We are IntechOpen, the world's leading publisher of Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

ANN Modelling to Optimize Manufacturing Process

Luigi Alberto Ciro De Filippis, Livia Maria Serio,
Francesco Facchini and Giovanni Mummolo

Additional information is available at the end of the chapter

Abstract

Neural network (NN) model is an efficient and accurate tool for simulating manufacturing processes. Various authors adopted artificial neural networks (ANNs) to optimize multiresponse parameters in manufacturing processes. In most cases the adoption of ANN allows to predict the mechanical proprieties of processed products on the basis of given technological parameters. Therefore the implementation of ANN is hugely beneficial in industrial applications in order to save cost and material resources. In this chapter, following an introduction on the application of the ANN to the manufacturing process, it will be described an important study that has been published on international journals and that has investigated the use of the ANNs for the monitoring, controlling and optimization of the process. Experimental observations were collected in order to train the network and establish numerical relationships between process-related factors and mechanical features of the welded joints. Finally, an evaluation of time-costs parameters of the process, using the control of the ANN model, is conducted in order to identify the costs and the benefits of the prediction model adopted.

Keywords: modelling, simulation, control and monitoring of manufacturing processes, simulation technologies

1. Introduction

The use of artificial intelligence and specifically artificial neural networks (ANNs) has allowed yielding revolutionary advances in manufacturing. However, most of the applications of artificial intelligence in the production field concerned expert systems and fewer attentions were paid to neural networks (NNs). Most important characteristics of the ANNs are:

- the self-adaptive behaviour that allows to adapt the forecast to changing of the environment, in this way improve the networks' ability to learn and to predict;

- the parallel computing architecture, that has a great impact in multiple disciplines and applications, from speech and natural language processing, to image processing or problems in bioinformatics and biomedical engineering.

Therefore, they could be of great help for today's computer integrated manufacturing and in smart factories, according to Industry 4.0 paradigm. Currently, the nature of the manufacturing process is changing with great speed, becoming more sophisticated and continuous variations are occurring due to changes in customer demand and reduced product life cycle. This requires manufacturing technologies that can easily adapt to such changes. In this context, artificial neural networks are a powerful technology to solve this problem. The use of ANNs is also widely used for process monitoring and control applications. The quality of a process can only be provided by in process monitoring through proper measurements. To ensure a high quality of a process, you must follow the following technological steps [1]:

- Identify the characteristic changes of a process;

- To estimate the changes of the product quality;

- Correct any process operations as a result of any anomalies detected from the comparison between the obtained and the desired quality.

These steps should be followed with minimal supervision and assistance from operators, if possible in unmanned manner. In addition, all processes should be implemented with special features such as: storing information, decision making, learning and integration. It should be noted that most manufacturing processes are regulated by many variable parameters and for this reason, such systems have a random, complex and uncertain nature.

This may be attributable to the fact that they are exposed to external disturbance and noise and often subjected to parameter variations. Furthermore, there is often a great interaction between variables and therefore it is not possible to properly define the final quality of the product and the variables that influence it. Due to these characteristics the quality often varies from product to product, impairing its uniformity and decreasing the yield of the product. So, all changes that may occur in manufacturing environments cannot be easily observed by an operator, so in recent years the use of neural networks applied to process monitoring and control has been of great interest. Indeed, it has been shown that the use of artificial intelligence can overcome the above-mentioned problems [2]. The research efforts in this direction will be accelerated with greater interest in the future and will lead to the development of truly intelligent manufacturing systems that are capable of producing products without the supervision or assistance of human operators [1]. **Figure 1** classifies the functionalities needed to imbed the artificial neural networks on manufacturing processes and summarizes the current developments in manufacturing application areas.

The real world applications here in manufacturing include the modelling, monitoring and control, identification, planning and scheduling associated with the processes. The purpose of this chapter is to present some applications of artificial neural networks in manufacturing process monitoring and control, among which particular attention will be paid to the study that has been published in international journals and that has investigated the use of the
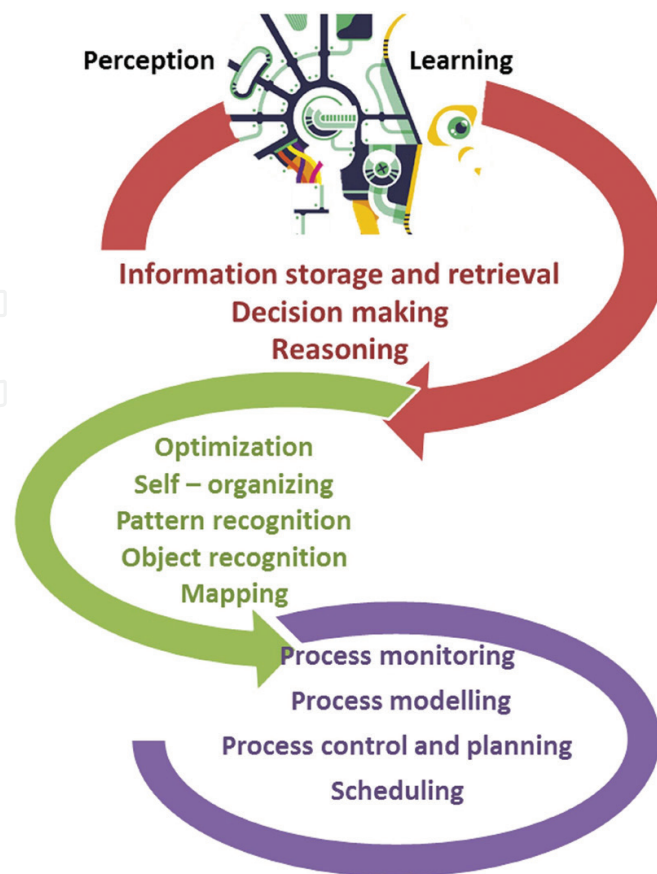
Figure 1. Functionality of the artificial neural networks and their manufacturing applications.

ANNs for the monitoring, control and optimization of a welding processes. Experimental observations were collected in order to train the network and establish numerical relationships between process-related factors and mechanical properties of the welded joints. Finally, an evaluation of the time-cost parameters of the process, using the control of the ANN model, is conducted in order to identify the costs and benefits of the prediction model adopted.

2. Manufacturing applications

Due to many external disturbances and many variations in process parameters, many production processes are complex and time-consuming. For these reasons, it is not always possible to identify the relationship between the product quality and the input variables of the process. Thus, there is interest to integrate the artificial intelligence into the production processes for storing, learning, reasoning and decision making. Such systems are able to adapt to changes in its environment and can truly realize unmanned operations of processes. The adoption of the neural network can be devoted to monitoring and to prediction of different parameters in many industrial areas, in order to solve issue relating to the manufacturing system design, process planning, as well as operational decision making. A summary of main NN applications field is shown in **Table 1**.

Category	Manufacturing		
Application topic	• Manufacturing system design	Author Ref. num.	Moon and Chi [53] Kaparthi and Suresh [54] Lee et al. [55] Moon [56] Wu [57]
	• Manufacturing process control		Cook and Shannon [58] Wu [57]
	• Robot scheduling		Yih et al. [59]
	• Manufacturing operational decision		Chryssolouris et al. [60]

Table 1. Summary of neural network manufacturing applications.

Manufacturing information, such as the sequence of operations, lot size; multiple process plans were given special consideration in their approach to solve the generalized part family formation problem. Many authors also point out that the method of artificial neural networks is flexible and can be efficiently integrated with other manufacturing functions. Below, from the literature, some important artificial intelligence applications to particular production processes are described.

2.1. Injection moulding processes

Injection moulding processes are characterized by dynamic characteristics since process input variables are the melting temperatures, the velocity of the cylinder, the holding, the pressure that produces the polymer flow into the model cavity and they vary in a complex manner. The phenomena occurring in the process are very complex, time-varying, nonlinear and uncertain. This complexity makes it difficult to relate the input operating variables to the product quality such as geometry accuracy and geometry surface smoothness. These processes have been implemented and optimized with the use of artificial intelligence with the use of *multilayer perceptron* which is found to be the most popular network and tries to model the process dynamics and based on this to predict the part quality [3–9].

2.2. Gas metal arc welding processes

In gas metal arc (GMA) welding processes [10–12], the flow of an electric current is generated by an electric arc that is maintained between the consumable wire electrode and the welding metal as shown in **Figure 2**.

Both the filling metal and the consumable electrode are automatically fed by a wire feeding device. A good quality of the welds is determined by the relatively high depth to width ratio of the molten welding pool. So, the monitoring and control of weld geometry and the surface temperatures that are strongly related with the formation of the weld pool, are very important for the penetration depth or the back bead width. For this purpose, the temperatures by noncontact were measured. Infrared temperature sensing system and recent studies conducted the ANN multilayer perceptron to detect and control with great success all the surface temperature information.
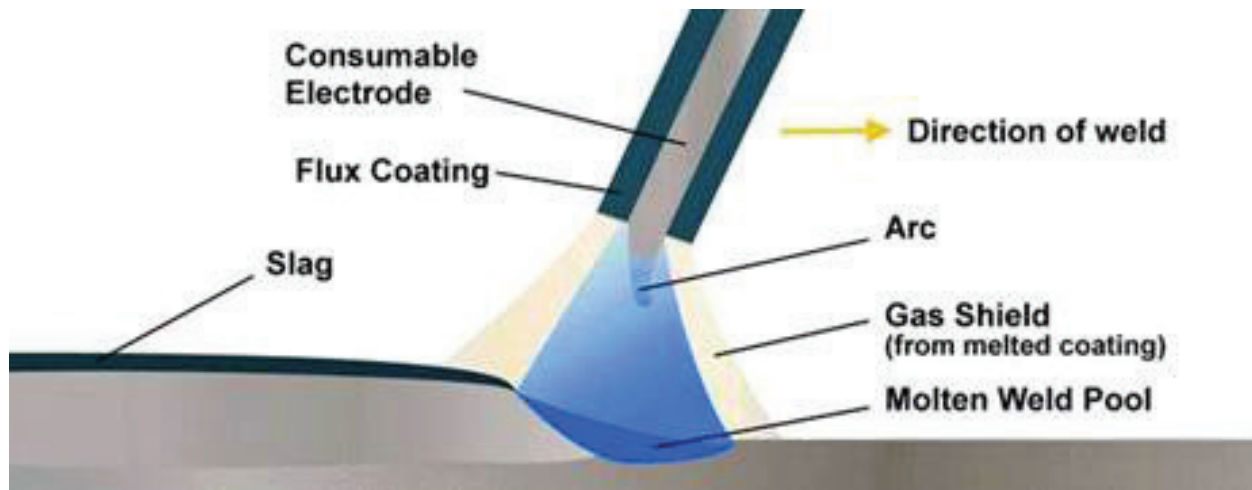
Figure 2. Characteristics of the GMA welding process.

