PERGAMON

Neural Networks 12 (1999) 767–775

Neural Networks

Contributed article

A new perspective in learning pattern generation for teaching neural networks

G.-P.K. Economou[a,*], D. Lymberopoulos[a], K. Spiropoulos[b], P.D. Goumas[b]

[a]*Department of Electrical and Computer Engineering, University of Patras, Patras, Achaia, Greece*
[b]*Regional University Hospital of Patras, University of Patras, Patras, Achaia, Greece*

Received 3 January 1997; accepted 4 January 1999

Abstract

This article deals with the Learning Patterns (LPs)' generation, a major aspect of Feed-Forward Artificial Neural Networks (FANNs)' learning process. Currently, more work is done to understand the mechanisms and improve the speed, learning accuracy, and implementation features of FANNs' teaching algorithms, though little is done towards the development of enhanced techniques that would extract experts' knowledge (from examples, rules, etc.) and obtain standardised LPs that would improve this learning process. A new approach in generating LPs is thereby introduced, that is used to train a new Medical Decision Support System (MDSS) based on FANNs, and its performance is analysed and compared with previous methods. It can handle incomplete data archives, individually boost any particular dataum's special characteristics, and its application induces the FANNs to show better convergent facets. The efficiency of the resulting MDSS was thoroughly tested by pulmonologists and haematologists using medical data archives of a regional hospital. © 1999 Elsevier Science Ltd. All rights reserved.

Keywords: Medical decision support systems; Feed-forward artificial neural networks and their learning process; Learning patterns' generation

1. Introduction

A Medical Decision Support System (MDSS) based on an architecture of Feed-Forward Artificial Neural Networks (FANNs) was developed recently and successfully applied to the whole spectrum of pulmonary and haematological diseases (PDs, HDs) (Chen, Ke & Chang, 1990; Economou et al., 1994a; Economou, Mariatos, Economopoulos, Lymberopoulos & Goutis, 1994b), in order to assist the process of completing medical diagnoses. Physicians face particular problems that introduce difficulties into the gathering and evaluation of medical data that include dealing with large number of patients who may be of various educational levels, thus augmenting the controversy of input medical data; selecting through complicated and incomplete data and assessing their value; working under time pressure; having to constantly improve their skills; facing the hazards of transmitted diseases, often being isolated from organised medical centres. In order to cope with these, a number of innovations were introduced to the MDSS: the integration of features of classic Expert Systems and FANNs (Economou et al., 1996); its explicit founding on the Clinical Differential Diagnosis Methodology (CDDM); its VLSI design in FPGA-based chips by approximating the FANNs' sigmoid function, implementing proper arithmetic operation modules, and dynamically managing and storing node weights (Economou et al., 1994b).

In addition, a new approach to extract and structure the FANNs' LPs using physicians' experience was devised, in order to better structure the system. It should be noted that the difficulty in building LPs is mainly a good understanding of the way this experience is conceived. More specifically, physicians often put together rules and make associations of interrelated data features involved in their particular field (i.e. associating symptoms to their findings) that cannot be described by largely accepted standards or through examples (Mulsant, 1990). Moreover, they can judge upon intricate cases without consciously perceiving and consequently describing the existing correlation between all (medical) data (Chen, Ke & Chang, 1990; Poli, Cagnoni, Livi, Coppini & Valli, 1987).

There are also diseases that cannot be identified by neither the cases related to the exhibited symptoms, nor the outcome of the *physical examinations* (PEs), but rather by the *clinical examinations* (CEs) that are performed upon a patient. Thus, the novel approach was utilised for the integration of all symptoms, PE and CE data within the

* Corresponding author. Tel./fax: + 30-61-99-61-85.
E-mail address: economou@ee.upatras.gr (G.-P.K. Economou)

PII: S0893-6080(99)00021-0

Table 1
Dependency of PD classes against 'Cough' symptom and its findings[a]

PD classes	Cough's findings dependencies																
	Dep	Rcn	Chr	Prd	nPrd	Prx	Exr	Day	Mrn	Evg	Ssn	Anx	Swt	Wg −	*Wg +	Vmt	Slp
COPD	****	*	*	*			*	*	*			*	*	*	*		*
Tuberculosis	****	*	*		*	*		*				**	**	*		*	
Interstitial PDs	****	*	*		*	*		*			*						
Abnormalities of the diaphragm																	
Cancer of the lungs	****	***	***	*	*	*		*				*		*			
Disorders of the mediastinum	**	*			*	*		*	*	*		*		*			
Infection diseases of the lungs	****	*		*	*	*		*				*					
Disorders of pleura	*	*			*			*		*		*		*			
Bronchial asthma	**	*	*	*	*	*	***	*		*	***						
Disorders of the pulmonary circulation	*	*		*	*	*		*				*					
Occupational disorders of the lungs	***		*	*	*	***		*	***	*							
Non PDs																	

