LEARNING PHENOMENA IN MANUFACTURING
AND ARTIFICIAL NEURAL NETWORK

Miroslav Car
University of Zagreb, FMENA/FSB, Zagreb, Croatia
miroslav.car@fsb.hr

Abstract: In the industrially advanced countries, that are different from our ex and present countries, to learning phenomena has been dedicated a significant attention for the last 60 years. One of more basic reasons is multiple purposes of results. Until now, there have been applied various approaches, methods and procedures for empirical data approximation, and in this article some possibilities of artificial neural network application are researched.

Keywords: manufacturing process, operation operator, learning process, artificial neural network, dynamic neural network.

1. INTRODUCTION

In the world, especially in industrial advanced countries, there has been dedicated significant attention to the research and application of learning phenomena in manufacturing: There were two reasons, more important than the rest: (1) multiple purposes of results, with direction of application results in a low lot size and work manufacturing and in equipment investment manufacturing, (2) establishing of time as a correct and objective measure of technological processes projecting. Relationships “quality-cost-term deliverance” consist of three affecting factors, two of which depend on the registering of work time and incentive (allowed) time, as well as about processes and time learning.

In preceding researches, there have been applied various approaches, methods, techniques and procedures for the time data establishing by a stopwatch and some other instruments registering, from classical models (Wright, De Jong, etc.) cross nonlinear models of polynomial regression to the models of time series. For analysis and partial forecastings (foreseeing, prediction) of the legality of the manufacturing phenomena, there were applied Hancock and BBC models for time data projection by the most important of PDTS, system MTM-1.

In the world, there have been developed two basic approaches of artificial neural network modelling: (1) neurocomputers with parallel processing design, and (2) simulating of the human like models, the way to artificial intelligency, that is the selected approach of this article.

Depending on selected criterias, there are various categories, types and models of artificial neural networks, learning modes with adequate procedure, as well as various algorithms. In this article, artificial neural network has been applied for analysis of learning
process data, and its comparison with preceding results.

2. THE DEVELOPMENT OF LEARNING PHENOMENA IN MANUFACTURING

Literature from learning phenomena in manufacturing coverage is especially extensive and rich in industrially advanced countries: USA, G. Britain, Germany, France, Sweden, etc. If, by example, person who is insufficiently acquainted with researching about phenomena has a possibility to list literature reviews of learning phenomena, as well as exhibitions of famous and acknowledgment authors of single modes and methods, it can easily list various names of curves or functions.

Nonacquainted person can easily come to the wrong conclusion about more various phenomena, but a well acquainted person, an expert, has to point on existing only one entire learning phenomena, divided by the criteria of (non) existing break or interruption in learning on two groups: (1) learning without breaks (continuous learning) and to (2) learning with breaks (discrete-discontinuous learning).


D. Marić is one of a rare authors from ex Yugoslavian areas who, beside articles [M. Car, 1990.; M. Car, D. Filipusic, 1991.; M. Car, 1993.] to [M. Car, 1995.] and belonging authors, have researched learning phenomena. He exhibited in 1979. In [D. Maric, 1979.,] the reasons of curve startup, which are seen in a fact that in the beginning of any process (production or manufacturing, technological, operation, etc.) in the belonging system and with purpose of doing products all participants-operators, and every operator especially, are uneducated or insufficient educated (without or with insufficient motivation, acknowledgment and skills) and the cost much more time than is the cost rate of repetitive sameness/like cycles. Opposite to preceding, establishing of a lot of partial non advanced approaches, with own researches, proclaimed in (M. Car, 1990.; M. Car, D. Filipusic, 1991.; M. Car, 1993.; M. Car, 1995), there are established basis of more rational design for different modes of classification and order researching phenomena by means of selected seven criterias. Those suggest a new shape and content of entire classification and order.

Consequences are that most frequently, the design lacks in following: too low estimation of requested investment resources at investment to capacity phase, optimistically contracted term delivery, later hardly realizable, unreal normatives and incentive (allowed) time, caused problems in production-manufacturing planning and realization. To have better and objective control upon lacks, it's needed to get acquainted with the learning process legality.

