Artificial Neural Network Structure Optimisation in the Pareto Approach on the Example of Stress Prediction in the Disk-Drum Structure of an Axial Compressor

Adam Kozakiewicz and Rafał Kieszek *

Faculty of Mechatronics, Armament and Aerospace, Institute of Aviation Technology, Military University of Technology, 00-908 Warsaw, Poland; adam.kozakiewicz@wat.edu.pl
* Correspondence: rafal.kieszek@wat.edu.pl

Abstract: The article presents the process of selecting and optimising artificial neural networks based on the example of determining the stress distribution in a disk-drum structure compressor stage of an aircraft turbine engine. The presented algorithm allows the determination of von Mises stress values which can be part of the penalty function for further mass optimization of the structure. A method of a parametric model description of a compressor stage is presented in order to prepare a reduced stress distribution for training artificial neural networks. A comparative analysis of selected neural network training algorithms combined with the optimisation of their structure is presented. A genetic algorithm was used to determine the optimal number of hidden layers and neurons in a layer. The objective function was to minimise the absolute value of the relative error and standard deviation of stresses determined by FEM and artificial neural networks. The results are presented in the form of the Pareto front due to the stochastic optimisation process.

Keywords: artificial neural networks; optimization; genetic algorithm; turbine engines; axial compressors; predictive model

1. Introduction

Artificial neural networks have been known to artificial intelligence researchers for over 70 years. W. McCulloch and W. Pitts developed the first formal description of the artificial neuron model in 1943, presented in paper [1]. The mathematical model of the McCulloch–Pitt neuron [1] is described in terms of the Equation (1):

\[ y_i = f(u_i) = f \left( \sum_{j=1}^{N} w_{ij} x_j(t) + w_{i0} \right) \]  

where:
- \( u_i \) —argument of the summation function;
- \( w_{ij} \) —value of the selected weighting factor;
- \( x_j(t) \) —input signal to the neuron;
- \( w_{i0} \) —value of bias.

As shown in the work of [2], the sigmoidal threshold function gives the best results described by the Equation (2):

\[ f(u_i) = \frac{1}{1 + \exp(-\beta u_i)} \]  

where:
- \( \beta \) —steepness coefficient of the function.
In Equation (1), the weights $w_{ij}$ can take any value. The case when $w_{ij} > 0$ means excitatory connection, when $w_{ij} < 0$ means suppressive connection, and when $w_{ij} = 0$ means no synaptic connection. The adoption of a discrete neuron model is justified by the phenomenon of refraction. This means that a biological neuron can only operate at a certain frequency.

In his book [3], D. Hebb discovered that the McCulloch–Pitt neuron can store information in its structure using weight values. In addition, he proposed a way to teach the network by manipulating the weight values. He derived his idea of teaching neurons from observations of nature. He observed that the weight of connections between nerve cells is amplified when they are activated at the same time. That is, the more often a connection is used, the stronger it becomes. Mathematically, Hebb’s model is represented as follows:

$$w_{ij}(k + 1) = w_{ij}(k) + \eta y_i(k)y_j(k)$$

(3)

where:

- $k$—cycle number;
- $\eta$—learning rate;
- $y_i, y_j$—output value from the $i$-th and $j$-th neuron.

In recent years, the possibility of combining Artificial Neural Networks (ANNs) with hard algorithms for numerical calculations is being studied. The aim is to reduce the time of numerical calculations, which is directly related to the design costs. An analysis of selected works presenting the problem of applying neural networks for optimisation using hard algorithms is presented below [4–7].

ANNs are increasingly being used in aviation [8–10]. They are used in reliability analysis [11,12], fault detection [13–16], identification [17], control [18,19], and design and optimisation [20–25].

Pierret and Van den Braembussche, in their work [26], presented the results of compressor palisades and turbines optimisation, using an algorithm containing an ANN. The authors were able to reduce the pressure loss coefficient by about 4% in relation to the output system. The algorithm proposed by the authors can be successfully used to optimise turbine or compressor blades, for both profiles alone and in 3D systems. The algorithm excludes the analysis of unreal blades, and the use of the ANN significantly increases the speed of calculations, which is a major advantage. Moreover, the applied algorithm made it possible to fully automate the optimisation process, ensuring that the algorithm does not indicate a local minimum, very distant from the global one.

In an article by Keshtegar [27], an ANN was used to predict stresses, displacements, and strains in turbine engines. The authors analysed selected parameters during a given mission in a BLISK turbine disc.

In the work [28], similarly to the works [29,30], the reliability of the turbine rotor in terms of creep was assessed for the clearances. The authors decomposed the rotor model into component models (i.e., disc, blades, and body), which were analysed using FEM. The data were used to train a neural network, which was able to achieve an RMSE accuracy of $6.32 \times 10^{-4}$ for the worst matching disc component. As a result, an overall solution accuracy of ~99.93% was achieved for 104 simulations. The authors believe that the demonstrated algorithm presents a promising approach in the field of dynamic reliability analysis of complex structures.