[a] COPD, Chronic Obstructive Pulmonary Disease; Dep, Dependence of PD class to Cough; Rcn, Recently exhibited Cough; Chr, Chronic Cough; (n)Prd, (non-)Productive Cough; Prx, Paroxismic Cough; Exr, Cough exhibited after Physical Exertion; Day, Cough exhibited all Day Long; Mrn, Cough exhibited in the Morning; Evg, Cough exhibited in the Evening; Ssn, Seasonary Cough; Anx, Cough followed by Anorexia; Swt, Cough followed by Sweating; Wg(±), Cough followed by Weight Loss/Increase; Vmt, Cough followed by Vomit; Slp, Cough followed by Sleepiness.

new MDSS. It supplies for enhanced components LP-generation that were difficult to obtain using previous methodologies, without causing eventual failures to the FANNs' convergence.

2. Medical data and medical decision support system structure

Physicians were asked to supply a typical procedure to obtain, classify, and judge the medical data. They proposed the CDDM (Economou et al., 1994a) as the commonly accepted standard technique for examining patients and gather data on diseases using his/her medical history, asking particular questions, and performing PEs/CEs, to reach to a diagnosis. According to the CDDM, each finding (either answers, symptoms, or examinations) is considered independently and no conclusion is reached before all procured medical data are properly evaluated. The whole rational of the CDDM should lead from the more general to the more specific diagnosis, e.g. first the general classes of possible diseases (i.e. *Disorders of the Pulmonary Circulation*), then the disease(s) (i.e. *Pulmonary Angiitis*).

Standardised tables of historical findings, major symptoms, and physical examinations, against their effect on the PDs/HDs were designed (e.g. 'Cough' symptom, Table 1). As diagnoses are based on the presence and/or absence of particular data and their significance for a disease (i.e. *Dependence* on Table 1), those tables represent a sound mapping of physicians' experience. Two different MDDSs, dealing separately with the PDs or the HDs, were structured. For the PD-MDSS, seven symptoms, four PEs, 15 CEs, and 30 (classification) levels regarding a patient's history, along with their data findings, are the external inputs of the system (a total of 142 inputs) (Economou et al., 1994a). For the HD-MDSS, the respective numbers are 12 symptoms, 4 PEs, 25 CEs, and 43 classification levels (a total of 172 external inputs).

Table 1 shows medical data for the *Cough* symptom related to the 12 PD classes (Economou et al., 1994b).

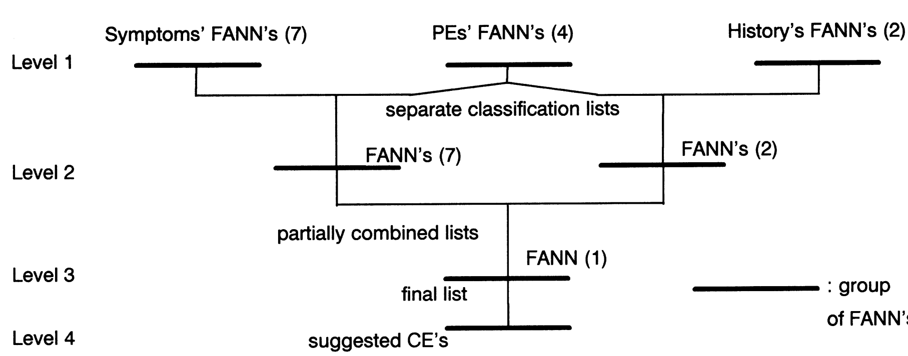

Fig. 1. Structure of Layers 1 and 2 of the new MDSS; the number of FANN's/level is given in brackets.

Table 2
Number of actual against physicians-created cases used to teach the PD FANNs

Classes of PDs	Actual cases	Physicians-created cases
COPD	48	2
Tuberculosis	11	39
Interstitial PDs	12	38
Abnormalities of the diaphragm	11	39
Cancer of the lungs	11	39
Disorders of the mediastinum	9	41
Infection diseases of the lungs	27	23
Disorders of pleura	13	37
Bronchial asthma	45	5
Disorders of the pulmonary circulation	13	37
Occupational disorders of the lungs	7	43
Non PDs	3	47
Sum	210[a]	390

[a] As a result of few PDs interpolation.

Column 1 of the table denotes these *Classes* (plus a non-PD one); column 2, the *Dependence* factor relating this symptom to each PD class; the remaining columns, whether the findings are *Recent*, *Chronic*, *Productive*, *non-Productive*, or/and *Paroxismic*; present after *Exertion*, all *Day* long, only in the *Morning*, or/and in the *Evening*; being *Seasonal*; followed by *Anorexia*, excessive *Sweating*, *Weight Loss/Increase*, *Vomiting*, or/and *Sleepiness*. The *Dependence*'s weight is given by a series of '*', according to the approximate and qualitative manner of physicians' reasoning; the more the number of '*', the larger the dependency. The absence of any findings implies the possibility of the disease not being a PD one (last row of Table 1). It is possible that two different PDs may have identical inputs for a given FANN, thus forbidding it to differentiate between them. However, the next FANNs in the MDDS structure can do so, because of their inputs. Thus, several pseudo-inputs are added to help the previous FANNs to converge, as the latter ones shall inform the two PDs apart (Section 4).