2.3. Arc welding processes

The complexity of the relationship between the process variable and weld quality is the common factor in all manufacturing processes and in particular arc welding processes. For these reasons, the literature documents with some interesting researches, the use of ANNs for quality monitoring and control of the process. In this type of welding process, the ANN input data are generally the surface temperature, the welding voltage and the current and torch speed. The majority type of ANN employed for this case was, again, the multilayer perceptron [10, 13–21]. Their use was found to be very satisfactory to predict the weld defects, the geometry such as bead width, head height and penetration.

2.4. Machining processes

To perform a correct quality control in all machining systems, it is very important to check particular parameters such as the cutting tool state, the vibrations, the forces and the temperature obtained during real-time machining operation. To optimize this particular type of supervision of the cutting tool state, some authors document in literature the use of the ANNs which used the above-mentioned process data to classify the status of tool wear, prediction tool life and detect tool failure in an on-line manner. Examples of typical sensors are tool dynamometers, acoustic emission sensor accelerometers and thermocouples. In this process, the networks in frequent use are the multilayer perceptron and Kohonen [22–46].

## 2.5. Semiconductor manufacturing processes

The complexity of plasma etching processes in integrated circuits fabrication promoted the use of ANNs for monitoring and control. In this field, the use of artificial intelligence brings advantages that could not be achieved with traditional open loop controls. Where used the multilayer perceptron networks that are the most popular for this process. The ANNs use the fundamental parameters that affecting process dynamics, such as power, gas flow rate, dc bias voltage and throttle positioning. Proper use of networks in this field

allows real-time monitoring and estimation of quality variables such as the etching thickness and the etching time [47–52].

3. Development and implementation of ANN

The origin of artificial neural networks is based on learning technique that mimics the biological learning process occurring in the brain. Neural networks present a robust way to predict an actual value after a learning activity from a supplied sample set [61]. The ANNs are based on a concept that combines a set of computational procedures with a theoretical basis in order to predict the unknown output parameter in various processes. Generally, neural networks are adopted to subordinate knowledge to observations or when data or activity is so complex that is not allows to identify an optimal solution in a reasonable time. It is difficult in each field of application and even for each task, to compare the use of neural networks versus other prediction techniques (e.g., statistical methods or a support vector machine) because, on the contrary to the conventional computational techniques, they are able to solve nonlinear and ill-defined problems. Many factors underlie this trend, most of them are related to reliability of the predictions, to the robustness and adaptability of the results as well as the learning ability of the neural process. In many cases the forecasts generated by ANNs, if correctly designed, significantly improve with increasing of dataset used as training subset. Consistently, in last years, the adoption of the ANNs in many business areas is increased exponentially and the number of publications, in high-level journals, was grown [62] (**Figure 3**).
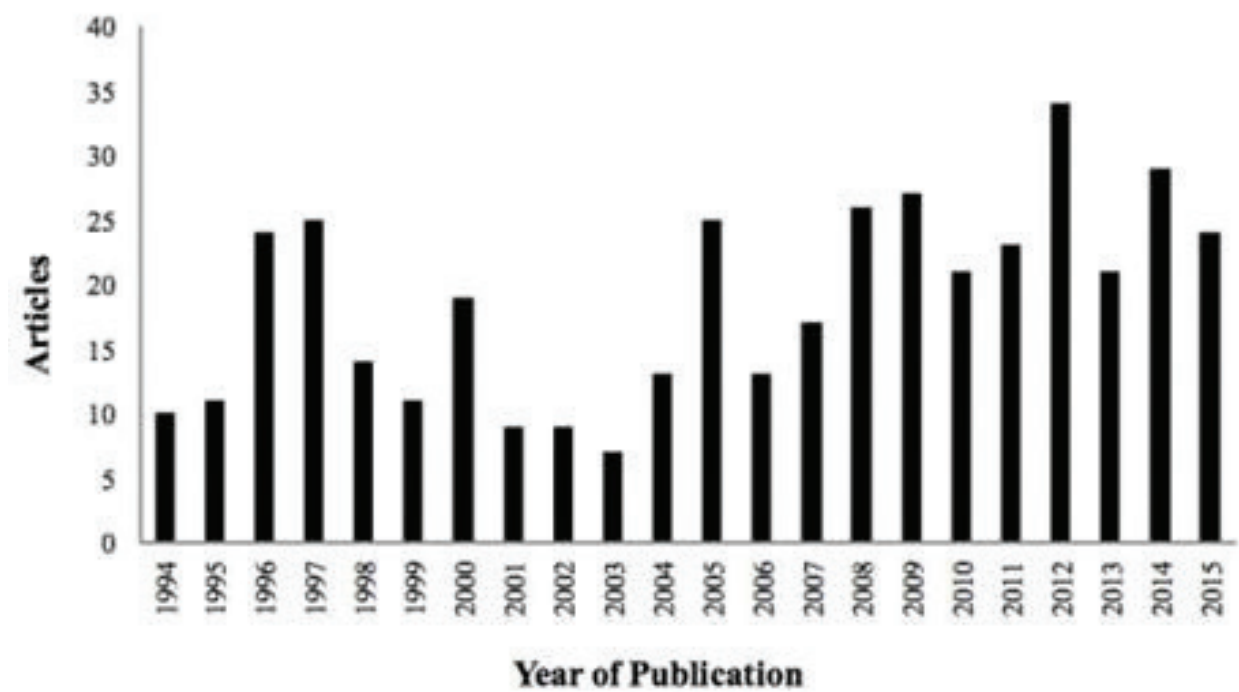
Figure 3. Distribution of ANN-papers by year.

Under engineering perspective, a 'good' ANN is based on models able to imitating the proprieties of natural systems, such as cognitive capabilities, flexibility, robustness, ability to learn and fault tolerance. At this scope the structure and the behaviour of the ANN required a study characterized by different hierarchical levels of organization as neurons, layers, synapses and cognition-behaviour functions. Different areas of application are interested by the ANNs, some of them are astronomy, mathematics, physics chemistry, earth and space sciences, life and medical science and engineering. In recent years USA and EU countries, have approved different initiatives for the study of the human brain in these cases, the ANN, in various forms and at different levels, has been included in thesis research projects. The inter-disciplinary given by system adopted for dataset analysis and by the complexity computational required by the elaboration of the data, allows to design and simulate systems capable to satisfying the needs and the challenges of the real world. Japan in 2014 has been developed a project based named as Brain Mapping by Integrated Neurotechnologies for Disease Studies (Brain/MINDS) [63], that will be integrated with new biomedical technologies and neural network systems. In Australia, a specific programme has also been set up with preliminary funds of around $250 million over 10 years with the goal of developing the world's first bionic brain (AusBrain) [64] based on multilayer perceptron (MLP) system. There is also another ambitious initiative in China (Brainnetome) [65], the goals of this are to simulate the brain networks for perception, memory, emotion and their disorders as well as to develop advanced technologies to achieve these goals.

3.1. Designing the ANN

An ANN is a computational model that establishes a relationship between process factors and output variables. Artificial neurons are combined through weights, which work as adjustable coefficients. There are many programs and frameworks, either of general purpose or that simulates functions or neural structures (e.g., IQR, NeuroSpaces, NNET, etc.) but there is not a specific simulator that is currently being used by the whole community since some different approaches are more suitable than others, on the basis of the research task being addressed. Moreover, most simulators can take full advantage of their computational capabilities on the basis of the features of the computer hardware to which it is installed [66].

This correlation depends by the fundamental features of the network, which define the way input and output are connected to each other [67]. The network includes input layer, output layer and a certain number of hidden layers. The fundamental features of the network are:

- Dataset splitting, which identifies the subset data to be adopted for the training, the testing and the validation of the ANN development;

- Architecture, which determines the connections between layers and neurons;

- Learning algorithm, which determines the weights of the links between neurons.

In the following sections the data splitting strategy, the architecture design approach and the learning algorithm identification, is described.