W. M. Hancock cited in (W.M.Hancock, 1971.), that operator often works with a low lot size in which the prediction of a learning rate usually has to be changed. Researched data shows that very short breaks with maximum time rate of one work day, for example, drinking a coffee, so different waitings and breaks appear with various reasons, but these haven't significantly affected the learning with a break. Longer breaks, from one to ten days and more, represent significant affect on forgetfulness, and that's why learning without breaks belongs to learning with breaks. For learning without breaks, a few authors exhibited in their own article in order and description of application.
Time (cost) and reasons of breaks startup is divided in 2 types: (1) external, where it consists of hiring of persons, as well as releasing of existing employees and employing a new employees, and (2) internal, as well as turnover of tasks-works in shift, and changeover tasks from shift to shift. Order and description of applications for learning with breaks have a few significant differences towards learning without breaks and there are caused by affecting of forgetfulness and repetitiveness on increasing of time and cost learning. The five most important applications are cited in [Ch. D. Bailey, 1989.]

Some authors gave more significant contribution in general approach analysis for purpose of break (time) determining, but some of them are in special application analysis. For example, 1974. in [G. L. Adler, R. Nanda, 1974.; two articles in the same journal] Adler and Nanda exhibited affecting of learning, breaks and forgetfulness on economical lot size determination, as well as it did Agrawal 1978. in [G. K. Agrawal, 1978.].

About definition of phenomena, famous researchers in preface of (R. Nanda, G. L. Adler, 1977.) cited the following: “Learning Curve is graphically or analytically representing of expecting decrease in entrance resources at manufacturing process repetition. In this matter, price or time decreasing, as well as increasing in manufacturing size, it is performed in one part by improvement of operator's performing direct work and in the other, by improvement of personal work managing and organizing.”.

Three basic ages of entire learning phenomena are: (1) age of introducing, (2) age of learning (framing), and (3) age of acquired (trained). Schendel and co-operators established 1978. (J. D. Schendel, J. L. Shields, M. S. Katz, 1978.), in a studied experiment, that significantly different “discrete” tasks and “continuously controlled” tasks exist. “Discrete” tasks consist of series of discrete motor reaction, as application of pressure of fitted buttons and montage of fitted discs and tracks for computer work. “Continuously controlled” tasks involve repeated movement without a clear start or an end, as driving a bicycle or visual control of parts on line of assemble. So, “discrete” task does not consist only of learning a motion, but it requests multiple makings of a decision about what has to be done. The application of the experiment involves a “discrete” task of mechanical assemble and disengage, because it represents more than a nature of “continuously controlled” task.

3. ARTIFICIAL NEURAL NETWORK

3.1. THE BASIC FEATURES OF DEVELOPMENT AND ACTION OF ARTIFICIAL NEURAL NETWORK, (B. NOVAKOVIC, D. MAJETIC, M. SIROKI, 1998.)

Long-term trials of man's biophysiology and man's brain work modeling with capability of information processing influenced the conception phenomena of artificial neural network. W. James set in (W. James, 1890.), the basis for basic structure of artificial neuron by means of acknowledgment of the existing activity at any point in brain with the purpose assumed in sum tendency to discharging residual points in selected point, depending on: number of neurons (relationships), (2) intensity (weight, complexity) of those relationships, (3) absence of a point towards which direction of a sum tendency to discharging residual points could be changed. On preceding basis McCulloh and Pitts made a simple model of artificial neuron (McCulloh and Pitts, 1943.), even today basic block for building of artificial neural network, which have a distinction one to another by relationships of structure between neuron and those with environment, as well as by methodology determining of action intensity, which makes a learning process.

In the world have been developed two basic approaches of artificial neural network modeling: (1) models of simulated human behavior, which leads to artificial intelligence,
(2) neurocomputers, designed with parallel processing. Approach (1) is here a subject of our interest. Structure of biological neuron (it consists body of neuron, axon, dendrite and synapses) assures the brain communication in both ways: (1) by chemical signals across synapses, and (2) by electrical signals inside the neuron, but this activity assures cellular membrane thickness about five nanometers. Speed of discharging of neuron is determined by cumulative effects for the great number of in excitable and prohibitive (inhibitive) entries, which are processed in short time in the body of a neuron. Excitable entries increase, but prohibitive decrease in speed of discharging of a neuron. That way, the neuron signals belong to FM (frequently modulated) impulse, just like radio-communication, with advantage of great decrease of inhibition (noise).