Medical data are fed into the two MDSS's FANNs which are connected by a three layered scheme that resembles the Artificial Neurons' feed-forward architecture (Rumelhart, Hinton & Williams, 1986; Lippmann, 1987). The first two layers of the system consist of four levels: Level 1 receives all medical data, i.e. a patient's partial symptoms, PEs, and history data in distinct FANNs (Fig. 1). Their outputs are classified lists of, and PDs/HDs (Layers 1 and 2, respectively). In order to make symptom and history data more objective (CDDM), in Level 2, PEs' FANN outputs are combined separately with the output data of the symptoms and history of FANNs. Their outputs are again as stated before. The single FANN of Level 3 accepts the lists of Level 2 and gives an output of the final classes of PD/HD lists. The single FANN of Level 4 receives this list (either of classes of, or diseases) and its output suggests the proper CEs that need to be performed by the patient.

Two blocks of four levels compose Layer 3 of the two MDSSs. Each block is identically structured as the FANNs in Layer 2, with the difference that Block 1 combines the results of the CEs in its Level 1 and Block 2 is input of new medical data in its Level 1. The reason for this 're-input' is the elapsed time between the first visit to a physician and the performance of CEs that necessitates a re-examination of a patient when returning to the physician (thus new medical data). The two blocks only deal with lists of PDs/HDs, as the MDSSs already converge, and their final classification are compared to consider more CEs or to suggest the proper medical treatment (medications and their dosages).

A total of 200 PD and 150 HD actual cases were used for learning/testing purposes in varied groups of 140/60 and 100/50 sets. As patient archives showed insufficient coverage of all PDs and HDs, physicians were asked to use their expertise to create model cases that would well supply missing ones, adjudicating qualitatively 50 different cases per disease for all PDs/HDs. On the other hand, according to the principle adopted by the collaborating physicians in the project, 'there are no easy or hard to diagnose diseases but there are patients'. Table 2 gives the number of physicians created against actual cases for the PD patient cases. Similar numbers stand true for the HD cases and their respective 11 classes.

It is clear that some PDs would be better 'represented' through actual cases should their data be taught to any MDSS. However, as the pulmonologists and haematologists involved in the systems' development supplied valid physicians-created cases to equal the teaching set, such a problem did not occur. In contrast, it is a standard practice to teach novice (internal) physicians both through theoretical examples and laboratory set experiments, and by having them treating real-world patient cases in a hospital ward. As a result, they learn by both studying books and by treating patients, a practice repeated in the adopted learning cases used for the new MDSS. In addition, physicians-created cases are always the ones based on their valuable experience on the PD/HD field.

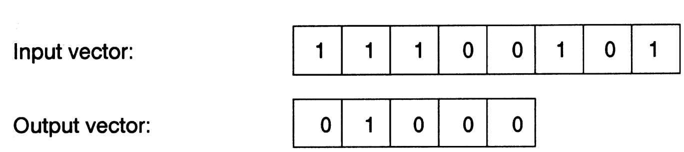

Fig. 2. A pair of input and output LP vectors (previous methodology).

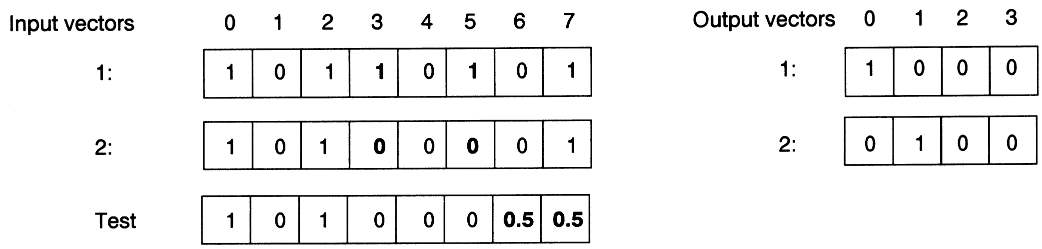

Fig. 3. Similar input learning patterns—A test pattern.

