given in table 2) does not motivate the choice of a complicated neural network topology. Thus the main criteria of choosing the neural model were the predictions on training and validation data and the compromise between performance and complexity.

Table 2

Different topologies tested for feed forward neural networks

No.	Type of modular neural network	Network topology	MSE	r	E_p (%)
1.	Two hidden slabs with possibility of two activation functions (**a**)	2MLP(4:60:30:10:4:1)	0.0199	0.988	2.22
2.	Two hidden slabs with possibility of two activation functions (**a**)	2MLP(4:60: 60:60:60:1)	0.0198	0.878	2.34
3.	Two hidden slabs with possibility of two activation functions (**a**)	2MLP(4: 60: 60:60:1)	0.0165	0.990	2.01
4.	Two hidden slabs with possibility of two activation functions (**a**)	2MLP(4:60:1)	0.0199	0.987	3.28

Table 2 (continues)

Table 2 (continued)

No.	Type of modular neural network	Network topology	MSE	r	E_p (%)
5.	Two hidden slabs with possibility of three activation functions and a jump connection (**b**)	2MLP(4:60:60:1)	0.0198	0.988	2.45
6.	Two hidden slabs with possibility of three activation functions and a jump connection (**b**)	2MLP(4:50:50:1)	0.0199	0.988	2.61
7.	Two hidden slabs with possibility of three activation functions and a jump connection (**b**)	2MLP(4:50:50:50:1)	0.0199	0.988	2.33
8.	Two hidden slabs with possibility of three activation functions and a jump connection (**b**)	2MLP(4:50:1)	0.0199	0.987	2.74
9.	Two hidden slabs with possibility of four activation functions and a jump connection between hidden layer and output layer (**c**)	2MLP(4:50:50:50:1)	0.0199	0.988	2.51
10.	Two hidden slabs with possibility of four activation functions and a jump connection between hidden layer and output layer (**c**)	2MLP(4:50:50:1)	0.0150	0.990	3.10
11.	Two hidden slabs with possibility of four activation functions and a jump connection between hidden layer and output layer (**c**)	2MLP(4:50:1)	0.0199	0.987	3.00
12.	Two hidden slabs with possibility of five activation functions and a jump connection between hidden layer and output layer (**d**)	2MLP(4:50:50:50:1)	0.0199	0.988	2.44
13.	Two hidden slabs with possibility of five activation functions and a jump connection between hidden layer and output layer (**d**)	2MLP(4:50:50:1)	0.0199	0.987	2.60
14.	Two hidden slabs with possibility of five activation functions and a jump connection between hidden layer and output layer (**d**)	2MLP(4:50:1)	0.0199	0.987	2.92

Table 3

Prediction of 2MLP(4:50:1) (modular neural network – **b**) on training data

M	L_{tot}	P	V/A	LC exp	LC net	LC net rounded
947	41.8	88.7	2.4	0	0.000008	0
1778	50.3	192.7	2.0	1	0.994136	1
2079	49.1	223.6	2.1	1	0.999305	1
750	30.8	66.5	2.4	0	-0.000036	0
722	30.0	66.3	2.4	0	0.000034	0
1570	44.7	165.4	2.1	0	0.000032	0
1737	48.4	187.2	2.1	1	0.999066	1
1736	49.5	187.2	2.0	1	1.008743	1
1542	33.6	161.7	2.2	0	0.001325	0
1518	50.8	165.3	2.0	0	-0.000047	0
1821	47.3	198.2	2.1	1	1.000137	1
1880	50.8	203.1	2.0	0	0.00003	0
2068	41.3	221.8	2.2	1	1.000022	1
1714	50.1	183.8	2.0	0	0.001609	0
1487	37.0	156.8	2.1	0	0.000028	0
858	38.5	83.6	2.5	0	-0.000007	0
1708	49.8	183.5	2.0	1	0.998513	1
2221	44.8	240.6	2.1	1	0.99894	1
1680	50.2	179.9	2.0	1	0.998941	1

Table 3 (continues)

Table 3 (continued)