There is an adequate likeness, but yet greater differences, between a brain and a classical digital computer. Generally, it could be said that two system, beside some similarities are different in many things. Both of them are optimized for solving of various problem types, both have basic different structure and were developed beside essential different criterias. In application of the design and application of more complex mathematical models of artificial neural networks, including approaches for description of fuzzy control and fuzzy logic.

Artificial neuron is often called, by the authors, McCulloh-Pitts neuron. Artificial neuron functions like biological neuron: goes out from other neurons and/or environment of the watched neuron, which are instructed to watch neuron, then multiply with weight factor and lead to a sumator. In sumator, products are summed and are led to the activity function input, which will give the neuron output at its belonging output, (B. Novaković, 1996).

For instance, output of neuron y, output of sumator x and activity function sinus, and result is y = sin (x). Depending on the setup criteria, we can get various categories of artificial neural networks, for instance towards: number of layers of neural network (one layer, more layers), the way of relationships of network layers, respective of the direction of moving signal (feed forward, for instance MADALINE, feedback or recurrent, by example Hopfield), time (time-continuous, time-discontinuous or discrete), main span of application (perceptions, associatives, adaptives, cognitrons, neocognitrons), method for learning (feedback propagate, opposite propagate, static). author's name (Kohonen’s, Hopfield’s, etc.), way of implementation-application (software, hardware, optics, etc.).

Now and in future, development of other criterias and categories of artificial neural network are possible. There exists a basic division on following types according to criteria of learning way of artificial neural network: learning beside supervisor - it requests external "teacher" of neural network, which is watching of network simulation, correcting till the desired network simulating is given: learning (training); process of network, in which the selected network structure is accepted and network beginning weights (parameters), then on the entry is led to a group of entry variables, that network generates in cluster of exit variables, which are compared with the desired group of exit variables, besides given network error assure calculating of new weights (parameters) by the means of selected algorithms. Procedure is repeated by iteration, till network error will not be less then feed forward of the requested rate, changing, if exists need, network structure; testing process of neural network; in which new group of network entry generates a new cluster of exit, and which is compared, beside in condition of constant weights (parameters), with the desired exit.

Network error rate services for appreciation of network generalization attributes (network capability for giving sufficient exits and for cluster of entries with which network wasn't learned, learning without a supervisor: for network learning isn't used external teacher, but neural network has been organized autonomously, and network that learned with this method is called a self-organized neural network. Procedure consists of: the
learning process, in which are lead to cluster of entry variables on network entry and network is self-organized by regulation of belonging parameters (weights) by selected algorithms. As the desired exit of the network isn't established during a learning time, learning results aren't predicted. This way, organized network can be applied for classification of entry, respectively for identification of sampler, testing process.

It exists as a series of various algorithms, for instance: generalized delta rule (network with feedback propagated; appendix with a moment; Kohonen's self-organized network), learning algorithm based on the acquainted Hebb law (Grossberg-Hebbian algorithm includes forgetfulness degree with positive-definitive Lyapun function; association network), optimal learning (setup of optimal criteria), statistical of learning method (Boltzmann's and Cauchy learning), adaptive learning (learning beside not-forgetting of acquainted), algorithms for very quick learning: (1) non-iterative learning procedure, (2) combination of non-iterative and iterative algorithm (typically representant is RBF network, (3) feed forward network with feedback propagated of error. Possibilities of implementation, beside advantages and lacks, are by means of analogue technique (better for feedback network), digital technique (better for feed forward network), hybrid, technique of impulse modulation and optics (optoelectronics).

Usage is for: signals classification, control and leading of system, robotics, identification of complex dynamic system, medicine, reading of manuscript, language translating, articulation of the English language, identification of 2 and 3 dimensions (neural vision), linear programming, optimized communication and making decision and conclusion. Interconnecting of development of the artificial intelligence and artificial neural network will cause great, now hardly assumed, changes in the first half of 21st century.

