Connectionist Models in Materials Science: Characterisation of the Sorption Properties of Hydrogen Storage Materials

Kuhu Pal (nee Dutta)[1] and Nikhil R. Pal[2]

[1]Pragatinagar, Chinsurah, Hooghly, India; [2]Electronics and Communication Sciences Unit, Indian Statistical Institute, Calcutta, India

We explain how connectionist models can be used in research and development in the areas of materials to make it more productive and useful at much lower cost and time. The basic idea is to identify a computational model *using neural networks to characterise the relation between the output characteristics, input ingredients and process parameters. As an illustration, we focus on the problem of characterising the sorption properties of hydrogen storage materials. We consider the composite materials* $La_2Mg_{17} - x$ *wt% Z with* $Z = LaNi_5$ *and* $Z = MmNi_{4.5}Al_{0.5}$ *for various values of x. We use training data on the desorbed amount of hydrogen for two different temperatures and different time of desorption. These training data are used to train a multilayer network which is then used to predict the amount of released hydrogen for new desorption temperature, and desorption time. Our results show that for both materials with different values of x, the network is able to learn the non-linear desorption characteristics quite successfully. Hence, for different temperatures and desorption time, we are able to predict the dehydriding kinetics and storage capacity of hydrogen storage materials without doing the actual experiments.*

Keywords: Connectionist models; Hydrogen storage material; Sorption property

1. Introduction

With significant advancement in computer technologies (e.g. artificial intelligence, neural networks, database management, computer graphics, virtual reality, etc.), the research and development activities in the area of *computer-assisted materials science* can be made more productive and useful. New computational tools such as Neural Networks (NNs) can be used to identify the inherent characteristics of materials, which can in turn guide the next experiments. Let us illustrate it with an example. Suppose we want to produce tyres with some target specification. The problem is to identify the proportions in which different ingredients like rubber, carbon, sulphur, etc. are to be mixed in order to achieve the desired specification. Apart from the ingredient materials, other factors such as the time of mixing and temperature at which mixing is done may also be important to determine the final characteristics of the mixed material. This is not an easy problem, as the nature of the highly non-linear relation between the percentage of different input materials and other process parameters is not only unknown, but it is also difficult to find an explicit relation for it. Consequently, the researchers conduct many experiments based on experience to get the desired material. There are many other materials science problems, which can also be viewed as finding an input-output relation like this. In such a case, if we could find a *computational model* describing the input-output relation implicitly, our task would be much easier. Connectionist models (neural networks) as well as fuzzy logic can indeed be used to realise this objective. In fact, this could be a highly promising area of materials science, and

we hope that it will lead to co-operative research endeavour between computer scientists and materials scientists in the use of powerful new computational models to result in significant advances in materials science. We can call this new area *computational materials science*. In this paper, we consider how neural networks can be used to characterise the desorption properties of different composite hydrogen storage materials.

Because of growing awareness of the earth's environmental problems, and our limited stock of fossil fuels, hydrogen energy is expected to become one of the clean energy sources to replace fossil fuel. To be able to make use of hydrogen energy, hydrogen absorbing alloys are being given attention for energy conversion, hydrogen purification and so on. Such alloys are key materials for the development of future pollution-free clean fuels. The technological and scientific interest in intermetallic compounds as hydrogen storage systems has increased during the last few decades.

In the search of viable storage materials, considering the cost, weight, storage capacity and materials availability, we make recourse to Mg-bearing materials as the most attractive option for hydrogen storage. The maximum storage capacity of pure magnesium is 7.6 wt%. However, it reacts with hydrogen only at a very high temperature and under high pressure, and even then the reaction is slow and the degree of hydrogenation is incomplete. Moreover, the rate of release of hydrogen from magnesium hydride is unacceptably low. To eliminate the limitations of magnesium, and also to find the optimum alloy for hydrogen storage, several workers have used various metal additives to magnesium [1–4]. One of the major objectives is to reduce the dissociation temperature of magnesium hydride while retaining the high hydrogen storage capacity. We attempted to eliminate or reduce these disadvantages by alloying Mg with La [5]. It was observed that the intermetallic compounds La_2Mg_{17} and $LaMg_{12}$ are very attractive materials for hydrogen storage. At this point, we explain the notations for intermetallic compound and composite materials that will be used. For example, La_2Mg_{17} represents an intermetallic compound obtained by melting a mixture of La (Lanthanum) and Mg (magnesium) in the ratio (by weight) of $\frac{2x}{2x+17y} : \frac{17y}{2x+17y}$, where x and y are the atomic weights of La and Mg, respectively. A composite material 'P − x wt% Q' means that before melting the ratio of the intermetallic compound P and Q is $100 : x$. By 'storage capacity of an intermatallic compound P is x wt %', we mean that 100 gm of P can store x gm of hydrogen. The symbol 'Mm' is used to represent a compound which is the mixture of five metals, Lanthanum, Cerium, Praseodymium, Neodymium and Samarium.

The hydrogen storage capacity of the intermetallic compound La_2Mg_{17} is 4.2 wt% at 400 ° ± 10°C, and these materials absorb and desorb hydrogen reversibly at high temperature. But again, its disadvantage is its slow hydrogenation-dehydrogenation rate – about 10 times slower than that of the well known hydrogen storage material $LaNi_5$.

In recent years, one of the most interesting fields of metallurgy and materials science has been the study of composite materials [6–8]. The greatest promise of composite materials lies in the concept itself, which can free us from the limited set of single monolithic materials. The goal is to maximise the desirable properties, while minimise the undesirable ones. Generally, most of the efforts to establish suitable hydrogen storage systems have been concentrated on individual intermetallic compounds $LaNi_5$, $FeTi$, Mg_2Ni, $LaMg_{12}$, La_2Mg_{17}, etc. and their substituted versions. There are some investigations, which suggest that mixed-type magnesium-based composite materials may possess better hydriding characteristics [9–11].

We observed in our previous investigations that if $LaNi_5$ or $MmNi_{4.5}Al_{0.5}$, a material of lower stability (having comparatively lower hydrogen storage capacity) is combined with La_2Mg_{17}, a highly stable material (with higher storage capacity), in a proper ratio, then it is capable of giving a promising material with higher storage capacity and lower stability (easier absorption/desorption) [12,13]. We used composite materials $La_2Mg_{17} - x\ wt\%\ Z$, where Z could be either $LaNi_5$ or $MmNi_{4.5}\ Al_{0.5}$. As the sorption rates of storage alloys are important from the application point of view, we have already carried out kinetic evaluations for La_2Mg_{17} based composite materials at different temperatures and the results of the investigation have been reported elsewhere [14,15]. The absorption and desorption characteristics of the composites are dependent on the amount (x) of the material Z, absorption temperature (T^A), desorption temperature (T^D), absorption time (t^A) and desorption time (t^D).