3.1.1. Dataset splitting

The appropriate data splitting can be handled as a statistical sampling problem. Therefore, various classical sampling techniques can be adopted in order to split the data in three subset for training, validation and testing of ANN, most commons are: Simple random sampling (SRS), Trial-and-error methods, Systematic sampling and Convenience sampling. The splitting strategy tries to overcome the high variance of the SRS by repeating the random sampling several times in order to minimize the mean square error (MSE) of the ANN. This technique is high time-consuming and requires significant computational costs. A subset (generally as big as 60% of the available experimental data composed by inputs/output pairs) is used for the ANN training. In this phase, the synaptic weights, which are the links between neurons, have a synaptic weight attached. They are updated repeatedly in order to reduce the error between the experimental outputs and the associated forecasts. A subset (generally as big as 20% of the available experimental data) is adopted for the ANN validation. In particular the validation sets allows to identifying the underlying trend of the training data subset. A subset (generally as big as 20% of the available experimental data) is adopted for testing the forecast reliability of the ANN in the learning phase. In order to deal with the overfitting problem that occurs when the network has memorized the training examples, but it has not learned to generalize to new situations, different approaches are suggested: reduce the number of hidden layers, improve the 'quality' of the training-subset adopted, introduce some noisy data into training set, etc. In Ref. [67] an efficient method is proposed for model establishment by means the identification of a low-dimension ANN learning matrix through the principal component analysis (PCA).

3.1.2. Network architectures

An ANN is a computational model that establishes a relationship between process factors and output variables. Artificial neurons are combined through weights, which work as adjustable coefficients. This correlation depends by the fundamental features of the network, which define the way input and output are connected to each other [67]. The network includes input layer, output layer and a certain number of hidden layers (**Figure 4**). The fundamental steps for the development of a network are:

- Dataset splitting, which identifies the subset data to be adopted for the training, the testing and the validation steps of the ANN development;

- Architecture, which determines the connections between layers and neurons;

- The learning algorithm, which determines the weights of the links between neurons.

Based on the connection pattern (architecture), ANNs can be grouped into two categories:

- Feed-forward networks (e.g., single-layer perceptron, multilayer perceptron, radial basis function nets) in which there are not network connection as loops (as shown in **Figure 4**);

- Recurrent (or feedback) networks, in which different loops occur in network connections (e.g., competitive networks, Kohonen's SOM, Hopfield network, ART models).
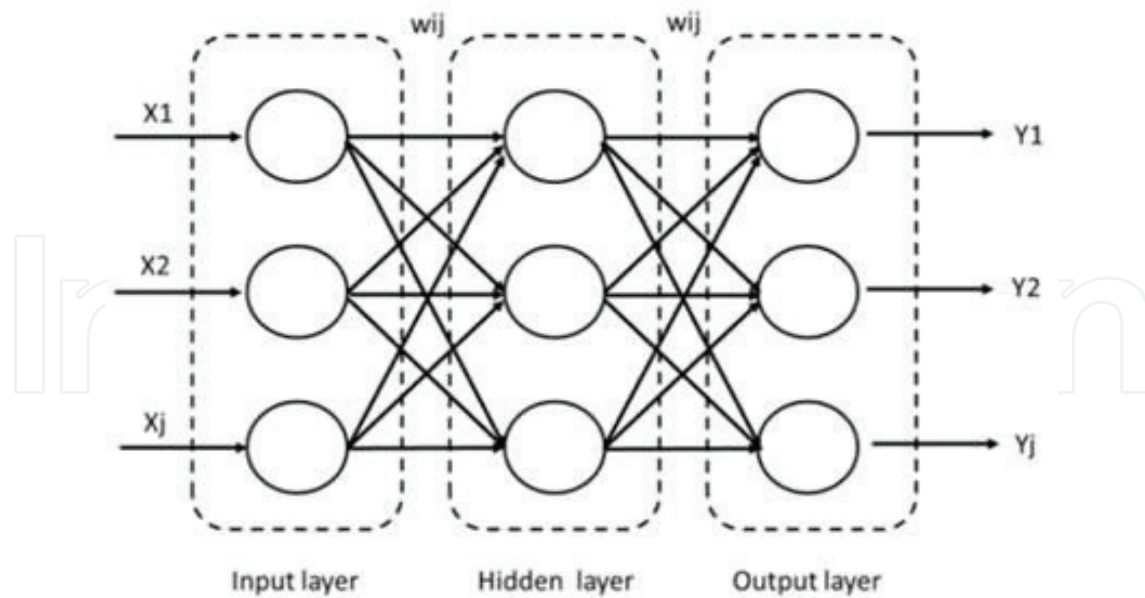
Figure 4. Structure of an ANN.

The first one network are considered "static", in fact they produce only one set of output values rather than a sequence of values from a given input, and they worked in memory-less condition, this means that their response to an input is independent of the previous network state. The 'recurrent networks', on the other hand, are dynamic systems, in which the input pattern leads the network to enter in a new state, when a new input is introduced. Most popular network architecture in use today is the multilayer perceptron neural network (feed-forward network) where the output of a previous layer is the input to the next layer. In this case a biased sum of the weights assigned to different inputs, allows identifying the activation level that, through a transfer function, produces the corresponding output. The network thus has a simple interpretation as a form of input-output model, with the weights and thresholds (biases) the free parameters of the model [68]. The design of the ANN architecture consists of identifying the kind of the structure (between feed-forward and recurrent architectures) and identifies the number of hidden layers and the number of neurons for each layer. On one hand, many neurons can lead to memorize the training sets with lost of the ANN's capability to generalize. On the other hand, a lack of neurons can inhibit the appropriate pattern classification. Many software allows to identify the best number of hidden layer and neurons (for each layer) through a 'trial-and-error' approach. In this case different architectures are iterative tested by software and for each of them, the software provide a "fitness bar" based on the inverse of the mean absolute error (MAE) computed on the testing set. In most cases the higher "fitness bar" identifies the best architecture.

3.1.3. Learning algorithm

The purpose of the learning algorithm is to train the network to predict the output parameter(s) given one or more input parameter(s). There are many types of neural network learning rules. There are three kind of learning algorithm, the first is known as supervised learning, in this

case the algorithm allows to predict the output parameter on the basis of a set of known input-output pairs [69, 70]. Second algorithm is unsupervised learning, in this case the output is not given, the aim consisting of inferring a function in order to describe a hidden structure (e.g., clustering, anomaly detection, etc.). Therefore the output parameters are considered 'unlabelled' (the observations are not classified) and is not provided any evaluation about the prediction reliability ensured by the ANN [71]. Third algorithm is named reinforcement learning, in this case a continue interaction between the learning system and the environment allows to identify the input-output mapping minimizing the performance scalar index. The approach is very similar to unsupervised learning (also in this case there are not given input-output pairs), reward or punishment signals are adopted for the prediction of output parameters [72]. In most cases, the unsupervised learning allows to ensuring lower cost function. Three different methods, usually considered to be supervised learning methods, are described in this work: Quick Propagation (QP), Conjugate Gradient (CG) and Levenberg-Marquardt algorithm (LM).

QP is a heuristic modification of the standard back propagation, the output of the mth output node for the pth input pattern is given by o_{pm} (Eq. (1)).

$$o_{pm} = f\left(\sum_{k=1}^{K} \overline{\omega}_{km} o_{pk}\right) \qquad (1)$$

where f is the activation sigmoidal function (Eq. (2)), $\overline{\omega}_{km}$ is the weight between the mth output neuron and the kth hidden neuron. The value of o_{pk} depends by two parameters: the first is given by the weight between kth hidden neuron and the nth input neuron ($\overline{\omega}_{nk}$). The second parameter is x_{pn} given by pth input pattern of nth neuron.

$$f(x) = \frac{1}{(1 + e^{-x})} \qquad (2)$$

All network weights are updated after presenting each pattern from the learning data set.

As far as concern CG method, the learning algorithm starts with a random weight vector that is iteratively updated according the direction of the greatest rate of decrease of the error evaluated as $\omega^{(\tau)}$ in Eq. (3).

$$\Delta\omega^{(\tau)} = -\eta\nabla E_{\omega(\tau)} \qquad (3)$$

where E is the error function evaluated at $\omega^{(\tau)}$ and η is the arbitrary learning rate parameter. For each step (τ) the gradient is re-evaluated in order to reduce E. The performance of the gradient descent algorithm is very sensitive to the proper setting of the learning rate, in case η is too high the algorithm can oscillate and become unstable, for η too small the algorithm takes too long to converge. In this case an adaptive learning rate allows to keep the learning step size as large as possible, ensuring, in this way, the learning rate stable. The LM algorithm allows to minimize the squares of the differences (E) between the desirable output, identified as $y_d(t)$, and the predicted output $y_p(t)$ [73]. 'E' is given by the follow equation:

$$E = \frac{1}{2} \sum_{t=1} \left(y_p(t) - y_d(t) \right)^2 \tag{4}$$

LM algorithm is also adopted, which blends the 'Steepest Descent' method and the 'Gauss-Newton', therefore it can converge well even if the error surface is much more complex than the quadratic situation; ensuring, in many cases, speed and stability. LM algorithm can be presented as:

$$w_{k+1} = w_k - \left(J_k^T J_k + \mu I \right)^{-1} J_k e_k \tag{5}$$

where J is Jacobian matrix, μ is the 'combination coefficient' (always positive), I is the identity matrix and e represents the error vector. When μ is very small (nearly zero), Gauss-Newton algorithm is used. On the other hand, when μ is very large, steepest descent method is used.

4. Case study: "Prediction of the Vickers microhardness and ultimate tensile strength of AA5754 H111 friction stir welding butt joints using artificial neural network"

Among the artificial neural networks applications to the production processes, this section describes the research of De Filippis et al. [74] in which a simulation model was developed for the monitoring, controlling and optimization of a particular solid-state welding process called friction stir welding (FSW). The approach based on the use of neural networks, using the FSW technique, has allowed identifying the relationships between the process parameters (input variable) and the mechanical properties (output responses) of the AA5754 H111 welded joints. The optimization of the technological parameters has been developed with the aim to produce a stable welding process that can provide welded joints with no defects. The experimental plans that were tested have been constructed by varying the following parameters:

- Tool rotation speed;

- Travel tool speed;

- Position of the samples extracted from the weld bead;

- Thermal data, detected with thermographic techniques for on-line control of the joints.