3.2. DYNAMIC NEURAL NETWORK

The basic idea of the dynamic neuron concept is to introduce some dynamics to the neuron transfer function, such that the neuron activity depends on the internal neuron states. In this study, an ARMA (Auto Regressive Moving Average) filter is integrated within the well-known static neuron model. Such filter allows the neuron to act like an infinite impulse response filter, and the neuron processes past values of its own activity an input signals. The structure of a proposed dynamic neuron model, the so-called Dynamic Elementary Processor (DEP) is plotted in Fig. 1. The filter input and output at time instant (n) are given in (1) and (2) respectively (D. Majetic, 1995):

\[ net(n) = \sum_{i=1}^{L} w_i u_i \]

\[ \tilde{y}(n) = b_r net(n) + b_j net(n-1) + b_k net(n-2) - a_r \tilde{y}(n-1) - a_j \tilde{y}(n-2) \]

The input of the neuron activation function is given in (3), and nonlinear continuous bipolar activation function is described in (4)

\[ \tilde{y}(n) = \tilde{y}(n) + w_i u_i \]

\[ y(n) = y(\tilde{y}(n)) = \frac{2}{1 + e^{-\tilde{y}(n)}} - 1 \]
The network (DNN) proposed in this study has three layers. The neuron in the first, the input layer \((i=1)\) has single input which represents the external input to the neural network. The second layer is consisted of 10 dynamic neurons, which are presented by Fig. 3. Each \(j\)-th dynamic neuron in hidden layer has an input from every neuron in the first layer, and one additional input with a fixed value of unity is usually named as Bias. The neuron in the third, output layer \((k=1)\) has an input from every neuron in the second layer and, like the second layer one additional input with fixed value of unity (Bias).

\[
\begin{align*}
E &= \frac{1}{2} \sum_{i=1}^{n} (O_j(n) - O(s))^2 \\
\end{align*}
\]  

where \(N\) is the training set size, and the error is the signal defined as difference between the desired response \(O_j(n)\) and the actual neuron response \(O(n)\). This error is propagated back to the input layer through the dynamic filters of dynamic neurons in hidden layer. Iteratively, the optimal parameters (both the weights and filter coefficients) are approximated by moving in the direction of steepest descent (J. M. Zurada, 1992):

\[
\begin{align*}
\theta_{new} &= \theta_{old} + \Delta \theta \\
\Delta \theta &= -\eta \nabla E = -\eta \frac{\partial E}{\partial \theta} \\
\end{align*}
\]

where \(\eta\) is a user-selected positive learning constant (learning rate). To accelerate the convergence of the learning algorithm given in (6), the momentum method can be applied. The method (J. M. Zurada, 1992) is given in (8) and involves supplementing the current learning parameter adjustment (7) with a fraction of the most recent parameter adjustment. This is usually done according to the formula

\[
\Delta \theta(n) = -\eta \frac{\partial E(n)}{\partial \theta} + \alpha \Delta \theta(n-1) \\
\]

where the arguments \(n\) and \(n-1\) are used to indicate the current and the most recent training step (instant time), respectively.

\[3.3. \text{DNN LEARNING ALGORITHM}\]

The goal of the learning algorithm is to adjust the neural network parameters \(\theta\) (both the weights and filter coefficients) in order to determine the optimal parameter set that minimizes a performance index \(E\) (J. M. Zurada, 1992) as follows:

\[
E = \frac{1}{2} \sum_{i=1}^{n} (O_j(n) - O(s))^2
\]
4. CURVES FITTING RESULTS

The research is performed on the object with the following features: an assembly of products is a stator of electro motor 4 AL 180 6/24, besides average lot size pieces 60-80, in selected shop factory work places and operations group are “Inserting and joining of spool in stator package” (integrity of assembly doing consists four interconnected operations in order, signed 1.1 to 1.4), and in this matter work places are stabilized according to familiar principles.

For research, realized and proclaimed 1999. in (D. Majetic & M. Car, 1999.), at first, it’s selected manual work with operation of sign 1.1 “Inserting of the first layer of spool” as a variant simplification of existing work mode (work is performed by only one operator, earlier there were two operators, besides of series improvement with purpose dodging of lacks). Between total of 41 operators on selected group of work places, an operator with 16 months of total experience on existing operations, unlike to researching, is selected. A sample of 11 cycles (stators) is established, with 15 performing each of them, total of 165 performing.