M	L_{tot}	P	V/A	LC exp	LC net	LC net rounded
2102	53.2	228.1	2.0	1	0.999767	1
2120	52.3	227.6	2.1	1	1.002162	1
1672	40.0	168.9	2.2	0	-0.001618	0
1578	52.3	172.9	2.0	1	1.000722	1
1980	45.8	216.2	2.1	1	0.999988	1
1129	45.3	114.2	2.1	1	1.000221	1

Table 4

Prediction of 2MLP(4:60:1) (modular neural network – **a**) on validation data

M	L_{tot}	P	V/A	LC exp	LC net	LC net rounded
1862	46.4	203.7	2.1	0	1.055555	1
1099	48.1	108.0	2.4	1	-0.055552	0
1099	48.6	108.0	2.4	0	-0.055553	0
1838	48.6	199.5	2.1	1	1.055229	1
1780	51.1	192.9	2.0	1	0.433418	0
1738	51.4	187.4	2.0	0	0.030419	0
1896	43.1	198.2	2.1	0	-0.055554	0

Table 5

Prediction of 2MLP(4:60:1) (modular neural network – **d**) on validation data

M	L_{tot}	P	V/A	LC exp	LC net	LC net rounded
1862	46.4	203.7	2.1	0	1.055556	1
1099	48.1	108.0	2.4	1	-0.055556	0
1099	48.6	108.0	2.4	0	-0.055556	0
1838	48.6	199.5	2.1	1	1.055553	1
1780	51.1	192.9	2.0	1	0.708862	1
1738	51.4	187.4	2.0	0	0.182334	0
1896	43.1	198.2	2.1	0	-0.055556	0

The results obtained in validation phase prove that modular neural network can be used to make predictions for different compounds (different structures and properties), substituting the experiments that are time and material consuming.

CONCLUSIONS

The prediction of the mesophase occurrence with machine learning methods (modular neural networks), as well as the choice and the codification (numerical and nominal) of different sets of parameters which characterize the behavior of some symmetrically derivatives with two ferrocene units represents a new approach in the field. The efficiency of modular neural networks with different types of connections between processing elements, seldom used in properties prediction, was demonstrated. The methodology exemplified here is relatively simple, provides accurate results in estimation of crystal liquid behavior and has a general character, so that it could be applied to other chemical systems.

Acknowledgements: This research occurred in the framework of the Project PN II ID_592/ no.59/1.10.2007 for which the authors acknowledge financing authority.

REFERENCES

1. C. T. Imrie, C. Loubser, P. Engelbrecht, C. W. McCleland and Y.F. Zheng, *J. Organomet. Chem*., **2003**, *48*, 665.
2. a) D. Apreutesei, G. Lisa, H. Akutsu, N. Hurduc, S. Nakatsuji and D. Scutaru, *J. Appl. Organomet. Chem.*, **2005**, *19*, 1022; b) unpublished results.
3. D. Apreutesei, D. Scutaru and D. G. Mehl, *Romanian International Conference on Chemistry and Chemical Engineering RICCE XIV*, vol. Proceedings, **2005**, 14.
4. P. Massiot, M. Imperor-Clerc, M. Veber and R. Deschenaux, *Chem. Mater.*, **2005**, *17*, 1946.
5. J. Homer, S. C. Generalis and J. H. Robson, *Phys. Chem. Chem. Phys.*, **1999**, *1*, 4075.
6. Z. Zhang and K. Friedrich, *Compos. Sci.Technol.*, **2003**, *63*, 2029.
7. M. S. Al-Haik, M. Y. Hussaini and H. Garmestani, *Int. J.Plasticity*, **2006,** *22*, 1367.
8. T. Kleinoder, S. Spycher and A. Yan, "Prediction of properties of compounds, in Chemoinformatics – A Textbook", Wiley-VCH, Weinheim, 2003.
9. L. A. Madsen, T. J. Dingemans, M. Nakata and E. T. Samulski, *Phys. Rev. Lett.*, **2004**, *92*, 145505.
10. R. Todeschini and V. Consonni, "Handbook of Molecular Descriptors", R. Mannhold, H. Kubinyi, T. Timmerman (Eds.), Wiley-VCH: Weinheim, New York, 2000**.**
11. D. Young, "Computational Chemistry: A Practical Guide for Applying Techniques to Real World Problems", John-Wiley & Sons: New York, 2001.