The quality of welded joints was evaluated through the following destructive and non-destructive tests:

- Visual tests;

- Macro graphic analysis;

- Tensile tests;

- Indentation Vickers hardness tests

- Thermographic controls.

The simulation model was based on the adoption of artificial neural networks (ANNs) using a back-propagation learning algorithm. Different types of architecture were analysed, which were able to predict with good reliability the FSW process parameters for the welding of the AA5754 H111 in Butt-Joint configuration.

### 4.1. About the friction stir welding process

The process of friction stir welding (FSW) is a solid-state welding method based on frictional and stirring phenomena, which was discovered and patented by the Welding Institute of Cambridge in 1999. In this process, a rotating non-consumable tool that plunges into the work piece and moves forward produces the heat necessary to weld the parts together. Therefore, given the particular geometry of the tool used, as shown in **Figure 5**, the following actions are performed in the process:

- The tool shoulder generates heat with the base material,

- The tool pin generates plastic deformation and mixing of the material.

The much lower temperatures compared with those achieved in traditional welding processes by melting, determine the main advantages of this process. In fact, there is minimal mechanical distortion, with minimal Heat Affected Zone (HAZ), and an excellent surface finish.

No crack formation and porosity right after welding thanks to the low input of total heat.

The main parameters of the friction stir welding process are the tool rotation speed (n) and the tool travel speed (v). The friction stir welding process enjoys major successful applications
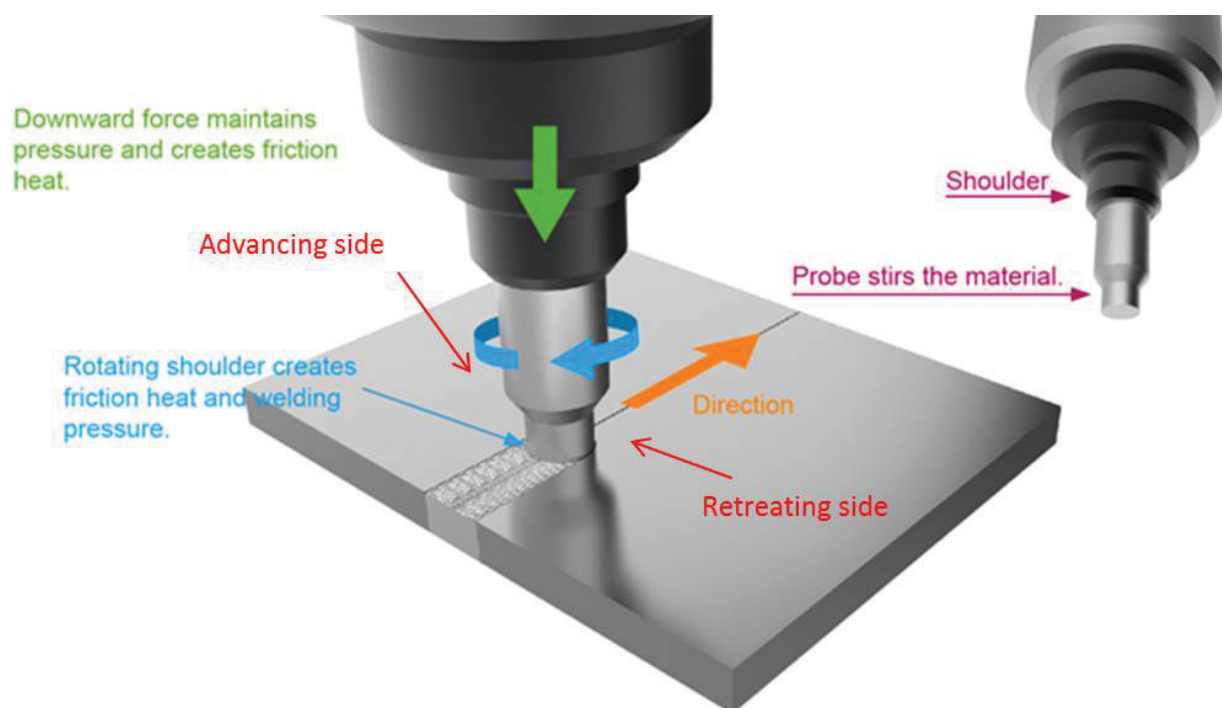

Figure 5. Schematic diagram of the FSW process.

in many fields, such as aeronautics, aerospace, rail, automotive, computer science, marine, chemical and petrochemical industries. Important advantages are also documented in the application of FSW processes on dissimilar materials, Al alloys, Cu alloys, Ti alloys and steel. The main applications are on aluminium alloys because these materials, due to their high strength-to-weight ratio, low density, forming properties, low cost and recyclability, are the main metals used in automotive, marine and aerospace applications. An aluminium alloy of large aeronautical and automotive interest is the AA5754 H111; on this material, there are still few researches about the advantages to apply the FSW. It has been shown that the mechanical properties of the AA5xxx friction stir welded joints depend mainly on the grain size and the dislocation density, due to the phenomena of plastic deformation and recrystallization occurring during the FSW process. To study and optimize properly the friction stir welding process, it is necessary to know the influence of process parameters on the mechanical properties of the joints. In general, the traditional process control techniques cannot provide information about the performance of the process during welding and require lengthy testing times, making them feasible for the industrial field. Therefore, in the production engineering, the control and the optimization of the manufacturing processes is becoming increasingly important. For the FSW process is necessary to carry out a control of the significant variables, in addition to the use of thermographic techniques. This justifies the deepening and use of information technology for enhancing the quality of manufacturing systems. The implementation of numerical and analytical models can reduce time and cost for experiment and analysis through quantitative solutions.

4.2. ANN simulation model for the monitoring of the friction stir welding process

The research interest to developing new technological tools for the control and optimization of manufacturing processes is growing. Such developments are crucial elements for the production engineering. Within the FSW process, many experiments are needed to understand the process-related dynamics and to control all the significant variables and the thermographic techniques are a valuable help but it is necessary to increase and optimize control techniques with new information tools for enhancing the quality of manufacturing systems. The reduction in time and cost of the experiments can be reduced by the implementation of numerical and analytical models. Thus the relationship between the process parameters and the quality of the weld can be easily identified by a model based on the adoption of one or more artificial neural networks (ANNs). Neural networks software packages are very common among scientists and manufacturing researchers. In particular, their applications in the field of welding have showed good success. As far as concern the FSW process, in scientific literature, there are only few papers that discuss the modelling of this welding process by a neural network [75–79]. In particular, a very interesting work is the study of Shojaeefard et al. [80] who studied the adoption of the neural network trained with Particle Swarm Optimization (PSO) for modelling and forecasting of the mechanical properties of the friction stir welding butt joints in AA7075/AA5083. A further contribution is provided by Asadi et al. [81], which with the use of ANN found a relationship between the grain size and the hardness of nanocomposites in FSW process. In this particular case study, an effective simulation model was developed for predicting, monitoring and controlling the mechanical properties of welded AA5754 H111

plates, using the ANNs with the FSW process parameters as input variables. The data set for training, testing and validation of the ANN were the results obtained by experimental cases [82–84], in which all welded joints were performed by non-destructive (visual inspection) and destructive testing (macrographic tests). These tests have been useful for detecting macro defects present on the surface and within the welded area. An accurate quantitative analysis of the FSW process was carried out using the results of the destructive tests of each welded specimen in terms of ultimate tensile strength (UTS) and Vickers micro hardness. Thermographic techniques were used to study the thermal behaviour of FSW process. In the thermal analyses, two thermal parameters were considered: the maximum temperature and the slope of the heating curve measured during the FSW process, along the two sides of the weld ($MSHC_{RS}$ and $MSHC_{AS}$, respectively). The analysis established that there is a correlation between the data derived from the thermographic controls and the quality of the welded joints, in terms of UTS. Thus this work defines the importance and effectiveness of the use of infrared technology for monitoring the FSW process in a quantitative manner, giving important information on the thermal behaviour of joints during the process. Finally, the purpose of this case study was to correlate the mechanical properties of welded joints in terms of UTS and microhardness to the thermal parameters with the use of the ANN. The results obtained have defined a model with the use of the neural networks that can predict quantitatively the mechanical behaviour of the FSW joints, as shown in **Figure 6**.
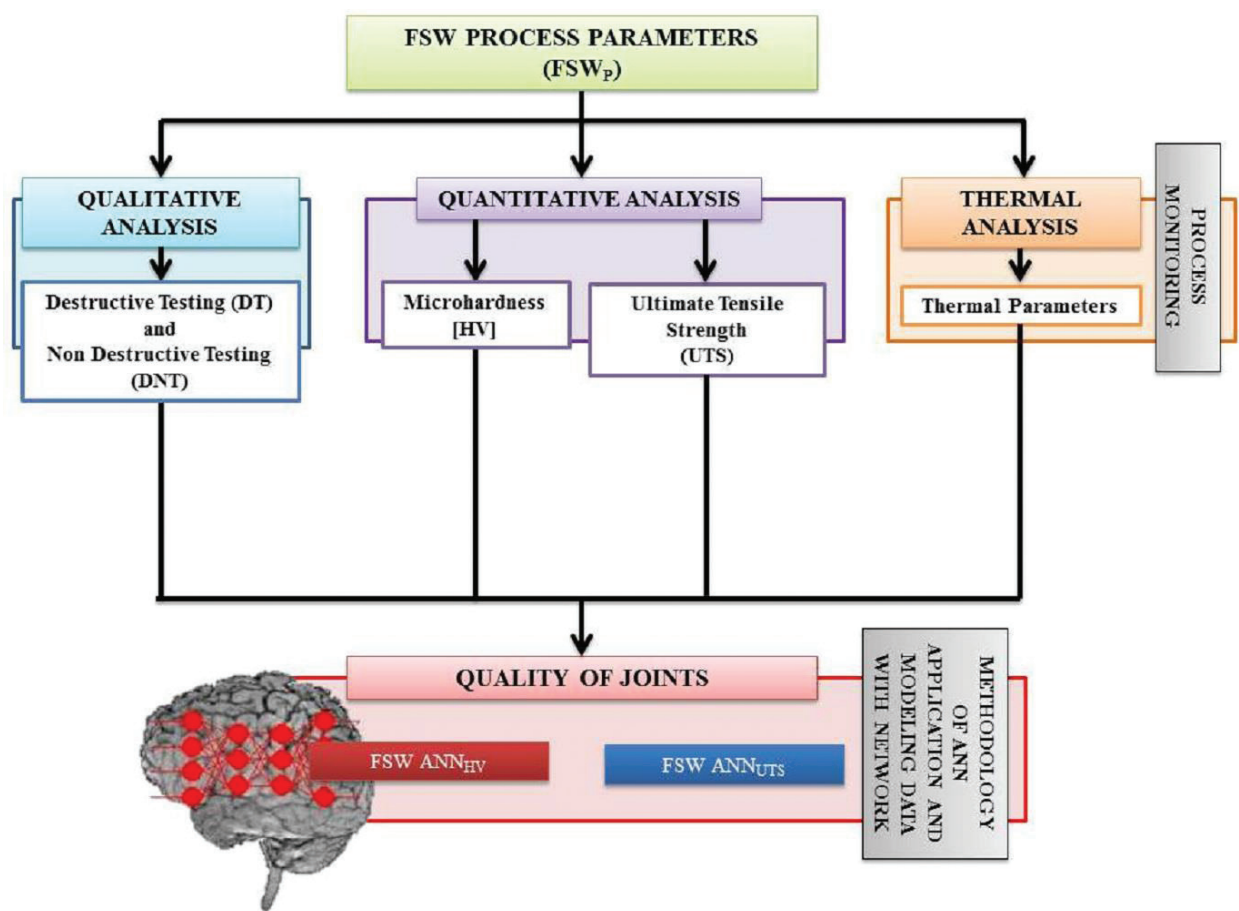
Figure 6. Approaches used for evaluating the quality of FSW process through destructive and non-destructive tests.