Since it is difficult to find an explicit relation between different interacting factors, our intention in this paper is to explore the possibility of finding a *computational model* using a connectionist framework (neural network) for the relation between the amount of desorbed hydrogen (y) and desorption temperature (T^D), composition (x) etc. We illustrate the general idea with a simple example: suppose for

temperatures $T = T_1$, T_2 and T_3, we observed from experiments the amount of hydrogen that comes out for different desorption time (t^D) from a particular composite material, i.e. for a particular choice of x and Z. Now we use a multi-layer neural network to learn the relation between y and T^D, t^D. Once this is known, we can find the amount of desorbed hydrogen for a completely different temperature $T = T_4$ for different desorption times, without doing the actual experiment. Such a scheme will cut down on the time and cost of research and development drastically. To the knowledge of the authors, no one has investigated the hydriding-dehydriding kinetics of hydrogen storage materials using neural networks. Note that for such problems, not many data points will usually be available. However, since the relation to be learnt is governed by the physics of metals, we expect a smooth but non-linear relation between output and various input parameters. If the relation is a smooth one, it may be possible to extract a fairly good computational description of the input-output relation, even using limited data. To establish the fact, we generated several noisy samples, and demonstrated that even in this case, the relation can be identified.

The rest of the paper is organised as follows. For the sake of completeness, we provide a summary of the experimental procedure in Section 2. A brief discussion on multilayer perception networks is provided in Section 3. Section 4 explains how an MLP can be used in the present context. The results are presented in Section 5, while the paper is concluded in Section 6.

2. Experimental Procedure

The materials used in this investigation are highly pure lanthanum (99.9%), magnesium (99.99%), and intermetallic compounds $LaNi_5$ and $MmNi_{4.5}Al_{0.5}$. High purity hydrogen (IOLAR grade) was used in this study. The mother alloy La_2Mg_{17} was synthesised through a solid state diffusion process, by taking the pellet of stoichiometric quantities of lanthanum and magnesium [12]. The pellet was melted in a quartz tube under a protective Ar-atmosphere with the help of a radio frequency induction furnace. In the next step, the composite materials $La_2Mg_{17} - x$ wt% Z, where Z = $LaNi_5$ or $MmNi_{4.5}Al_{0.5}$ were synthesised for various values of x by melting pellets of homogeneous mixtures of the intermetallic compounds La_2Mg_{17} and Z in appropriate ratios. The synthesis procedure of the intermetallic compound La_2Mg_{17} and composite materials La_2Mg_{17} − x wt% $LaNi_5$ and La_2Mg_{17} − x wt% $MmNi_{4.5}Al_{0.5}$ has been demonstrated in detail by Pal [12,13].

Generally, at room temperature, the hydriding reaction of La_2Mg_{17}-based composites is very difficult to nucleate. Some activation procedure is necessary for this. The activation as well as the hydriding characteristics of the composite materials were performed using a reactor vessel of known volume. The amount of materials used was about 3 gms. After evacuating the vessel, the sample was brought in contact with hydrogen under a pressure of nearly 33 kg/cm^2. The material was heated at nearly 360°C for several hours (depending on the sample), and then cooled to room temperature. Then at 360° ± 10°C, hydrogen was desorbed and the vessel containing the material was evacuated. Again, hydrogen gas was introduced into the reactor vessel. We continuously heated and degassed the composite material under study for several cycles at high temperature, and applied high hydrogen pressure for full activation. When the sample is in contact with hydrogen, it usually takes some time before any noticeable change in hydrogen pressure can be observed. Depending on the stability of the composite material, the composition of the samples, the activation temperature and the pressure, the activation period might be anything between a few seconds and a few days.

It was found that the initial reaction of hydrogen with the fresh sample is a very slow process, and the hydriding behaviour improves with subsequent absorption-desorption cycles. After a number of hydrogenation-dehydrogenation runs, there was no appreciable change with cycling in the dehydriding kinetics and in the amount of desorbed hydrogen by the composite material. This indicates that the activation treatment is over. The kinetic experiments were carried out after full activation of the material for various values of x. The quantities of desorbed hydrogen were measured by a volumetric method after preliminary outgassing of the reactor vessel. The dehydriding kinetics of the La_2Mg_{17}-based composite materials were studied between 350°C and 400°C under a pressure of 4 kg/cm^2. The temperature was kept constant within ±10°C. To calculate the desorption rate, data were collected after removing hydrogen from the synthesised materials. The procedure consisted of (i) putting the material under study into the reactor vessel, (ii) sealing the reactor so that there is no leakage, (iii) evacuating the reactor vessel and then pressurising the system under 33 kg/cm^2 with hydrogen, (iv) activating the sample for several cycles before measuring the dehydriding rate to ensure that the sample is fully hydrided, (v) heating the material at high temperature under high

hydrogen pressure, and (vi) measuring the amount of hydrogen coming out of the material with time at a particular pressure.

Next, with an introduction to neural networks, we describe the model required for our study.

3. Neural Networks

Although the concept of artificial Neural Networks (NN) has been inspired by biological neural networks, the heart of this emerging technology is rooted in different disciplines. Biological neurons are the structural constituents of the brain, and they are much slower than silicon logic gates. But inferencing in biological NN is faster than the fastest computer available today. The brain compensates for its relatively slower operation by a large number of neurons with massive interconnections between them. Biological neural networks

(i) are nonlinear devices, which are highly parallel, robust and fault tolerant;
(ii) have built-in capability to adapt their synaptic weights to changes in the surrounding environment;
(iii) can handle imprecise, fuzzy, noisy and probabilistic information; and
(iv) can generalise from known tasks or examples to unknown ones.