In (M. Car, 1993.) it is shown that it would be applied to priority models of linear regression Y3 (standard type of exponential function – model T. P. Wright) for establishing on time registering by chronometer at request of less accuracy fitting theoretical data to empirical. Besides, there is applied single data, with less economic than the average ones. For Y7, legality is mathematically assumed by the means of:

\[ Y_7 = 7.651165 - 0.878493 \ln x \]  \hspace{1cm} (9).

Results of fitting natural logarithmic function NLF by the means of Y7 from (T. P. Wright, 1936.), as well as dynamic neural network DNN researched in (D. Majetić, 1995.), to empirical (real) data are shown graphically in diagrams in figure 4.

Adequate attributive conclusion about fitting to empirical (real) data is possible to get also with a “graphic comparison” of single parts and entire legality in preceding indicating diagrams. Complementary to numerical data, as a supplement to the preceding diagrams, more successful comparison analysis is assured, more expressive deviation of values one
data to another (of single or series of single in some interval-age) data, as well as more accuracy and more precision concluding about comparison.

Because of cited reason, there are shown comparative numerical, empirical (real) and estimated for NLF and DNN, data in separated tables, depending on ordinal number of performing. From a cluster of empirical (real) data for single samples-operations (1.1 with n=165, 1.2 with n=131 and 132, 1.3 and 1.4 with n=132 data), there are selected, according to expectations of possible deviations in the entire learning legality, two the most affective sub clusters – intervals from age (2) (learning age) of entire legality. The first is initial, the most steep-sheer, part in which are selected all of single performances from 1 to 10, while is the second, yet always steep-sheer but less than the first, part in which are showed performances in interval from 11 (relatively / 15) to 80, with single interval by 5 units.

Because of too great extent of tables for all of 4 operations, it's separated and showed only a table 1 for operation 1.1.

Features showed in table 1.

Table 1. Comparison of empirical and estimated data for operation 1.

<table>
<thead>
<tr>
<th>Operation 1.1: time t, min, for performing ordinal number</th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
<th>5.</th>
<th>6.</th>
<th>7.</th>
<th>8.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>7,18</td>
<td>6,75</td>
<td>6,18</td>
<td>6,25</td>
<td>5,76</td>
<td>6,78</td>
<td>5,78</td>
<td>5,9</td>
</tr>
<tr>
<td>9.</td>
<td>4,4</td>
<td>3,72</td>
<td>4,1</td>
<td>4,13</td>
<td>4,08</td>
<td>3,93</td>
<td>3,53</td>
<td>3,95</td>
</tr>
<tr>
<td>45.</td>
<td>7,651</td>
<td>7,042</td>
<td>6,686</td>
<td>6,433</td>
<td>6,237</td>
<td>6,077</td>
<td>5,941</td>
<td>5,8244</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>7</td>
<td>5,941</td>
<td>7</td>
</tr>
<tr>
<td>9</td>
<td>5,720</td>
<td>5,628</td>
<td>5,272</td>
<td>5,019</td>
<td>4,823</td>
<td>4,663</td>
<td>4,527</td>
<td>4,4105</td>
</tr>
<tr>
<td>4,307</td>
<td>4,214</td>
<td>4,130</td>
<td>4,054</td>
<td>3,984</td>
<td>3,918</td>
<td>3,858</td>
<td>3,8016</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>4,307</td>
<td>4,214</td>
<td>4,130</td>
<td>4,054</td>
<td>3,984</td>
<td>3,918</td>
<td>3,858</td>
<td>3,8016</td>
</tr>
<tr>
<td>6,898</td>
<td>6,625</td>
<td>6,454</td>
<td>6,253</td>
<td>6,255</td>
<td>6,380</td>
<td>6,296</td>
<td>6,1601</td>
<td></td>
</tr>
<tr>
<td>351</td>
<td>538</td>
<td>53</td>
<td>721</td>
<td>14</td>
<td>32</td>
<td>251</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>5,839</td>
<td>5,681</td>
<td>5,352</td>
<td>5,189</td>
<td>4,981</td>
<td>4,665</td>
<td>4,414</td>
<td>4,2950</td>
<td></td>
</tr>
<tr>
<td>294</td>
<td>051</td>
<td>127</td>
<td>009</td>
<td>976</td>
<td>43</td>
<td>413</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>4,245</td>
<td>4,216</td>
<td>4,148</td>
<td>4,054</td>
<td>3,830</td>
<td>3,825</td>
<td>3,778</td>
<td>3,7479</td>
<td></td>
</tr>
<tr>
<td>636</td>
<td>13</td>
<td>563</td>
<td>757</td>
<td>656</td>
<td>544</td>
<td>264</td>
<td>72</td>
<td></td>
</tr>
</tbody>
</table>