The results of all tests are summarized in **Table 2** and the same data were used to train the ANN.

In order to establish a relationship between the mechanical properties of the FSW joints and the process parameters, a simulation model was developed. In the development of the model two different ANNs were used, as follows: in the first network, called "ANN$_{HV}$", was used as output variable, Vickers micro hardness of HAZ; in the second network, called "ANN$_{UTS}$" was used as the output variable the ultimate tensile strength. Both have used process parameters as inputs. The ANNs were implemented using Alyuda NeuroIntelligence™-Neural networks software (2.2, Alyuda Research Company, LLC., Cupertino, CA, USA). The first network ANN$_{HV}$ was developed with five input nodes (n, v, p, $MSHC_{RS}$, and $MSHC_{AS}$) and only one response node

Input					Output			
n^1[RPM]	v^2[cm/min]	p^3[mm]	$MSHC_{RS}{}^4$ [°]	$MSHC_{AS}{}^5$ [°]	$HV_{haz}{}^6$	$HV_{haz\ norm.}{}^7$	UTS^8[MPa]	$UTS_{norm.}{}^9$[MPa]
20	500	20	86,05	85,83	60,88	0,50	166,69	1,00
30	700	20	87,37	87,37	61,93	0,70	70,25	0,21
20	700	20	87,12	87,92	63,33	0,97	120,75	0,62
30	500	20	86,88	86,85	61,72	0,66	80,05	0,29
20	500	120	87,25	86,80	60,88	0,50	90,66	0,38
30	700	120	88,14	88,15	61,93	0,70	44,29	0,00
20	700	120	87,23	87,91	63,33	0,97	56,06	0,10
30	500	120	87,97	87,91	61,72	0,66	71,99	0,23
20	500	20	86,60	86,16	62,02	0,72	132,43	0,72
30	700	20	87,74	88,23	62,72	0,85	114,87	0,58
20	700	20	86,53	88,00	58,23	0,00	51,86	0,06
30	500	20	89,07	87,23	63,50	1,00	97,43	0,43
20	500	120	87,59	87,43	62,02	0,72	99,06	0,45
30	700	120	88,53	88,32	62,72	0,85	59,95	0,13
20	700	120	87,48	87,74	58,23	0,00	46,55	0,02
30	500	120	87,58	88,45	63,50	1,00	113,98	0,57

[1]Tool rotation speed.

[2]Tool travel speed.

[3]Position of the sample along the welding direction.

[4]Maximum Slope of Heating Curve of thermal profiles evaluated on the surface of joints along the retreating side.

[5]Maximum Slope of Heating Curve of thermal profiles evaluated on the surface of joints along the advancing side.

[6]Vickers microhardness values measured in the HAZ.

[7]Vickers microhardness normalized values measured in the HAZ.

[8]Ultimate tensile strength values.

[9]Ultimate tensile strength normalized values.

Table 2. Measured data used to train the ANN.

(output node), identified as the micro hardness of the Heat Affected Zone of the welds, HVhaz. "Trial-and-error approach" was used to investigate and analyze more than 1000 different network architectures to identify the best architecture for the first network (ANN$_{HV}$). The network fitness score was calculated for each network with different design (number of hidden layers, number of nodes, etc.), based on the inverse of the mean absolute error (MAE) on the testing set. The best network architecture has been identified with the higher fitness score. The best accuracy and minimum prediction error was obtained by adopting an ANN$_{HV}$ characterized by only one hidden layer with 12 neurons as shown in **Figures 7–9**.

The second network (ANN$_{UTS}$) was developed with six input nodes (n, v, p, $MSHC_{RS}$, $MSHC_{AS}$ and HV_{haz}) and one response node (output) identified as the ultimate tensile strength of the welds (UTS). Even in this second analysis the methodology chosen for identifying the architecture of the network are the same of the ANN$_{HV}$. The "best" reliability of the prediction was achieved, adopting an ANN$_{UTS}$ characterized by only one hidden layer with four neurons, as shown in **Figures 10–12**.