Research on Artificial Neural Networks (ANN) started with a view to mimicking some or all of these characteristics [16–18]. But at present, the focus is not just to mimic the brain, but to use principles from nervous systems to solve complex problems in an efficient manner. This new computational paradigm is different from programmed instruction sequences. Here information is stored in the synaptic connections. A neuron is an elementary processor with primitive types of operations, like summing the weighted inputs coming to it, and then amplifying or thresholding the sum. The computational neuron model of Mc-Culloch–Pitts is a simple binary threshold unit where a neuron has a number of input links and one output link. The ith neuron computes the weighted sum of all its inputs from other units and outputs a binary value, zero or one, depending on whether this weighted sum is greater than equal or less than a threshold θ_i. Thus:

$$x_i(t+1) = f\ (\Sigma_j\ w_{ij}\ x_j(t) - \theta_i)$$

$$\text{where } f(x) = 1 \text{ if } x \geq 0$$

$$= 0 \text{ otherwise}$$

Here w_{ij} is the connection weight between neuron i and neuron j, and x_j is the input to the jth neuron. Often the hard threshold function f of Mc-Culloch–Pitts neuron is replaced by a sigmoidal type nonlinear function [16,17].

A neural network is characterised by the network topology, connection strength between pairs of neurons (weights), node characteristics and updating rules. The updating or learning rules may be for weights and/or states of the processing elements (neurons). Normally, an objective function is used to represent the status of the network, and its set of minima corresponds to different stable states of the network. The adaptability of a neural network comes from its capability of learning from 'environments' (as represented by the training data). There are different models of NNs [17], but our interest is in the multilayer perceptron network (MLP).

An MLP (Fig. 1) is a layered network which can be trained to learn an unknown relationship between a set of inputs and outputs. The first layer is the input layer and the last layer is the output layer. All remaining layers are hidden layers and nodes in them are called hidden nodes. Each node in a hidden layer is connected to all nodes in its immediately preceding and following layers via connection weights. At each neuron (nodes except those in the input layer), all incoming signals are summed up and transformed by an activation function to give the output of the node.

For an MLP the problem is to select an appropriate set of weights to minimise the error between the network outputs and the target outputs. Among the variety of training schemes [19,20] that have been developed, the most celebrated one is the back-propagation algorithm, which is a gradient descent in the weight space over a surface generated by the sum of the squared errors between the target outputs and the network outputs.

Let $S = (X,Y)$ be the training data, where $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$; $\mathbf{x}_i \in \Re^p$ and $Y = \{\mathbf{y}_1, \ldots, \mathbf{y}_N\}$; $\mathbf{y}_i \in \Re^q$; $\mathbf{y}_i$ is the q-dimensional output corresponding to the p-dimensional input $\mathbf{x}_i$.

We use the following notation:

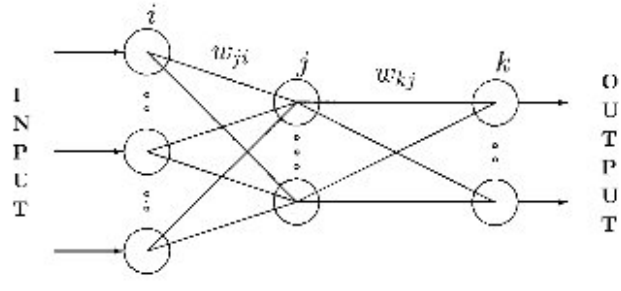

Fig. 1. Multilayer perceptron network.

x_i : ith component of an input vector $\mathbf{x}$ in the training set.

y_i : ith component of the target vector $\mathbf{y}$ corresponding to $\mathbf{x}$.

f : Activation function.

n : Number of hidden layers in the network.

n_i : Number of nodes in the ith layer.

o_i^k : Output of the ith node in the kth hidden layer and $o_i^0 = x_i$ and $o_i^{n+1} = y_i$.

f'^k_i : Derivative of the activation function of the ith node in the kth layer.

w_{ij}^k : Weight connecting the jth node of the kth hidden layer to the ith node of the $k+1$st layer.

$\epsilon_{\mathbf{x}}$: Error corresponding to the pair $(\mathbf{x},\mathbf{y})$.

η : learning rate or step size.

The back-propagation algorithm consists of a *forward* pass and a *backward* pass, the latter being the weight adjustment pass. The forward pass computes the output at each node by:

$$o_i^k = f(\Sigma_j\,(o_j^{k-1} w_{ij}^k)) \quad \text{when, } k = 1, \ldots, n+1; \quad (1)$$

$$o_i^k = x_i \text{ when, } k = 0.$$

In the backward pass, with a view to minimising $\Sigma\epsilon_{\mathbf{x}}$ over the entire training data-set S weights are adapted using gradient descent on each $\epsilon_{\mathbf{x}}$ in the training set, where

$$\epsilon_{\mathbf{x}} = 1/2 \sum_i^q (y_i - o_i^{n+1})^2 \quad (2)$$

Using the gradient descent technique one can obtain the following learning rule [19]:

$$\Delta w_{ij}^k = \eta \delta_i^{k+1}\, o_j^k \text{ where}$$

$$\delta_i^k = f'^k_i \sum_{l=1}^{n_{k+1}} (\delta_l^{k+1}\, w_{li}^k) \text{ when } k = 1, \ldots, n; \quad (3)$$

$$\delta_i^k\,(y_i - o_i^{n+1})\, f'^k_i \text{ when } k = n+1$$

The incremental changes Δw_{ij}^k may be summed up over the entire data-set S and then w_{ij}^ks may be updated by the resulting sums, or the updating may be done separately for each pair $(\mathbf{x}_i, \mathbf{y}_i)$ in S. The former method is called the *batch* mode, while the latter is called the *online* or pattern mode of training [17].

4. Use of MLP in Materials Science

Suppose we want to characterise composite materials $La_2Mg_{17} - x$ wt% Z. Our interest is to find out how the absorption rate (y_1), and desorption rate (y_2) are related to wt% of Z (i.e. x, absorption temperature T^A, desorption temperature T^D, time of absorption t^A and time of desorption t^D). We assume that there is a non-linear relation between (x, T^A, T^D, t^A, t^D) and (y_1, y_2). Let us denote the set (x, T^A, T^D, t^A, t^D) by a vector $\mathbf{x}$ in $\Re^5$ and (y_1, y_2) by a vector $\mathbf{y}$ in $\Re^2$. Here our assumption is that there exists a smooth, but non-linear relation F between $\mathbf{x}$ and $\mathbf{y}$, such that $\mathbf{y} = F(\mathbf{x})$. In the NN terminology, $\mathbf{x}$ will be called the input vector while $\mathbf{y}$ the output vector. The objective is to find a computational scheme (NN) to identify F. To do this we need some data. So we find $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ for several values of $\mathbf{x}$. Let there be N such experimental results: $X = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\} \subseteq \Re^5$ and $Y = \{y_1, y_2, \ldots, \mathbf{y}_N\} \subseteq \Re^2$, where $\mathbf{y}_i$ is the output for $\mathbf{x}_i$. In general, let us assume that $x_i \in \Re^p$ and $\mathbf{y}_i \in \Re^q$. Given (X, Y), we can train an MLP with p input node, q output node and $n \geq 1$ hiddenlayers, where the ith hidden layer has n_i nodes, $i = 1, 2, \ldots, n$. Let the computed output of the net for input $\mathbf{x}_i$ be $\mathbf{y}_2'$. The connection weights of the net are then adapted with a view to minimising $\Sigma_{i=1}^N \|\mathbf{y}_i - \mathbf{y}_i'\|^2$ using the back propagation algorithm described in the previous section.