3. Previous methodology for the generation of learning patterns

During a FANNs' learning process, a one to one and on to mapping of a vector of inputs and a vector of outputs is performed. A literature-standard input vector (Rumelhart et al., 1986; Lippmann, 1987), consists of series of logical '1's and '0's, denoting the presence or absence of a component, while the output consists of one logical '1' followed by a series of logical '0's that fill the rest of the vector, denoting a matching of only this class of outputs to the input vector. Thus, the interrelation of logical '1's and '0's (components) and the input to the output LPs (vectors), is critical to the teaching and the identification of patterns. As a result, a slight diversity from an already existing input LP that leads to a different output, is fed by using another pair of LP; otherwise, the FANN could fail to classify them correctly. Fig. 2 shows an LP pair, built according to Rumelhart et al. (1986) and Lippmann (1987).

Consequently, the input vectors that map one to one and on to a particular output have to be structured very strictly. Also, as this previous LP generation could also be used for a particular finding, symptom, or disease mapping, a slight variation of an input LP could result in a totally different output as a result of the FANN's operation, a fact that is not acceptable by the medical point of view. In addition, through using these LPs is not possible to boost a particular LP component without altering their interrelation with the other components. As a further step, three LPs built according to Rumelhart et al. (1986) and Lippmann (1987) are considered as an example (Fig. 3); these LPs are illustrative sub-sets of the actually fed ones.

In Fig. 3, the first two input vectors differ only in the third and fifth component; thus their Hamming distance is 2. Consequently, should a FANN be taught with the two input vectors and later tested by the insertion of the first input vector, both its outputs would be set to a logical '1', the first to be 100% and the second to be 80% of its numerical representation (experimental data; the FANN's arithmetical convergence error is not taken into consideration). The difference in the numerical representation of the two outputs is not as large as it is perceived by physicians, given the different learning input vectors, hence posing a problem.

Also, the medical archives are often not compiled to their full extent as the personality of physicians is involved in the classification and description of the patients' symptoms and findings (i.e. they 'compress' information in a personal code-like manner) (Mulsant, 1990; Chen et al., 1990; Poli et al., 1991; O'Kane, 1988). These need rendering, should the appropriate physician(s) be out of duty or transferred elsewhere. Nevertheless, the problem of using old medical archives on which possible new-sprung symptoms are not accounted, remains. The substitution of unknown data by mapping them into logical '0's (Chen et al., 1990), according to a physician's expertise (Poli et al., 1991), or with statistically elaborated ones (Kampschöer et al., 1989), was considered inappropriate for this project.

To solve this problem, several researchers propose the use of the logical '0.5' as a *don't care* or *don't know* term (Mulsant, 1990) when new data is input (after the FANNs' teaching). If this be the case, the same FANN fed by the test input vector of Fig. 3, would have its two outputs set to logical '1's up to the 65 and 80% of their numerical representation (experimental data; convergence errors are again not considered). Once again, physicians considered this inadequate, as it alters the significance of the input component.

The order of the output classification formed another problem. In Table 1, the *Cough* symptom uses 16 findings for the correct classification of the 12 PD classes. Supposing

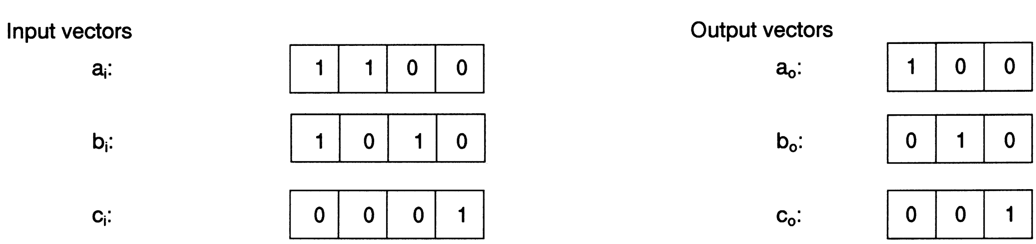

Fig. 4. Three input/output vectors of the previous LP-generation approach.

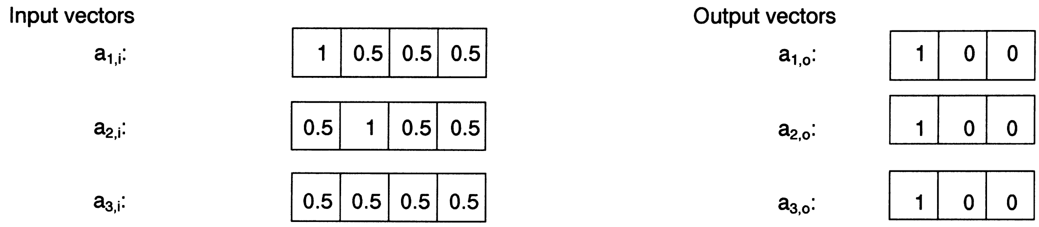

Fig. 5. New representation of ‘a’ input vector (first intermediate format).

there are at least two input LP vectors that have some logical ‘1’s and ‘0’s combined in a (partly) similar manner, what ought to be the output of a FANN should a test pattern consisting of only their common components’ combination be fed into it? What about the similarity of outputs triggered by input vectors with no components at all?