The reliability of the estimation of the mechanical properties predicting by the ANN simulation model was evaluated by comparing the data with the experimental results. For this purpose, the mean absolute percentage error (MAPE) was calculated for the two ANNs modelled. **Table 3** summarizes the results of this analysis that demonstrated that the values derived from ANN simulation have a higher level of reliability. Therefore, the neural networks were able to predict, with significant accuracy, the mechanical properties of the friction stir welding joints, under a given set of welding conditions.
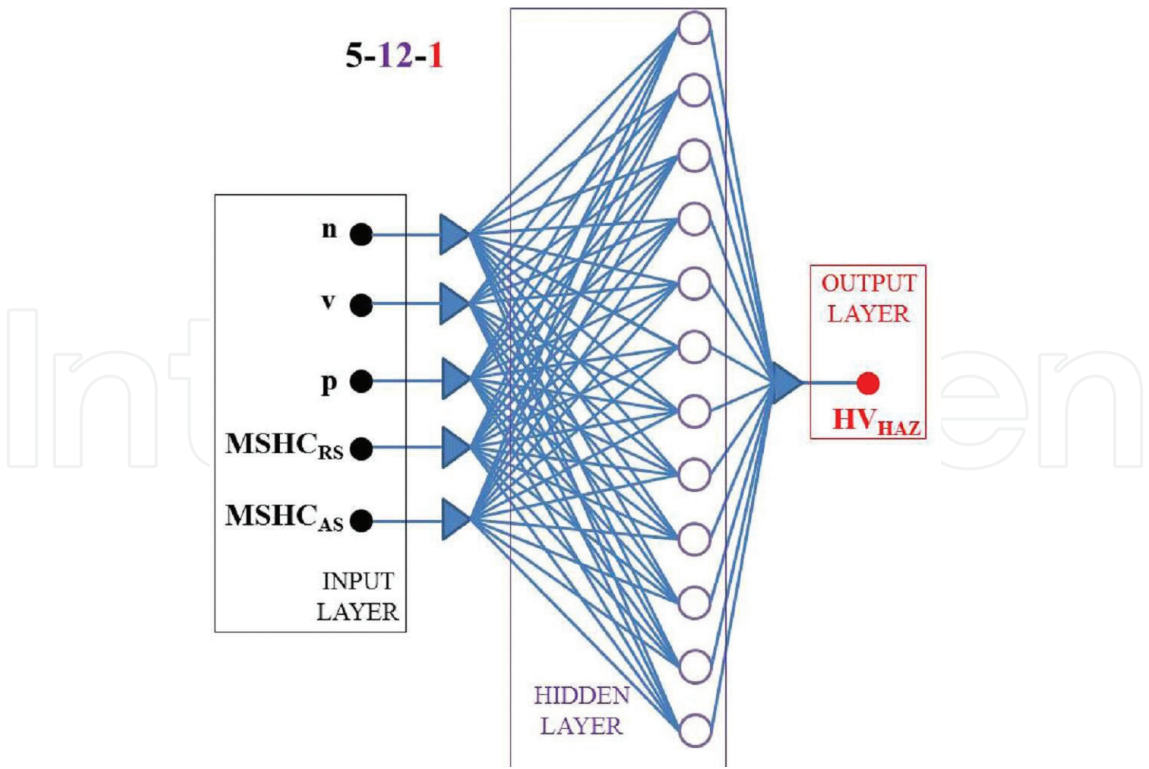
Figure 7. Back-propagation neural network used to foresee the Vickers micro hardness of the Heat Affected Zone (HAZ).
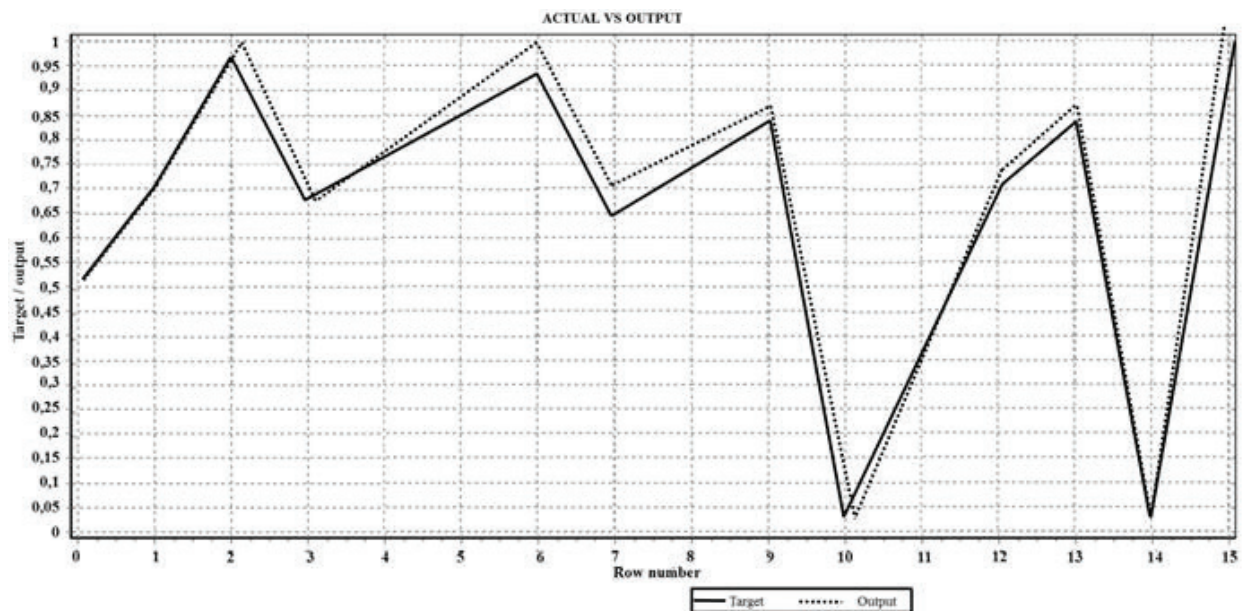
Figure 8. Predicted Vickers HAZ micro hardness by ANN versus the experimental data.
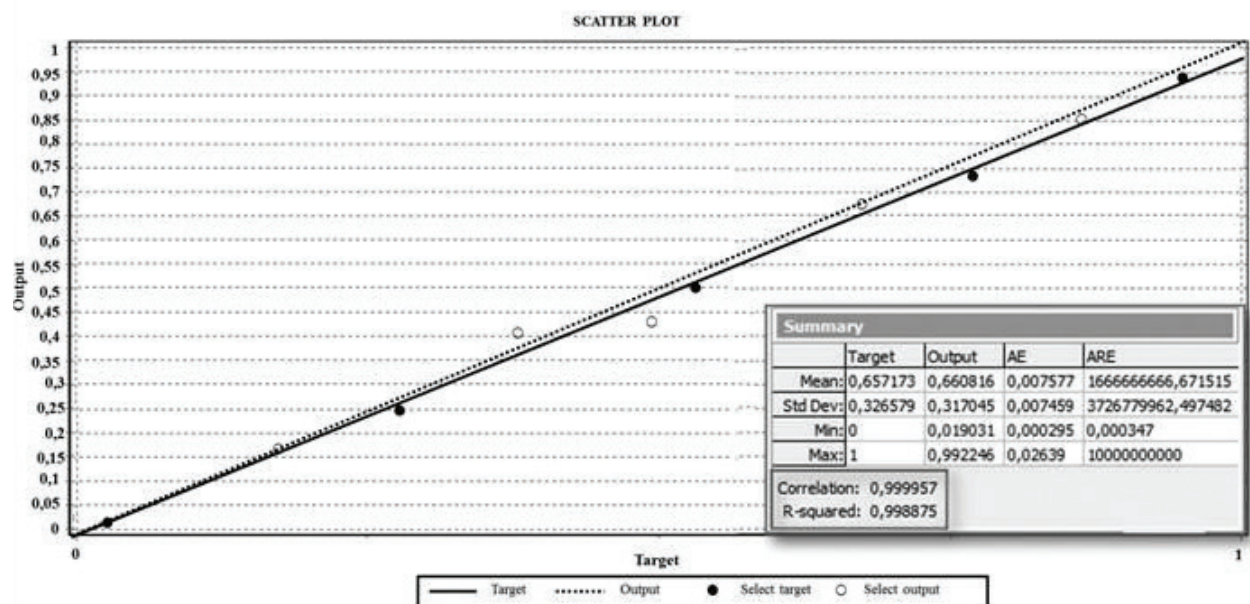

Figure 9. Regression line at the training stage.

Starting from the results obtained in previous researches [84] where the most significant FSW process parameters for AA5754 H111 plates were identified, the development of this ANN model could be used to identify the optimal process parameter setting in order to achieve the desired welding quality.
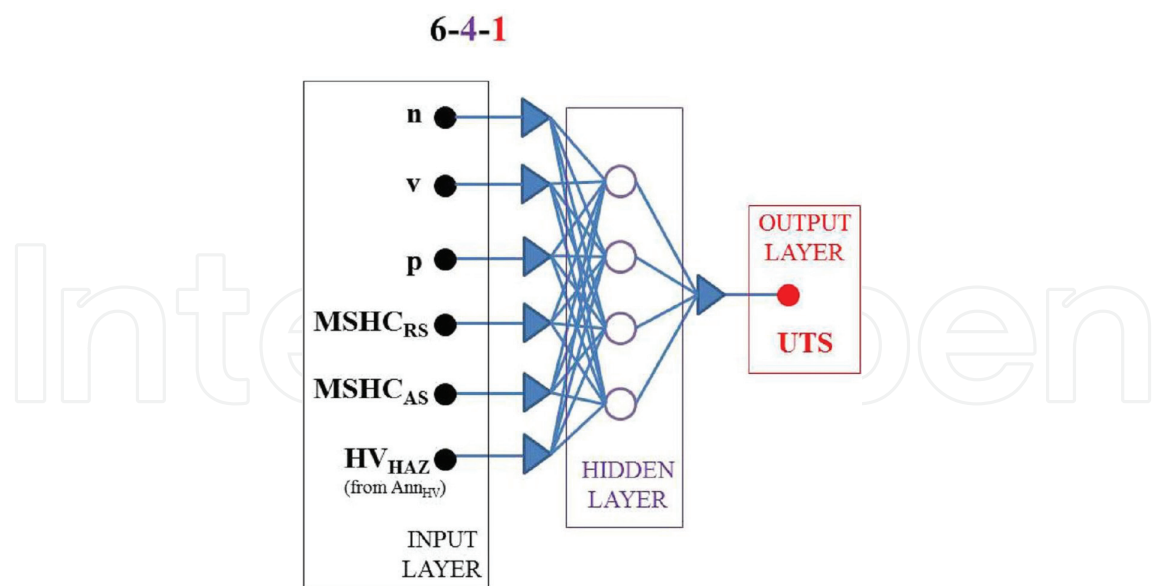
Figure 10. Back-propagation neural network used to foresee the ultimate tensile strength.
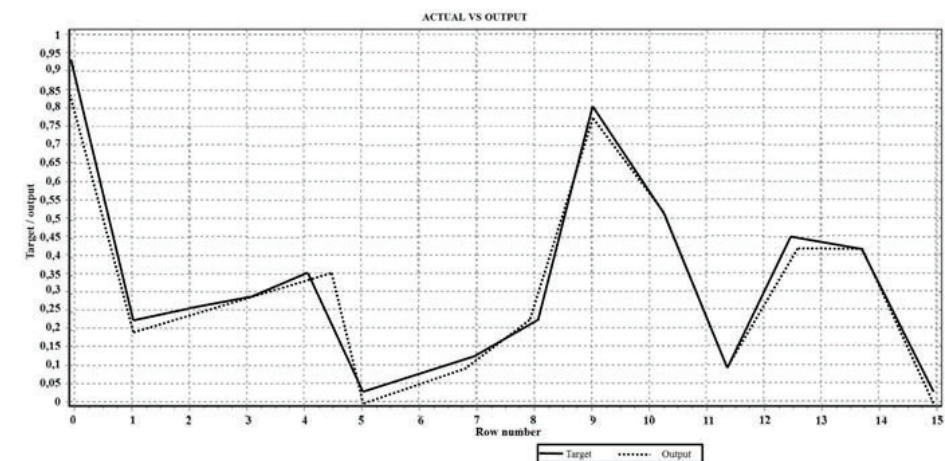

Figure 11. Predicted ultimate tensile strength by ANN versus the experimental data.
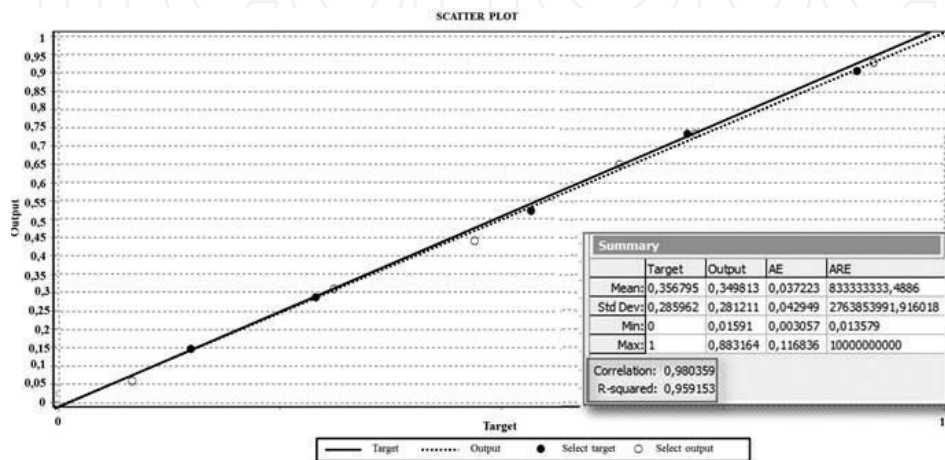

Figure 12. Regression line at the training stage.