In the present investigation, we would like to learn the relation between the amount of desorbed hydrogen from La_2Mg_{17} − x wt% Z and desorption time and temperature for a particular x and Z. In other words, for a fixed x and Z, we perform experiments for several t^D and T^D, use these results to train an MLP, and finally, use the trained MLP to predict the desorbed hydrogen for different t^D and T^D, not used in the experiment.

In this case, each $\mathbf{x}_i$ has two components $x_{i1} = t_i^D$ and $x_{i2} = T_i^D$ while the output vector has only one component $\mathbf{y}_i = y_i =$ amount of desorbed hydrogen. Thus, a net with two input nodes, one output node and $n \geq 1$ hidden nodes will be used. In the subsequent discussion, we use the notation (p-h-q) to represent a net with p input nodes, one hidden layer with h hidden nodes and an output layer with q nodes. Thus, (2–5–7–1) will represent a net with two input nodes, one output node and two hidden layers with five and seven nodes, respectively.

5. Results and Discussion

Figures 2 and 3 show the kinetic curves of hydrogen desorption for the composite materials La_2Mg_{17} − x wt% $MmNi_{4.5}Al_{0.5}$ at 350° ± 10°C, 375° ± 10°C

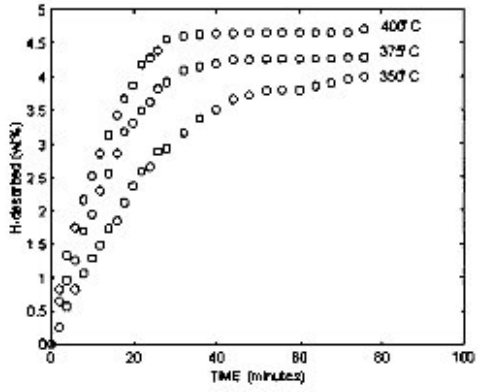

Fig. 2. Kinetic curves of hydrogen desorption at different temperatures and at a pressure of 4 kg/cm^2 for La_2Mg_{17} − 10 wt% $MmNi_{4.5}Al_{0.5}$.

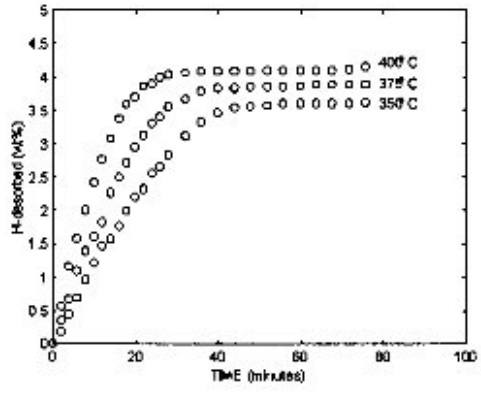

Fig. 3. Kinetic curves of hydrogen desorption at different temperatures and at a pressure of 4 kg/cm^2 for La_2Mg_{17} − 20 wt% $MmNi_{4.5}Al_{0.5}$.

and 400° ± 10°C temperature, and at a pressure of 4 kg/cm^2 for x = 10 and 20, respectively. From hereon, we will omit writing the temperature range ± 10°. The rates of hydrogen desorption for La_2Mg_{17} − 5 wt% $LaNi_5$ are shown in Fig. 4. They

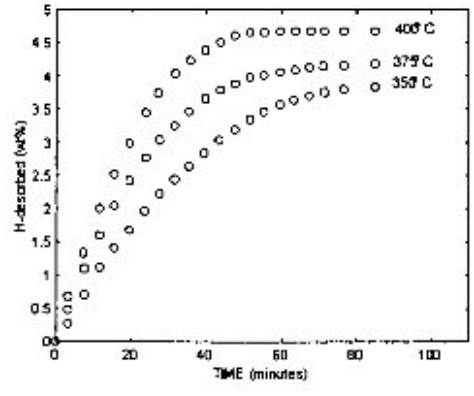

Fig. 4. Kinetic curves of hydrogen desorption at different temperatures and at a pressure of 4 kg/cm^2 for La_2Mg_{17} − 5 wt% $LaNi_5$.

are measured at three different temperatures 350°C, 375°C and 400°C under a hydrogen pressure of 4 kg/cm^2. In the initial stage the rate increases rapidly with time at an almost uniform rate. Subsequently the rate decreases rapidly. In all cases, desorption increases rapidly in the beginning and then it slows down. The detailed analysis of these results has been reported elsewhere [14,15].

Each series of experiments (for a fixed x, Z, activation procedure, the set of observations for a particular T^A and T^D is called a series) takes at least 10 days (of course, if the experimental procedure is optimised, otherwise it may even take more than a month). Let us pretend that we have only two series of experimental data for T^D = 350°C and T^D = 400°C. Given such two sets of observations, is it possible to learn the characteristics of the composite material so that we can predict the desorption kinetics at T^D = 375°C without doing the experiment? To explore this possibility, we first consider Z = $MmNi_{4.5}Al_{0.5}$ and x = 20 and a network with (2–4–6–1) architecture. We trained the net for 25,000 iterations with 26 observations for each T^D (T^D = 350°C and 400°C). Figure 5(a) shows an excellent match between the actual output and the network output for the training data at two different temperatures 350°C and 400°C. Although the net is able to learn the training data quite well, it does not necessarily mean that it has acquired a good generalisation capability. To check this, we pretend as if we do not have observations for T^D = 375°C and predict the description kinetics at T^D = 375°C for different t^Ds. Figure 5(b) shows the network predicted output and the actual output for T^D = 375°C, and we indeed find a very good match. The network predicted output is shown by ooooo, while **** indicates the output of the actual experiment. In this case, we have training data for T^D = 350°C and T^D = 400°C, and the net could predict (interpolate) values for the desorption kinetics at T^D = 375°C.