As for the first case, it is often encountered, i.e. some LPs being ‘sub-sets’ of larger ones with respect to the number or positions of their ‘1’s. According to the FANN learning operation, all outputs corresponding to these ‘similar’ learning inputs, would be set to logical ‘1’. However, the taught LP with the smaller number of different logical ‘1’s to those of the test input, through a FANN’s generalising operation (after their teaching procedure), will *make* its attached output to have a numerical representation of this logical ‘1’ greater than the numerical representation of the other LPs’ one (Rumelhart et al., 1986; Lippmann, 1987). As for identical input LPs (with the same number of ‘1’s or totally devoid of them), will make the FANN(s) not to converge.

This performance had to be expected as the formation and partition of the hyper-space defined by the use of the previous LP building, is a matter of constructing distances from a LP’s components. The more the logical ‘1’s in an LP, the greater the Hamming distance of a test input vector from it. However, this might lead to a faulty result as medical teams diagnose a disease according to ‘the special characteristic of its symptoms’ findings (attributes) and not the number of them’.

4. Proposed methodology for generating learning patterns

The new LP-generation methodology is specially made to enhance the *independent input vector* components’ features, contrary to the previous approach which promotes a *single-class output vector*. It is described by a number of rules, deduced from FANNs’ teaching- and the CDDM application-experiments.

Rule 1: The size of input vectors is decided by enumerating each distinct symptom’s findings. All input vectors have only one logical ‘1’ to mark the presence of only one component (*finding*), setting all the other elements of the input vector to logical ‘0.5’s. Hence, *each* input component of the previous approach will generate *one* input LP. This preserves the individual finding’s value in forwarding a class of, or disease.

Rule 2: The number of the input LPs per given output equals the maximum number of present findings of the previous approach. Outputs that lack this norm are supplied with a number of *pseudo-findings* (via pseudo-inputs) that will be set to the logical ‘1’, and augment the size of input vectors. Pseudo-inputs enhance the data equilibrium between LPs (aiding FANNs to converge), and, on the contrary, they could be proven to be new medical data that ought to be considered and utilised as medicine progresses.

Rule 3: The output vectors are compiled by setting all their components that correspond to the firing output neurons (because of an input vector’s component) to the logical ‘1’. Thus, every finding is mapped to a *set of outputs*, and each disease is directly attached to a particular finding.

Rule 4: Logical ‘1’s of the proposed LPs can be replaced by *weighting factors* or attributes (if supplied) that induce the significance of a sole finding to its output(s) and boost specific LP mappings.

Rule 5: Logical ‘0’s are used only if the *absence* of some findings needs to be pointed out, or else, they are implied by the lack of logical ‘1’s. Moreover, even a small evidence of a likely output (disease), due to the presence and not the absence of a finding, is crucial in medical decision support systems.

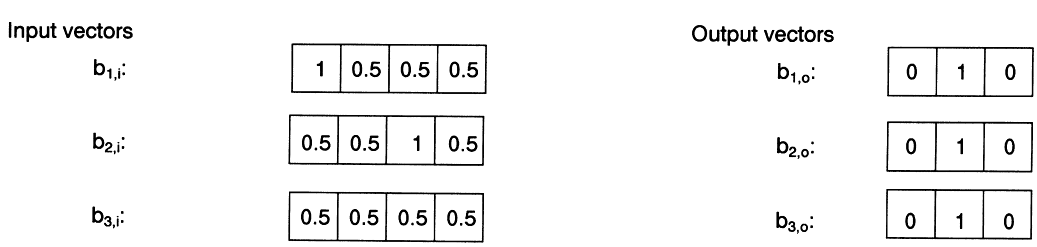

Fig. 6. New representation of ‘b’ input vector (first intermediate format).

Input vectors					Output vectors			
$c_{1,i}$:	0.5	0.5	0.5	1	$c_{1,o}$:	0	0	1
$c_{2,i}$:	0.5	0.5	0.5	0.5	$c_{2,o}$:	0	0	1

Fig. 7. New representation of ‘c’ input vector (first intermediate format).

Rule 6: An *input vector* full of logical ‘0.5’s is fed and mapped one to one on to an output vector full of logical ‘1’s, signifying that in case of doubt all possible classes of diseases should be promoted.