ANN model	Input parameters	Output parameters	MAPE (%)
ANN_{HV}	- n - v - p - $MSHC_{AS}$ - $MSHC_{RS}$	HV_{haz} Vickers micro hardness of the Heat Affected Zone	0.29
ANN_{UTS} (in cascade)	- n - v - p - $MSHC_{AS}$ - $MSHC_{RS}$ - HV_{haz} (predicted with the ANN_{HV} model)	UTS ultimate tensile strength	9.57

AA5754 H111 plates welded with the friction stir welding process.

Table 3. Mean absolute percentage error computed for hardness and ultimate tensile strength of AA.

5. Conclusions

The interest in ANN is growing so much that its models and algorithms are becoming standard tools in computer science and information engineering. This highlights the fact that after a long and productive youth, neural networks have formed a robust set of computation procedures with a robust theoretical base and undeniable effectiveness in solving real problems in different fields of information processing. In case study discussed in this work, the analysis performed has shown that the ANN simulation model can be used as a further effective method for predicting the FSW process. The MAPE obtained for the outputs micro hardness (HAZ) and ultimate tensile strength (UTS) were, respectively, 0.29% and 9.57%; R2 values were, in all cases, bigger than 0.90. Although the prediction of UTS was characterized by more high level of MAPE, if it is compared to HAZ estimated value, it was considered acceptable to ensure a model characterized by high reliability. The adoption of the simulation model can be very useful for the friction stir welding process. In fact, the use of tools for predicting the mechanical properties of the welds and for controlling the welding process, allows the production of welds with fewer defects. This reduces the number of repairs and costs, associated with the reiteration of the process. In the context of neural networks, the biggest question is "are we currently capable of building a human brain?" [85]. Undoubtedly, the achievement of these challenges is very ambitious, considering that the human brain has around 90 billion neurons shaping an extremely complex network, but in many cases the ANN can support the human behaviour, simplifying the decision making process, increasing the level cognition, under stress conditions, and increasing the capacity in evaluation and analysis of complex processes.

Author details

Luigi Alberto Ciro De Filippis, Livia Maria Serio, Francesco Facchini* and Giovanni Mummolo

*Address all correspondence to: francesco.facchini@poliba.it

Department of Mechanics Mathematics & Management, Polytechnic University of Bari, Italy

References

[1] Cho HS, Leu MC. Artificial neural networks in manufacturing processes. IFAC Proceedings Volumes. 1998;**31**(15):529-537. DOI: S1474-6670(17)40607-0

[2] Huang SH, Zhang H-C. Artificial neural networks in manufacturing: Concepts, applications, and perspectives. IEEE Transactions on Components Packaging and Manufacturing Technology Part A. 1994;**17**(2):212-228. DOI: 10.1109/95.296402

[3] Lilly KW, Melligeri AS. Dynamic Simualtion and Neural Network Compliance Control of an Intelligent Forging Center. Journal of Intelligent and Robotic Systems. 1996;**17**:8199

[4] Kim DJ, Kim BM, Choi C. Determination of the initial billet geometry for a forged product using neural networks. Journal of Material Processing Technology. 1997;**72**:86-93

[5] Lee WH, Cho HS. Design of a CMAC controller for hydro-forming process. Journal of Institute of Control, Robotics and Systems. 2000;**6**(3):329-337

[6] Hyun BS, Cho HS. Prediction of fonning pressure curve for hydroforming processes using artificial neural network. Proceedings of the Institution of Mechanical Engineers. 1994;**208**(part I):108-121

[7] Park HJ, Cho HS. A CMAC-based learning controller for pressure tracking control of hydroforming processes. In: ASME Winter Annual Meeting, Dollars, Texas, USA. 1990

[8] Hamdi Demirci H, Coulter JP. Neural network based control of molding processes. Journal of Materials Processing & Manufacturing Science. 1994;**2**:335-353

[9] Lim TG, Cho HS, Chang HS. The use of artificial neural networks in the monitoring of spot weld quality. Journal of Korean Welding Society. 1993;**11**(2):39-53

[10] Kim YS, Kim BM, Cho HS. A robust on-line quality monitoring method for butt welding process under gap size change using artificial neural network. Pacific Conference on Manufacturing. 1996;**2**:639-644

[11] Lim TG, Cho HS. Estimation of weld pool sizes in GMA welding process using neural networks. Proceedings of the Institution of Mechanical Engineers. 1993;**207**:15-26

[12] Cho HS. AI applications in welding process monitoring and control. In: Japan-Korea FA Sym, Tokyo, Japan. 1991

[13] Lim TG, Cho HS. Weld pool size control based upon direct estimation of weld pool geometry in GMA welding processes. In: Asian Control Conference, Japan. 1994

[14] Ulrich D, Jens H. Using weld Al-methods for parameter geometry determination in scheduling, quality control GMA-welding. ISIJ International. 1999;**39**:1067-1074

[15] Mirapeix J, García-Allende PB, Cobo A, Conde OM, López-Higuera JM. Real-time arc-welding defect detection and classification with principal component analysis and artificial neural networks. NDT & E International. 2007;**40**(4):315-323

[16] Anderson K, Cook GE, Karsai G, Ramaswamy K. Artfical neural networks applied to arc welding process modeling and control. IEEE Transactions on Industry Applications. 1990;**26**(5):824-830

[17] Ohshim K, Yamane S, Yabe M, Akita K, Kugai K, Kubota T. Controlling of torch attitude and seam tracking using neuro arc sensor. Proceedings of the International Conference on IECON. 1995;**2**:1185-1189

[18] Quero JM, Millan RL, Franquelo LG. Neural network approach to weld quality monitoring. Proceedings of the International Conference on IECON. 1994;**2**:1287-1291

[19] Matteson MA, Morris RA, Raines D. An optimal artificial neural network for GMAW arc acoustic classification. international trends in welding science and technology, In: Proceedings of the 3rd International Conference on Trends in Welding Rearch, Gatlinburg, Tennessee, USA. 1992. pp. 1031-1035

[20] Li P, Fang MTC, Lucas I. Modelling of submerged arc weld beads using self-adaptive offset neutral networks. Journal of Materials Processing Technology. 1997;**71**:288-298

[21] Ko KW, Kim H, Cho HS, Kong WI. A bead shape classification method using neural network in high frequency electric resistance welding. Journal of KSPE. 1995;**12**(9):86-94

[22] Teshima T, Shibasaka T, Takuma M, Yamamoto A. Estimation of cutting tool life by processing tool image data with neural network. Annals of the CIRP. 1993;**42**:59-62

[23] Monostori L. A step towards intelligent manufacturing: modeling and monitoring of manufacturing processes through artifitial neural networks. Annals of the CIRP. 1993;**42**(I):485-488

[24] Malakooti BB, Zhou YQ, Tandler EC. In-process regressions and adaptive multi-criteria neural networks for monitoring and supervising machining operations. Journal of Intelligent Manufacturing. 1995;**1**(6):53-66

[25] Rangwala S, Dornfeld D. Sensor integration using neural networks for intelligent tool condition monitoring. Journal of Engineering for Industry. 1990;**112**:219-228

[26] Sunil Elanayar VT, Shin YC. Design and implementation of tool wear monitoring with radial basis function neural networks. In: Proceeding of American Control Conference. 1995. pp. 1722-1725

[27] Ali Zilouchian and Oren Masory. Neural network controller for turning operation of a metal cutting process. Proceedings of the American Control Conference. 1995;713-717

[28] Huang S-J, Ching Chiou K. The application of neural networks in self-tuning constant force control. International Journal of Machine Tools and Manufacture. 1996;**36**(I):17-31

[29] Chryssolouris G, Domroese M, Beaulieu P. Sensor synthesis for control of manufacturing processes. Journal of Engineering for Industry. 1992;**114**(I):158-174

[30] Leem CS, Dreyfus SE. A Neural network approach to feature extraction in sensor signal processing for manufacturing automation. Journal of Intelligent Material Systems and Structures. 1994;**5**:247-257

[31] Leem CS, Dornald DA, Dreyfus SE. A customized neural network for sensor fusion in on-line monitoring of cutting tool wear. Journal of Engineering for Industry. 1995;**117**:152-159

[32] Chao PY, Hwang YD. An improved neural network model for the prediction of cutting tool life. Journal of Intelligent Manufacturing. 1997;**8**:107-115

[33] Balazinski M, Czogala E, Sadowski T. Control of metal-cutting process using neural fuzzy controller. In: 2nd IEEE International Conference Fuzzy Systems. 1993. pp. 161-166

[34] Venkatesh K, Zhou M, Caudill R. Design of artificial neural networks for tool wear monitoring. Journal of Intelligent Manufacturing. 1997;**8**:215-226