In approximative interval n=1 to 15, time values of DNN are continuously larger than NLF ones, for 16 to 30 time values of DNN varying about NLF ones, for 31 to 60 time values of DNN are continuous less than NLF ones, while, for 61 to 80 time values of DNN continuous varying about NLF ones;

(12) operation 1.2.

Origin sample with 132 data: time values of DNN have larger amplitudes than NLF ones for interval to n=100 (single data with extreme deviation towards legality), while from 101 to 132 time values do more smooth down function.

Sample with 131 data (excluded single data on n=100), table 1.: in interval n=1 to 15 time values of DNN have small amplitudes around NLF ones (more smooth out curves), for 16 to 35 time values of DNN are less than NLF ones, for 36 to 80 time values of DNN varying around NLF ones, besides beyond n=100 they have greater amplitude, especially

32
on the end;

(13) **operation 1.3.**

In interval from 1 to 30 time values of DNN varying around NLF ones, while for 31 to 80 time values of DNN are greater than NLF ones;

(14) **operation 1.4**

In interval from 1 to 10 time values of DNN varying around NLF ones, for 11 to 20 time values of DNN are less than NLF ones, for 21 to 80 time values of DNN are larger than NLF ones;

(20) **Common (joint) for all of 4 operations ;**

(21) for empirical (real) data; boundary of time/data deviation are significantly greater in approximative interval : n=1 to 80, than resumption towards the end of legality,

(22) analogous to boundaries of time data deviation from (21), amplitudes estimated with DNN are in interval n=1 to 80 significantly larger than in prolongation towards the end of legality,

(23) time values of DNN much better follow (fitted itself, exhibit) the empirical (real) data than NLF data.

5. **CONCLUSION**

Learning phenomena is in the industrial production-manufacturing and wider of multiple importance. Because of the preceding fact, in this article have been exhibited theoretical hypothesis and serves of basic features of learning phenomena, including such features as approaches, methods and techniques and procedures of learning, as well as the most important results of the same researches.

Also, there have been exhibited theoretical hypothesis and series of following basic features of artificial neural network: definition of the artificial neural network; a biological neuron; similarity and differences of a brain and the computer; the artificial neuron; the kinds of artificial neural networks; learning of artificial neural network; implementation, application and the trend of development of artificial neural network and artificial intelligence.

The research is realized on a group of four interconnected operation 1.1. to 1.4., called “Inserting a spool in a stator package of electro motor”. There is suggested, comparatively by graphic and analytical ways, the data of natural logarithmic function NFL and dynamic neural network DNN. It can conclude the following : (1) in the first, steeped part of the entire learning legality (learning age), empirical (real) data have greatest boundary of deviation, (2) relationships of legality NFL and DNN shows that in a single parts of learning legality DNN has larger or less time values than NFL ones, or DNN time values are varying around NFL ones, (3) NFL is a good estimation for the empirical (real) data, but it makes better, with greater sensitivity, estimates of DNN, and it can be said that DNN estimates justify NFL data with a larger sensitivity.

It's desirable, in the research resumption, to check the same or some future samples with various kinds of models and algorithms of the artificial neural networks and soon, it's time to begin with the research of the learning legality predicting.
REFERENCES


[18] B. Novaković, D. Majetić, M. Široki: Artificial Neural Networks, Faculty of
Mechanical Engineering and Naval Architecture, University in Zagreb, Zagreb 1998.


Research is supported by Croatian Ministry of Science and Technology.

Received: 16 March 2007
Accepted: 16 September 2007