The aforementioned rules are clarified later, considering the input/output LPs of Fig. 4:

Each of the LP input vectors will span to a number of new input vectors according to the number of their findings (Rule 1). Another input vector full of logical ‘0.5’s will be added (Rule 6; Figs. 5–7):

As the maximum number of input logical ‘1’s (per input LP) with regard to all LP’s is 2 and the *c* input vector has only one logical ‘1’ (as shown in Fig. 4), one pseudo-input has to be added to all the input vectors (Rule 2). Note that this addition is made only to the input LPs and not to the input patterns that are fed to the FANNs after their learning process’ end. A more precise rationale behind the use of this rule is that when the MDSSs generalise into unknown inputs, the Hamming distances between learnt LPs and these unknown vectors have to be calculated only between existing components and not between the pseudo ones. The pseudo-findings only help to create more Hamming distances between given LPs, but not to actually differentiate between them, especially in the case of ‘similar’ pairs of input/output vectors which are the outcome of using the previous LP-generation approach. These inputs that are used for the MDSS testing or generalisation purposes, will be structured according to the previous building of input vectors. Figs. 8–10 display with detail the intervention suggested by Rule 2:

As a result, all the initial input vectors were mapped *one to one* and *on to* the new LPs and thus the output LPs’ components are triggered by the same number of logical ‘1s’ per input finding. In addition, all new input vectors comprise a pseudo-input (in order to satisfy teaching needs) and logical ‘0.5’s complete the vectors. The important feature is the presence of only one logical ‘1’ per input LP, the major characteristic that distinguishes the proposed from the previous LP-generation methodology.

The final LPs are produced (Rule 4), by considering the significance each finding has for each class of outputs (class of diseases, diseases, etc.), shown in Table 1 (Economou et al., 1996). This dependence is included as weighting factors ranging from 1 to 4 (corresponding to the number of ‘*’) while building the PD- and HD-MDSSs. Nevertheless, it would be too detail to illustrate these factors here, particularly as no attribute files are related to these exemplified *a, b*, and *c* LPs.

Fig. 11 shows the final input/output vectors for the new LP generation, after Rule 3 has taken effect. The example given here does not require the use of Rule 5.

5. Overall analysis and comparative results

A number of learning and testing experiments were conducted in order to establish the efficiency of the rules given in Section 4. As already discussed, the 200 PD and 150 HD cases were divided into varied batches of 140/60 and 100/50 learning/testing case sets. These cases were randomly selected, except for the first experiment batch; its cases were chosen by equally distributing diseases per case, thus obtaining a more balanced learning space. A total of 20 experiments were conducted using a combination of the back-propagation algorithm and the Kalman filter error prediction equations (selected after extensive trials) (Economou et al., 1994a). Their testing results are shown in Table 3 for PDs and in Table 4 for HDs.

The data on these tables show the percentages with which the MDSSs correctly classified the PD/HD test cases; column 1 illustrates the methodology implemented to generate the LPs. First the previous representation both with and without the addition of ‘don’t care/don’t know’ terms, and then the results of the proposed methodology are shown,

Input vectors						Output vectors			
$a_{1',i}$:	1	0.5	0.5	0.5	**0.5**	$a_{1',o}$:	1	0	0
$a_{2',i}$:	0.5	1	0.5	0.5	**0.5**	$a_{2',o}$:	1	0	0
$a_{3',i}$:	0.5	0.5	0.5	0.5	**0.5**	$a_{3',o}$:	1	0	0

Fig. 8. New representation of ‘a’ input vector (second intermediate format).

Fig. 9. New representation of ‘b’ input vector (second intermediate format).

again with and without the use of the pseudo-inputs and the supply of an input LP formed of a series of ‘0.5’s. NC stands for no convergence.

A convergence did not occur when one or more diseases happened to have findings that were ‘similar’ (Section 3) to the findings of other diseases and appeared in fewer cases in the specific batch; this was resolved using the proposed methodology. Also, the addition of the ‘don’t know/don’t care’ terms, certainly improved the performance of the MDSSs taught with the previous LP-generation approach. This fact along with the addition of the other steps of the new methodology, greatly improved the testing results, which is attributed to a better partitioning of the learning space (Economou et al., 1996). In addition, the experiments proved that this methodology always led the FANNs to converge to the requested learning error (1%), this error appears to be smaller given the same number or epochs, and to drastically reduce the learning time. LP generation by means of the previous approach sometimes prevented the convergence to a smaller error.

All experiments were conducted by writing software that conforms to the ANSI-C specifics. Typical learning parameters included FANNs of 4–146 inputs (a total of 741 inputs), 12–35 outputs (a total of 792 outputs), 0–0.54 hidden artificial neurons (a total of 288 hidden neurons), 18–235 total neurons (a total of 2077 neurons), and 32–9774 synapses (a total of 18165 synapses). Learning times for a FANN varied between 1 and 5 min., i.e. 2000–3000 learning cycles. Still, when the MDSSs are left to generalise, less than a second’s processing time for the (partial or final) output is required per level.

6. Discussion

The new methodology for the LP generation emerged from the necessity of structuring an MDSS applied in the area of human diseases. The nature of medical data and the model of reasoning from two different teams of physicians led to the conclusion that enhancement of an independent input vector components had to be adopted. With experimentation, the previous LP generation was found inadequate to promote the actual classification of input and output vectors both by not fully exploiting the experts’ specifications and by not always ensuring the FANNs’ convergence (Economou et al., 1996; Mulsant, 1990; Poli et al., 1991).