Usually, MLP is not good for extrapolation unless the input-output relation is nice and the test input are not far from the training data. Keeping this in mind, let us see whether the net could do extrapolation. Since the input-output relation to be learnt is governed by the physics of materials, the underlying relation is expected to be smooth and extrapolation might work. And, if it does, it will be of great help to the material scientists. With this hope, we train the net using the data for T^D = 350°C and T^D = 375°C and predict the dehydrogenation rate for T^D = 400°C. Figure 6(a) shows the match with the training data for an architecture (2–6–5–1), while Fig. 6(b) displays that the network predicted dehyd-

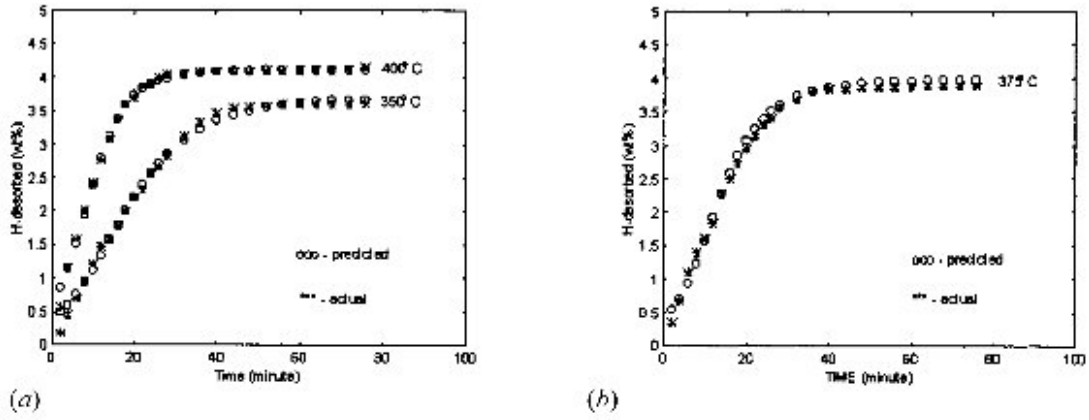

Fig. 5. Predicting the dehydriding kinetics of the composite material La_2Mg_{17} − 20 wt% $MmNi_{4.5}Al_{0.5}$. (a) Agreement between training data and network suggested output with T^D = 350°C and T^D = 400°C with an architecture (2–4–6–1); (b) agreement between actual desorption rate at T^D = 375°C and the network interpolated output when the trained net corresponding to Fig. 5(a) is used.

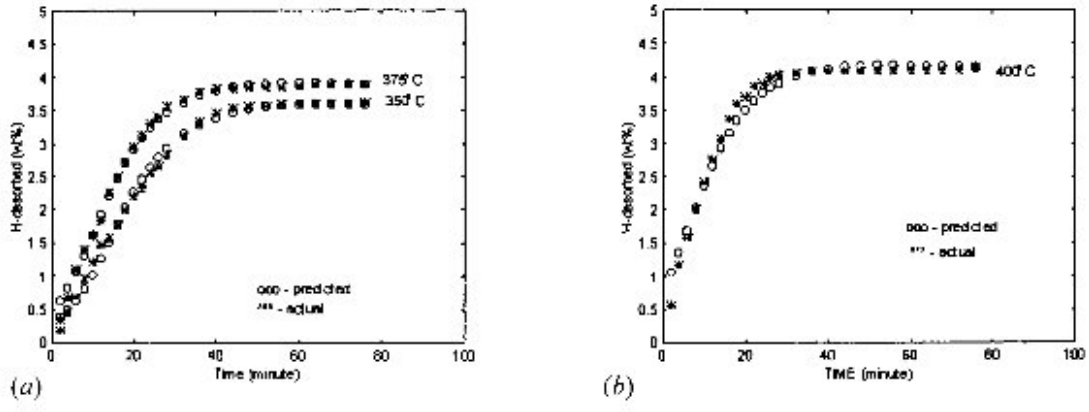

Fig. 6. Prediction of the dehydriding kinetics for the composite material La_2Mg_{17} − 20 wt% $MmNi_{4.5}Al_{0.5}$. (a) Agreement between training data and network suggested output with T^D = 350°C and T^D = 375°C with an architecture (2–6–5–1); (b) agreement between the actual desorption rate at T^D = 400°C and the network extrapolated dehydriding kinetics when the trained net corresponding to Fig. 6(a) is used.

riding kinetics for T^D = 400°C agrees quite nicely with the actual desorption amount at 400°C.

The preceding two computational exercises were done with MLPs with two hidden layers, but our experience shows that in most cases, a MLP with a single hidden layer is sufficient. As an illustration, we report the results for the case when a net of (2–9–1) size was trained with data for T^D = 350°C and T^D = 400°C. Figure 7(a) shows the match between the predicted output and the actual output for the training data, while Fig. 7(b) shows the agreement between the network suggested desorption kinetics and the actual. Figure 7 clearly reveals that even a single layer net is enough to capture the desorption characteristics of La_2Mg_{17} − 20 wt% $MmNi_{4.5}Al_{0.5}$.

Experiments were also conducted using the same material for x = 10% with both one hidden layer and two hidden layers, and for both interpolation and extrapolation cases. In all cases, we got quite satisfactory results. We just report here only one result with an architecture (2–7–1) for the interpolation case (Figs 8(a) and 8(b)).

Next we consider a different composite material La_2Mg_{17} − 5 wt% $LaNi_5$ i.e. $Z = LaNi_5$ and x = 5. Figure 9(a) shows the performance of a (2–7–7–1) architecture on the training data when training was done using data for T^D = 350°C and T^D = 400°C for 25,000 iterations. Figure 9(b) depicts an excellent match between the actual amount of desorbed hydrogen and the predicted amount of desorbed hydrogen at T^D = 375°C. Similarly, we trained a net with a single hidden layer having nine nodes for T^D = 350°C and T^D = 400°C and predicted the amount desorbed hydrogen as a function of time. The results obtained after 25,000 iterations with the (2–9–1) net are shown in Figs 10 (a) and 10(b). Figure 10 exhibits that even a single layered