In the publication [31], an innovative method of temperature field predicting in a three-dimensional model of a turbine rotor rim was presented. For this purpose, an algorithm combining the FVM (Finite Volume Method) and an ANN was used. Using the ANSYS program, the temperature distribution was determined, which serves as input data for ANN training. This was able to predict the temperature distribution for selected nodes of the turbine model with the change of rotor rotational speed, temperature, and inlet pressure. The maximum absolute temperature error was 44.7 K, and mean absolute error (MAPE) was 0.5%. A similar problem was solved in paper [31] where the main goal was
minimisation time of temperature prediction. Elapsed time using the CFD method is one h, 42 min, and 44.7 s, and elapsed time using the ANN method was 0.0027645 s.

Stress analysis using an ANN in wind turbine blades is the subject of papers [32] and [33]. The first study uses Genetic Algorithm–Back Propagation (GA-BP) and the second Particle Swarm Optimization–Back Propagation (PSO-BP). In both cases, a relative error of about 6% (on the leading edge, on the overpressure side) was achieved compared to full-scale static tests. For the trailing edge, the relative error was 6.5% and 18% for the PSO-BP and GA-BP algorithms, respectively. There are also other examples of ANN use available in the literature, e.g., in wind turbine optimisation [34,35]. There are also works questioning the validity of approximating the objective function with an ANN [36], which is a counterexample to the previously cited works.

In the analysed area of the literature, the authors did not come across any publications related to the use of an ANN in disk systems or disk-drum turbine engines optimisation. Previous attempts of the authors to use an ANN in compressor disc calculations are presented in publications [9,10]. This paper is an extension of these works, aimed at optimising and extending the research of the previously used algorithm.

The article presents the results of applying an ANN to reduce the time of determining the penalty function value in the case of the structural system strength analysis of the compressor stage. The constructed model was based on a disk-drum system of a turbojet engine compressor. The stresses necessary to train the ANN were obtained using the APDL language. Additionally, the network structure was optimised due to the relative error and standard deviation of stress prediction, and the solutions were presented in the form of the Pareto front. In the classic Pareto approach, two objective functions are optimised, which are, for example, related to weight values [37], or a solution is sought in the solution space using, for example, the cone separation method [38]. In the article, the optimal solutions front was determined by the simultaneous minimisation of two objective functions using the stochastic method, and therefore, the Pareto front was outlined on the basis of many stochastically optimal solutions.

2. Use of the APDL Language to Determine the Disk Stresses

The greatest advantage of APDL (Ansys Parametric Design Language) is its adaptation to parametric modelling, which provides a set of solutions. Using the APDL language, it is also possible to discretise the model, set parametric boundary-initial conditions and perform numerical analyses. It is also possible to generate paths along which, regardless of the model variant, the selected results of numerical analysis are interpreted. Due to these advantages, it was decided to use the APDL language. The aim of the work was to create a parametric model of the disk-drum compressor structure, an aircraft turbine engine, whose single stage is shown in Figure 1. The parametric model of the stage is shown in Figure 2.

![Figure 1. Optimised compressor disc: (a) real object, (b) CAD model.](image-url)
Figure 2. Parametric model of the disc with indication of the main dimensions. Colours indicate the paths along which the stresses for optimisation were determined.

The model shown in Figure 2 was constructed to ensure the tangency of the curves (G1 connection class) at the points of the system element connections, with the exception of the hub-drum connection, where the G0 connection class was used, i.e., no derivatives continuity. The model used is axisymmetric to the vertical axis (Figure 3), as the APDL language structure requires this.

Figure 3. Discretised disc model with marked boundary conditions, where: A and C—symmetry condition zone, B—rim pull zone.

The discretisation of the system was carried out using the “AMESH” function with a set mesh size of 0.005.
Based on Equation (4) [39] in the polar coordinate system, summing the forces in the radial direction, including body force $R$ per unit volume in radial direction and dividing by $dr \, d\Theta$, we obtain the equation for equilibrium:

$$\frac{(\sigma_r)_1 - (\sigma_r)_3}{d\Theta} - \frac{1}{2}[(\sigma_\Theta)_2 + (\sigma_\Theta)_4] + \frac{(\tau_{r\Theta})_2 - (\tau_{r\Theta})_4}{d\Theta} + Rr = 0$$

(4)

where:
- $\sigma_r$—radial stress component;
- $\sigma_\Theta$—circumferential stress component;
- $\tau_{r\Theta}$—shearing stress components;
- $R$—body force per unit volume in radial direction;
- $r$—radius.

Index 1, 3 is normal to the plate, and 2, 4 is radial direction.

For axisymmetric stress distribution, stress components do not depend on $\Theta$. After a few simple transformations, the end of the disc (Figure 3, B) is loaded by pulling according to the Equation (5) [40–42]. This equation binds centrifugal force from blade and lock mass and centrifugal (radial) stress on the outside surface of the disc.

$$\sigma_w = \frac{(F_1 + F_2) \, n}{2\pi \, r_{\text{max}} \, h_{\text{min}}}$$

(5)

where:
- $\sigma_w$—disc centrifugal stress;
- $F_1$—centrifugal force from the lock;
- $F_2$—centrifugal force from the blade;
- $n$—number of blades;
- $r_{\text{max}}$—outer radius of the disc;
- $h_{\text{min}}$—thickness of the disc.