Using the new LPs, each input component is made to individually excite an appropriate output component, providing a part of its arithmetic value. This value is a function of the number of logical ‘1’s enumerated in the new input vector, their Hamming distances from the already taught LPs’ vectors, and the special attributes already given to the new LPs’ components. The previous LP building rather led an output to the logical ‘1’ or ‘0’ by relating the Hamming distances of LP and input vectors as a whole.

Evenmore, the advantages provided by generating LPs with the new methodology, reflect the physicians’ specification that an ‘MDSS applied to medicine has to always reach a possible (correct) diagnosis, up to a level, than to exclude improbable diseases’ (Poli et al., 1991). In other words, a similarity of an input finding to an already taught one has to always excite an output, as the CDDM imposes. Further, using the previous LP building, the separate component’s weight in exciting an output could not be highlighted.

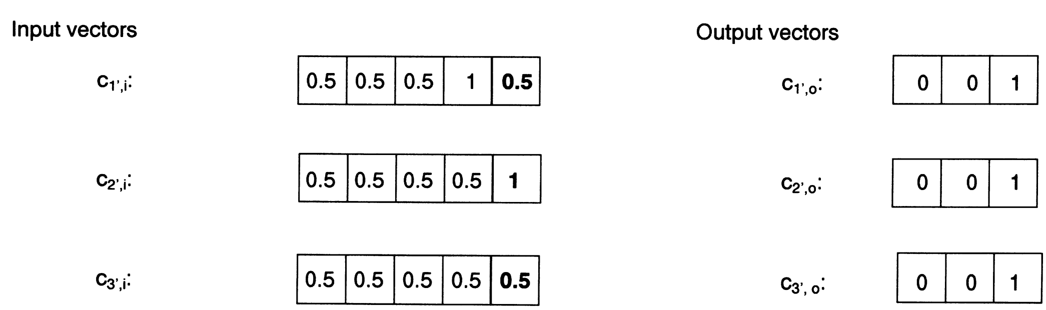

Fig. 10. New representation of ‘c’ input vector (second intermediate format).

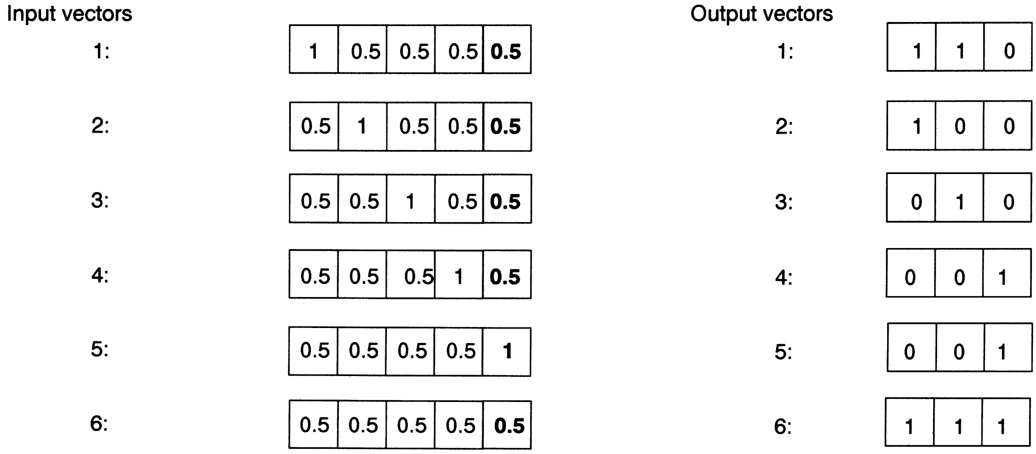

Fig. 11. Final input/output LP's (proposed methodology).

The isolation of each logical '1' or '0' from the rest of the components of an input vector, as the CDDM requires, allows them to be treated as separate entities. The number of separate '1's can then be balanced, adding pseudo-inputs to symmetrically form the sub-regions into which the learnt hyper-space is divided. Additionally, interrelations between components of a previous built LP input vector are not lost, as their contribution to the final output is later on re-assessed in the operating flow of the FANN's structure, as each separate component is made to promote a particular output's component.

An alternative solution for handling medical data would be to divide the actual inputs into groups of more important/frequent findings, thus forming an architecture consisting of trees of linked primary- and secondary-FANNs. This solution would tackle the problem of having learning patterns structured from unknown data (Section 3), as missing findings would simply not be used to form branches of secondary-FANNs when implementing the MDSS's learning procedure. In contrast, these secondary-FANNs could be used at a physician's choice whenever decided as appropriate.