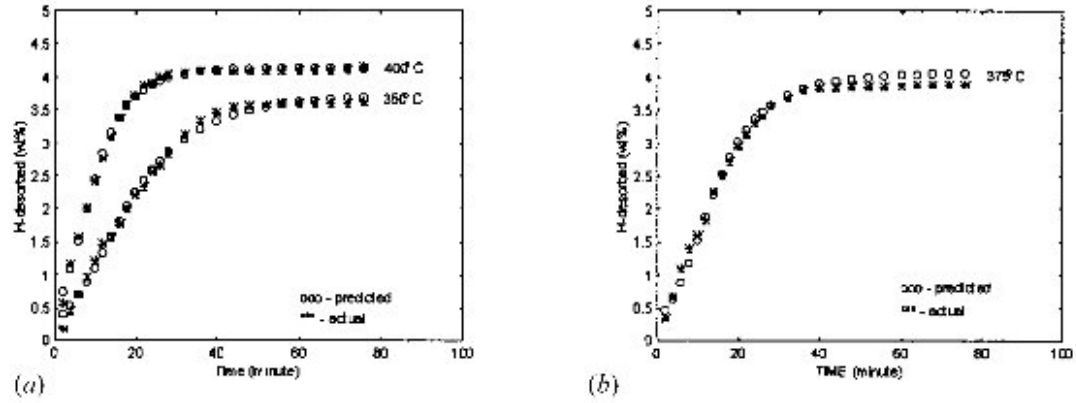

Fig. 7. Predicting the dehydriding rates of the composite material La_2Mg_{17} − 20 wt% $MmNi_{4.5}Al_{0.5}$. (a) Agreement between training data and network suggested output with T^D = 350°C and T^D = 400°C with an architecture (2–9–1); (b) match between the network interpolated desorption kinetics and the actual at T^D = 375°C when the trained net corresponding to Fig. 7(a) is used.

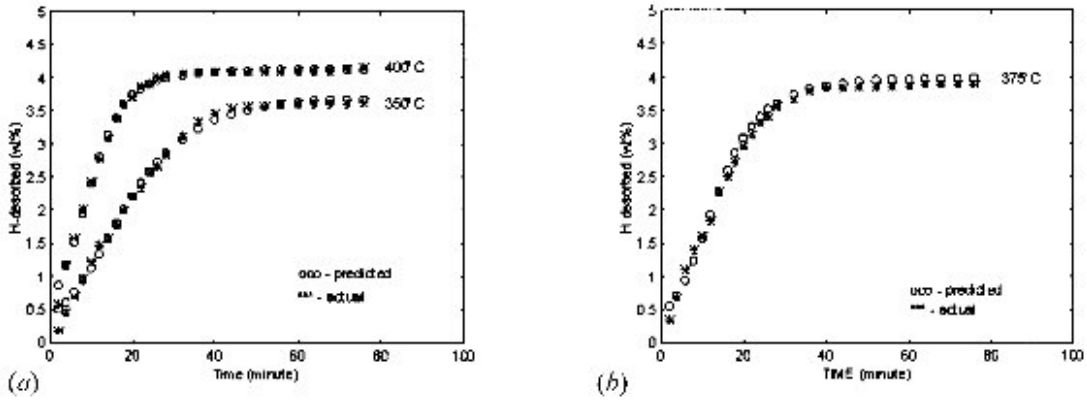

Fig. 8. Predicting the dehydriding kinetics of the composite material La_2Mg_{17} − 10 wt% $MmNi_{4.5}Al_{0.5}$. (a) Agreement between training data and network suggested output with T^D = 350°C and T^D = 400°C with an architecture (2–7–1); (b) agreement between actual desorption rate at T^D = 375°C and the network interpolated output when the trained net corresponding to Fig. 8(a) is used.

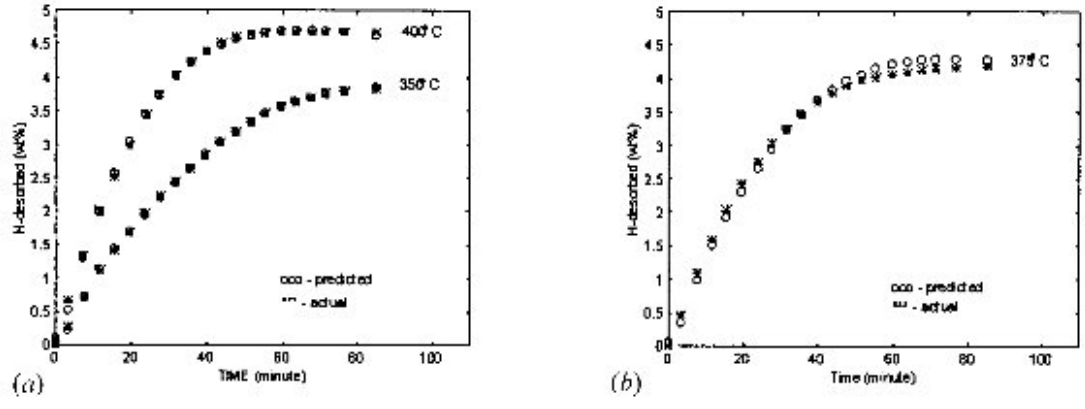

Fig. 9. Predicting the dehydriding kinetics of the composite material La_2Mg_{17} − 5 wt% $LaNi_5$. (a) Performance of a (2–7–7–1) architecture on the training data with T^D = 350°C and T^D = 400°C with 24 observations for each T^D; (b) agreement between actual desorption rate at T^D = 375°C and the network interpolated output when the trained net corresponding to Fig. 9(a) is used.

net can identify a computational model for the desorption kinetics.

Let us now explore how good the connectionist model is for extrapolation with La_2Mg_{17} − 5 wt% $LaNi_5$. Here the training data consist of 21 observations at each of T^D = 350°C and T^D = 375°C. Figure 11(a) displays the performance of the net for the training data. Again, we see a good agreement

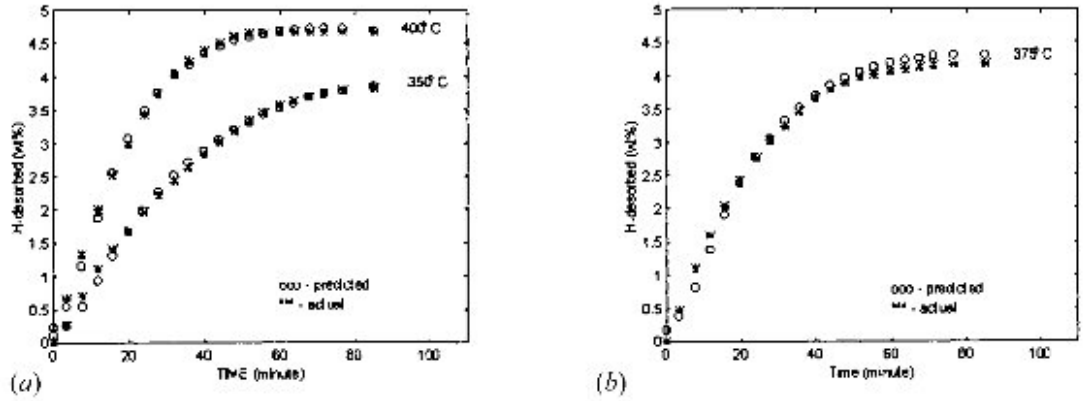

Fig. 10. Predicting the dehydriding kinetics of the composite material La_2Mg_{17} − 5 wt% $LaNi_5$. (a) Agreement between training data and network suggested output with T^D = 350°C and T^D = 400°C with an architecture (2–9–1); (b) agreement between the actual desorption rate at T^D = 375°C and the network interpolated dehydriding kinetics when the trained net corresponding to Fig. 10(a) is used.