[35] Mou I. A method of using neural networks and inverse kinematics for machine tools error estimation and correction. Journal of Manufacturing Science and Engineering. 1997;**119**:247-254

[36] Grabec I, Kuljanic E. Characterization of manufacturing processes based upon acoustic emission analysis by neural networks. Annals of the CIRP. 1994;**43**:77-80

[37] Govekar E, Grabec I. Self-qrganizing neural network application to drill wear classification. Journal of Engineering for Industry. 1994;**116**:233-238

[38] Tamg YS, Li TC, Chen MC. On-line drilling chatter recognition and avoidance using an ART2-A neural network. International Journal of Machine Tools and Manufacture. 1994;**34**(7):949-957

[39] Lin SC, Ting CI. Drill wear monitoring using neural networks. International Journal of Machine Tools and Manufacture. 1996;**36**(4):465-475

[40] Ko TI, Cho DW. Cutting state monitoring in milling by a neural network. International Journal of Machine Tools and Manufacture. 1994;**34**(5):659-676

[41] Tamg YS, Hseih YW, Hwang ST. Sensing tool breakage in face milling with a neural network. International Journal of Machine Tools and Manufacture. 1994;**34**(3):341-350

[42] Ko TI, Cho DW, Jung MY. On-line monitoring of tool breakage in face milling using a self-organized neural network. Journal of Manufacturing Systems. 1995;**42**(2):80-90

[43] Tansel IN, Mekdeci C, McLaughlin C. Detection of Tool Failure in End Milling with Wavelet Transformations and Neural Networks. International Journal of Machine Tools and Manufacture. 1995;**35**(8):1137-1147

[44] Liu H, Chen T, Qui L. Predicting grinding bum using artificial neural networks. Journal of Intelligent Manufacturing. 1997;**8**:235-237

[45] Kao IY, Tang YS. A neural-network approach for the on-line monitoring of the electrical discharge machining process. Journal of Materials Processing Technology. 1997;**69**:112-119

[46] Himmel CD, Kim TS, Krauss A, Kamen EW, May GS. Real-time predictive control of semiconductor manufacturing processes using neural networks. In: Proceedings of the American Control Conference, 1995. pp. 1240-1244

[47] Bushman S, Edgar TF, Trachtenberg I. Modeling of plasma etch systems using ordinary least squares, recurrent neural network, and projection to latent structure models. Journal of the electrochemical society. 1997;**144**(4):1379-1389

[48] Rietman EA. A neural network model of a contact plasma etch process for VLSI production. IEEE Transactions on Semiconductor Manufacturing. 1996;**9**(1):95-100

[49] Nami Z, Misman O, Erbil A, May GS. Semi-empirical neural network modeling of metal-organic chemical vapor deposition. IEEE Transactions on Semiconductor Manufacturing. 1997;**10**(2):288-294

[50] Rietman EA, Lory ER. Use of neural networks in modeling semiconductor manufacturing processs: An example for plasma etch modeling. IEEE Transactions on Semiconductor Manufacturing. 1993;**6**(4):343-347

[51] Baker MD, Himmel CD, May GS. In-situ prediction of reactive ion etch endpoint using neural networks. IEEE Transactional on Components, Packaging and Manufacturing Technology-Part A. 1995;**18**(3):478-483

[52] Wang Q, Sun X, Golden BL, DeSilets L, Edward AWH, Luco S, Peck A. A neural network model for the wire bonding process. Computers & Operations Research. 1993;**20**(8):879-888

[53] Moon YB, Chi SC. Generalized part family formation using neural network techniques. Journal of Manufacturing Systems. 1992;**1**(3):149-159

[54] Kaparthi S, Suresh NC. Machine-component cell formation in group technology: A neural network approach. International Journal of Production Research. 1992;**30**(6):1353-1367

[55] Lee H, Malave CO, Ramachandran S. A selforganizing neural network approach for the design of cellular manufacturing systems. Journal of Intelligent Manufactuting. 1992;**3**:325-332

[56] Moon YB. Establishment of a neurocomputing model for part family/machine group identification. Joumal of Intelligent Manufacturing. 1992;**3**:173-182

[57] Wu B. An introduction to neural networks and their applications in manufacturing. Journal of Intelligent Manufacturing. 1992;**3**:391-403

[58] Cook DF, Shannon RE. A predictive neural network modelling system for manufacturing process parameters. International Journal of Production Research. 1992;**30**(7):1537-1550

[59] Yih Y, Liang T-P, Moskowitz H. Robot scheduling in a circuit board production line: A hybrid OR/ANN approach. ZZE Transactions. 1993;**25**(2):26-33

[60] Chryssolouris G, Lee M, Domroese M. The use of neural networks in determining operational policies for manufacturing systems. Journal of Manufacturing Systems. 1991;**10**(2):166-175

[61] Facchini F, Mossa G, Mummolo G. A model based on artificial neural network for risk assessment to polycyclic aromatic hydrocarbons in workplace. In: 25th European Modeling and Simulation Symposium. 2013. pp. 282-289

[62] Verner MTR. Artificial neural networks in business: Two decades of research. Applied Soft Computing. 2016;**38**:788-804

[63] Brain Mapping by Integrated Neurotechnologies for Disease Studies [Internet]. Available from: http://brainminds.jp/en/

[64] Available from: https://www.science.org.au/publications/inspiring-smarter-brain research-australia

[65] Brainnetome Project [Internet]. Available from: http://www.brainnetome.org/en/brain-netomeproject.html

[66] Prieto A, Prieto B, Ortigosa EM, Ros E, Pelayo F, Ortega J, Rojas I. Neural networks: An overview of early research, current frameworks and new challenges. Neurocomputing. http://dx.doi.org/10.1016/j.neucom.2016.06.014

[67] Casalino G, Facchini F, Mortello M, Mummolo G. ANN modelling to optimize manufacturing processes: The case of laser welding. IFAC-PapersOnLine. 2016;**49**(12):378-383

[68] Jin L, Kuang X, Huang H, Qin Z, Wang Y. Study on the overfitting of the artificial neural network forecasting model. Acta Meteorologica Sinica. 2004;**19**:216-225

[69] Karthick Anand Babu AB. Design and development of artificial neural network based tamil unicode symbols identification system. International Journal of Computer Science Issues (IJCSI). 2012;**9**(1):388

[70] Jordan MI, Rumelhart DE. Forward models: Supervised learning with a distal teacher. Cognitive Science. 1992;**16**(3):307-354

[71] Bousquet O, Raetsch G, von Luxburg U. Advanced Lectures on Machine Learning. Lecture Notes in Artificial Intelligence. Berlin: Springer Verlag; 2004

[72] Sutton RS, Barto AG. Reinforcement Learning. An Introduction. MIT Press. 2014, 2015, 2016 A Bradford Book The MIT Press Cambridge, Massachusetts London, England

[73] Casalino G, Losacco AM, Arnesano A, Facchini F, Pierangeli M, Bonserio C. Statistical analysis and modelling of an Yb: KGW femtosecond laser micro-drilling process. 10th CIRP Conference on Intelligent Computation in Manufacturing Engineering—CIRP ICME 2016. 2017;**62**:275-280

[74] De Filippis LAC, Serio LM, Facchini F, Mummolo G, Ludovico AD. Prediction of the vickers microhardness and ultimate tensile strength of AA5754 H111 friction stir welding butt joints using artificial neural network. Materials. 2016;**9**(11):915

[75] Ghetiya ND, Patel KM. Prediction of tensile strength in friction stir welded aluminium alloy using artificial neural network. Procedia Technology. 2014;**14**:274-281

[76] Yousif YK, Daws KM, Kazem BI. Prediction of friction stir welding characteristics using neural network. Jordan Journal of Mechanical and Industrial Engineering. 2008;**2**:151-155

[77] Fratini L, Buffa G, Palmeri D. Using a neural network for predicting the average grain size in friction stir welding processes. Computers and Structures. 2009;**87**:1166-1174

[78] Boldsaikhan E, Corwin EM, Logar AM, Arbegast WJ. The use of neural network and discrete Fourier transform for real-time evaluation of friction stir welding. Applied Soft Computing. 2011;**11**:4839-4846

[79] Shojaeefard MH, Akbari M, Tahani M, Farhani F. Sensitivity analysis of the artificial neural network outputs in friction stir lap joining of aluminum to brass. Advances in Materials Science and Engineering. 2013;7

[80] Shojaeefard MH, Behnagh RA, Akbari M, Givi MKB, Farhani F. Modelling and Pareto optimization of mechanical properties of friction stir welded AA7075/AA5083 butt joints using neural network and particle swarm algorithm. Materials and Design. 2013;**44**:190-198

[81] Asadi P, Givi MKB, Rastgoo A, Akbari M, Zakeri V, Rasouli S. Predicting the grain size and hardness of AZ91/SiC nanocomposite by artificial neural networks. International Journal of Advanced Manufacturing Technology. 2012;**63**:1095-1107

[82] Serio LM, Palumbo D, Galietti U, De Filippis LAC, Ludovico AD. Analisi del processo di Friction Stir Welding applicato alla lega AA5754-H111: Comportamento meccanico e termico dei giunti. Riv. Ital. Saldatura. 2014;**66**:509-524

[83] Serio LM. Control, Monitoring and Optimization of the Friction Stir Welding Process [PhD thesis]. Bari, Italy: Polytechnic of Bari; 2015

[84] Serio LM, Palumbo D, De Filippis LAC, Galietti U, Ludovico AD. Effect of friction stir process parameters on the mechanical and thermal behavior of 5754-H111 aluminum plates. Materials. 2016;**9**:122

[85] Costandi M. How to buil a brain July. Sciencefocus.com. 2012:32-38