This solution was considered unfeasible because of the large number of connections it requires. Conversely, synapses made by FANNs treating all learning data, grow in lesser numbers as they do not follow a general but strict architectural scheme, but an internal mapping mechanism. Likewise, trees of primary/secondary-FANNs would not

Table 3
PD-MDSS diagnosis efficiency against various LP-generation methodologies

LP generated by	Batches #									
	1	2	3	4	5	6	7	8	9	10
Previous representation	67	65	NC	NC	70	55	NC	78	70	NC
-"- with the addition of '0.5'	70	66	67	NC	73	66	NC	80	75	66
Proposed methodology	85	80	83	82	78	85	80	85	88	80
-"- with the addition of pseudo-inputs	90	85	88	87	85	90	85	90	93	85
-"- with the addition of last pattern composed of '0.5's	93	88	90	88	88	92	88	92	95	88

Table 4
HD-MDSS diagnosis efficiency against various LP-generation methodologies

LP generated by	Batches #									
	1	2	3	4	5	6	7	8	9	10
Previous representation	NC	56	68	72	NC	NC	NC	58	NC	76
-"- with the addition of '0.5'	NC	64	68	76	68	NC	NC	60	62	78
Proposed methodology	82	78	72	84	74	68	84	64	70	82
-"- with the addition of pseudo-inputs	86	84	80	88	78	78	90	76	78	90
-"- with the addition of last pattern composed of '0.5's	88	88	88	90	88	88	95	88	88	94

correlate all important/frequent data, thus denying the full exploitation of FANN's operation, which can reach to stretch out the unknown correlation of input data (Poli et al., 1991).

7. Conclusions

A new methodology for generating learning patterns for the teaching of FANNs is presented. This approach was mainly developed in order to satisfy a number of problems posed by using FANNs to build efficient decision support systems in the field of human diseases. It proved to efficiently overcome the major drawbacks the previous LP-generation posed, while given better generalisation performances.

Further, the new methodology can be applied wherever expertise is not clearly stated, as it generates LPs that provide for the response of the FANNs even when lacking some input vector components, using an inner mechanism for efficiently mapping the actual available data into sub-regions of the hyper-space being formed by the teaching procedure. Also, incomplete medical archives can be used both for training and testing purposes, contributing to the overall MDSS performance.

The new methodology promotes particular input vector components, forcing each of them to specifically contribute to a firing output. The experiments conducted using patient cases from a regional hospital's archives, in the fields of pulmonary and haematological diseases, showed an overall improvement in the FANN's convergence characteristics, while the MDSSs resulted in an efficiency of 88–95%, when being tested and left to generalise into new symptoms' findings they were not taught. Therefore, the novel methodology better fits medical data taken from physicians' expertise into LPs.

Acknowledgements

The authors gratefully acknowledge the support and contributions of the referees and the staff of the respective laboratory and hospital clinics for their creative remarks.

References

Chen, S. M., Ke, J. S., & Chang, J. F. (1990). An effficient algorithm to handle medical diagnostic problems. *Cybernetics and Systems: An International Journal, 4*, 1012–1016.

Economou, G.-P. K., Spiropoulos, K., Economopoulos, N. M., Charakopos, N., Zikos, P., Lymberopoulos, D., & Goutis, C. E. (1994a). Medical decision making systems in pulmonology: a creative environment based on artificial neural networks. *IEEE International Conference on SMC* (pp. 975–980). San Antonio, USA.

Economou, G.-P. K., Mariatos, E. P., Economopoulos, N. M., Lymberopoulos, D., & Goutis, C. E. (1994b). FPGA implementation of artificial neural networks: an application on medical expert systems. *Fourth International Conference on Microelectronics for Neural Networks and Fuzzy Systems* (pp. 287–293). Torino, Italy.

Economou, G.-P. K., Goumas, P. D., & Spiropoulos, K. (1996). A novel medical decision supporting system. *Computing and Control Engineering Journal (IEE Press), 7* (4), 177–183.

Kampschöer, G. H. M., Maruyama, K., Van de Velde, C. J. H., Sasako, M., Kinoshita, T., & Okabayashi, K. (1989). Computer analysis in making preoperative decisions: A rational approach to lymph node dissection in gastric cancer patients. *Br. J. Surg.*, *76*, 905–908.

Lippmann, R. P. (1987). An introduction to computing with NN. *IEEE ASSP Mag.*, *5*, 4–22.

Mulsant, B. H. (1990). A NN as an approach to clinical diagnosis. *MD Comp.*, *7* (1), 25–36.

O'Kane, K. (1988). Computers in the health sciences. *Advances in Computers*, *27*, 211–253.

Poli, R., Cagnoni, S., Livi, R., Coppini, G., & Valli, G. (1991). An NN Expert System for Diagnosing and treating hypertension. *IEEE Comp.*, *24*, 64–71.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). *Learning internal representations by error propagation, parallel distributed processing: explorations in the microstructure of cognition*, 1. Cambridge, MA: MIT Press.