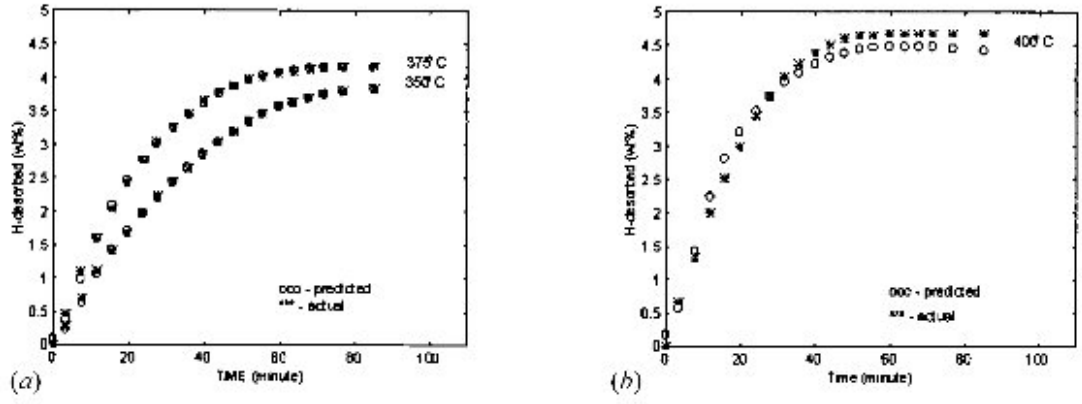

Fig. 11. Predicting the dehydriding kinetics of the composite material La_2Mg_{17} − 5 wt% $LaNi_5$. (a) Agreement between training data and network suggested output with T^D = 350°C and T^D = 375°C with an architecture (2–9–1); (b) agreement between the actual desorption rate at T^D = 400°C and the network extrapolated dehydriding kinetics when the trained net corresponding to Fig. 11(a) is used.

between the actual output and the computed output by the net. Figure 11(b) depicts the network output and the actual experimental output for T^D = 400°C. In this case, the computed desorption amount for $t^D > 40$ is little less than the actual output, while for $t^D < 40$, the computed output is little more than the actual, but still the agreement is quite satisfactory.

For such applications, it is not possible to get many observations and so was our case – we did not have too many observations. However, since the underlying relation that we are trying to learn is governed by the laws of physical sciences, we could identify the correct relation and that is established through both interpolation and extrapolation. To demonstrate robustness and reliability of such systems, we generated some noisy data. We assumed that instruments, which measure time and desorbed hydrogen, have less that 5% error on the measured value. For a particular reading the error is random. There are two types of randomness: the error (noise) can increase or decrease the observed value with probability 0.5 and the amount of error is also random but less than equal to 5%. It is a two step process. In step 1 we draw a random number in [0,1], and if it is > 0.5 then the noise will be added or else it will be subtracted. In step 2, we draw another random number, say r, in [0, 0.05]. If the observed value is x, then the noise is e = x.r. Now e is either added to or subtracted from x depending on result of step 1. This is done for all observations for both x = time and x = amount of hydrogen desorbed. This way we generate an augmented data set and the network is trained with this. Figure 12 shows the noisy data set corresponding to Fig. 4 (considering only the curves for 350°C and 400°C), while Fig. 13(a) compares the actual desorbed amount with the output obtained from the net trained using the noisy data. We find a very nice agreement between the net output and the original data, and it is quite comparable to Fig. 9(a). For this experiment we used the same architecture, initialisation, etc. as used for the result shown in

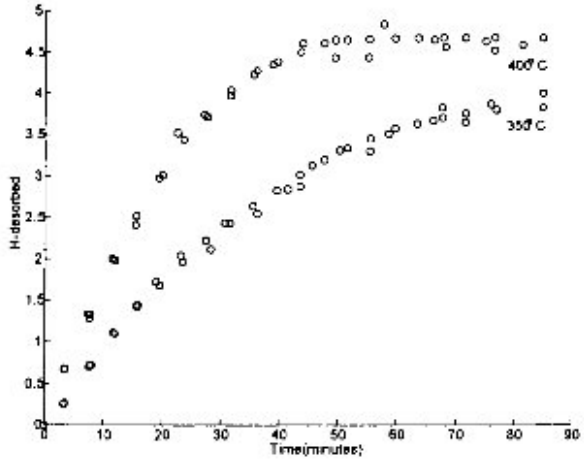

Fig. 12. The noisy data generated from data corresponding to Fig. 4, i.e. for the composite material $La_2Mg_{17} - 5$ wt% $LaNi_5$.

Fig. 9(a). To test the generalisation ability of the trained net, we predicted the desorbed amount of hydrogen with time for 375°C and Fig. 13(b) shows that there is a very good agreement between the predicted value and the actual. Comparing Fig. 13(b) with Fig. 9(b), we see that the noisy data did not influence much the generalisation ability of the network. We repeated such experiments for several other noisy data sets generated using the same philosophy, and results obtained were pretty consistent with the one reported.

These observations show that the connectionist framework is very effective to identify quite accurately and consistently the relation characterising the amount of desorbed hydrogen as a function of T^D and t^D, even though we have a limited number of data points. As a result, we could predict the amount of desorbed hydrogen for given T^D and t^D without doing the actual experiments, and this saves both time and cost of research. This is possible, mainly because, the relation is smooth and it is governed by laws of physical science.

6. Conclusions

We have explained how connectionist models, in particular, multilayer perception networks can be used in the materials science research and development. *Computational materials science* appears to have high potential to make research and development in the areas of materials science more productive and useful at low cost and time. As examples, we have shown that for $La_2Mg_{17} - x$ wt% $LaNi_5$ and $La_2Mg_{17} - x$ wt% $MmNi_{4.5}Al_{0.5}$ the desorption kinetics for a fixed x can be identified (learnt) easily by an MLP with limited experimental data. Once the computational model of the kinetics is obtained (i.e. after the neural net is trained), it can be used to find the desorbed amount of hydrogen at new desorption temperature T^D and desorption time t^D without actually doing the experiment. The success of the scheme has been demonstrated on two different composite materials $La_2Mg_{17} - x$ wt% $LaNi_5$ and $La_2Mg_{17} - x$ wt% $MmNi_{4.5}Al_{0.5}$ for different values of x. It would be interesting to investigate the utility of this new approach to identify the relation of the amount of desorbed hydrogen to x, T^D and t^D. A further more complex study would be to train the net with five dimensional input $\mathbf{x} = (x,$

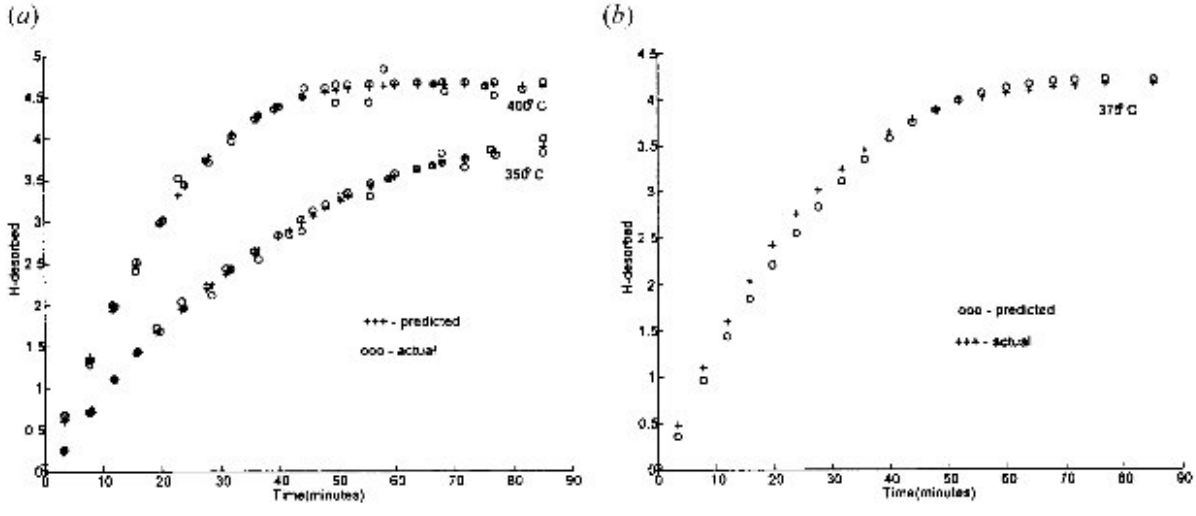

Fig. 13. Predicting the dehydriding kinetics of the composite material $La_2Mg_{17} - 5$ wt % $LaNi_5$ on the noisy data in Fig. 12. (a) Performance of a (2–7–7–1) architecture on the noisy training data with $T^D = 350$°C and $T^D = 400$°C; (b) agreement between actual desorbed ammout at $T^D = 375$°C and the network interpolated output using the trained net corresponding to Fig. 13 (a) (the net was trained using the noisy data in Fig. 12).

T^A, t^A, T^D, $t^D)^T$ and two dimensional output $\mathbf{y} = (y_1, y_2)^T$, where y_1 is the amount of absorbed hydrogen and y_2 is the amount of desorbed hydrogen. We would like to explore these possibilities in future. Note that the proposed methodology is not only applicable to hydrogen storage materials, but also in many other materials science problems, like crystal growth, in medicine to realise desirable features in mixed chemicals, in rubber industries to produce tyre-material with specified characteristics, and so on.

Acknowledgements. We gratefully acknowledge the partial support by Department of Non-conventional Energy Sources, India and the Council of Scientific and Industrial Research, India for completion of this work. Thanks are also due to Professor O.N. Srivastava, Department of Physics, Banaras Hindu University, Varanasi, India for his interest in this work.

References

1. Selvam P, Viswanathan B, Swamy CS, Srinivasan V. Int. J. Hydrogen Energy 1986; 11: 169
2. Anderson AF, Maeland AJ. Hydrides for energy storage. Pergamon Press, Oxford 1987
3. Suda S. Int. J. Hydrogen Energy 1987; 12: 323
4. Bogdanovic B. Int J. Hydrogen Energy 1987; 12: 863
5. Dutta K. Thesis: Investigations on synthesis, structural - microstructural characterizations and hydrogenation behavior of some hydrogen storage materials. Banaras Hindu University, India, 1993
6. Singer C, Holmyard EJ, Hall AR. In A History of Technology, Vols I (1954) and II (1956). Oxford
7. Forbes RJ. Studies in Ancient Technology, Vols. I–III. Leiden 1955
8. Hodges H. Technology in the Ancient World. Pelican Books 1971; 153: 35, 61, 67, 69, 83, 97, 110
9. Tanguy, B, Soubeyroux JL, Pezat M, Portier J, Hagenmuller P. Mater. Res. Bull. 1976; 11: 14441
10. Nagai H, Tomizawa H, Ogasawara T, Shoji K. J. Less-Common Mets. 1990; 157: 15
11. Fujii H, Orimo S, Yamamoto K. J. Less-Common Mets. 1991; 157: 243
12. Pal K (nee Dutta). J Materials Science 1997; 32: 5177
13. Pal K (nee Dutta). Int. J. Hydrogen Energy 1997; 22: 903
14. Dutta K, Srivastava ON. J. Materials Science 1993; 28: 3457
15. Dutta K, Srivastava ON. Int. J. Hydrogen Energy 1993; 18: 397
16. Hertz J, Krogh A, Palmer RG. Introduction to the Theory of Neural Computation. Addison-Wesley, Redwood City, CA, 1991
17. Haykin S. Neural networks - a comprehensive foundation. Macmillan College, Proc. Con. Inc, NY, 1994
18. Rosenfeld R, Anderson J (Ed.). Neuro Computing MIT Press, 1988
19. Pao YH. Adaptive Pattern Recognition and Neural Networks. Addison Wesly Publishing Company, Massachusetts, 1989
20. Fahlaman S. Fast-learning variations on back-propagation: An imperical study. Connecionist Models Summer School. Proc 1998