The pull was performed using the SFL (Specifies Surface Loads) function. The rotation speed of the disc system is set using the OMEGA function. The symmetry condition is set at the point where the drum is connected to the next stage drum (in Figure 2, zones A and B). This approach makes it possible to decompose the analysis of the entire compressor rotor (Figure 4) into individual stages. It was assumed that the connection of the stages is sufficiently rigid, and the symmetry condition in this area does not introduce significant errors.

![Figure 4. Compressor optimisation rotor.](image-url)
Solving the numerical problem, three paths were generated in the post-processor along which the data necessary for the calculations were interpreted. Each path is defined by two points. The first path is based on the centre of the hub base and the centre of the disc tip (Figure 2—blue line); the second runs through the centre of the left drum section (Figure 2—red line), and the third through the centre of the right drum section (Figure 2—green line). This made it possible to obtain the stress distribution calculated in accordance with the hypothesis of the maximum energy of shear deformation (Huber–Mises–Hencky theory) in the sections necessary for the profiling of the structure. An example of the stress pattern for the first path is shown in Figure 5.

![Stress distribution along the first path.](image)

The disc section path was divided into twenty nodes and the drum section path into five. The created numerical model was validated using the stochastically selected geometric parameters of the model and loads. The results of the numerical analysis were compared with the analytical calculations bringing the model geometry closer to the straight disc. This convergence is achieved by reducing the thickness of the hub to 5.1 mm at the disc thickness of 5 mm, reducing the drum thickness to 0.5 mm and the $R_i$ radii of curvature to 5 mm. It can be seen that the obtained model corresponds to the straight disc arrangement. The stress distribution for the analytical model ($\sigma_{AN}$) and APDL ($\sigma_{APDL}$) and relative error ($\delta$) for the obtained models are presented in Figure 6.

For $r = 0.4$ m, the greatest error occurred due to the presence of an additional structural element in this area (due to the disk–drum connection). For the remainder, the error does not exceed 0.6%.

In a further part of the paper, the APDL algorithm is used to generate the input training parameters described in Figure 2: $r_{\text{max}}$—radius of disk, $r_{\text{min}}$—radius of central hole, $h_0$—hub width, $h_{\text{min}}$—disc width, $\text{l}_{\text{bl}}$—length of left drum, $\text{r}_{\text{bl}}$—radius of left drum, $\text{l}_{\text{br}}$—length of left drum, $\text{r}_{\text{br}}$—radius of left drum, and $\text{rb}$—radius of location of the centre of the circle round the disc-hub joint. Additionally, there are a few parameters connected with material properties and blade geometry such as: Young module, density, Poisson’s ratio, mass of single blade, number of blades, radius of the centre of mass of the lock, radius of the centre of mass of the blade, volume of lock, and rotational speed. The Von Mises stresses along the normalized length of the three paths (Figure 2, blue, red, and green lines) were calculated with the APDL algorithm and used as target training data. The disc part path (blue) was divided into twenty-one nodes, and the drums paths (green and red) were divided into six nodes.
Solving the numerical problem, three paths were generated in the post-processor along which the data necessary for the calculations were interpreted. Each path is defined by two points. The first path is based on the centre of the hub base and the centre of the disc tip (Figure 2—blue line); the second runs through the centre of the left drum section (Figure 2—red line), and the third through the centre of the right drum section (Figure 2—green line). This made it possible to obtain the stress distribution calculated in accordance with the hypothesis of the maximum energy of shear deformation (Huber–Mises–Hencky theory) in the sections necessary for the profiling of the structure. An example of the stress pattern for the first path is shown in Figure 5.

Figure 5. Stress distribution along the first path.

The disc section path was divided into twenty nodes and the drum section path into five. The created numerical model was validated using the stochastically selected geometric parameters of the model and loads. The results of the numerical analysis were compared with the analytical calculations bringing the model geometry closer to the straight disc. This convergence is achieved by reducing the thickness of the hub to 5.1 mm at the disc thickness of 5 mm, reducing the drum thickness to 0.5 mm and the $R_0$ radii of curvature to 5 mm. It can be seen that the obtained model corresponds to the straight disc arrangement. The stress distribution for the analytical model ($\sigma_{ANA}$) and APDL ($\sigma_{APDL}$) and relative error ($\delta$) for the obtained models are presented in Figure 6.

![Figure 6. Results of program validation in APDL to determine the stress distribution in the disc.](image)

3. Selection of the Training Method

The selection of the training method resulted from the literature review. It indicated that ANN is a tool capable of supporting or replacing classical numerical calculations. As a result of the preliminary numerical analysis presented in [10], the authors obtained knowledge in the field of optimal network selection. It provided knowledge on the network structure used. NARX (nonlinear autoregressive neural network with external input) and FeedForward neural networks with extended structure were tested. The NARX network was unable to solve the problem of profiling compressor disc and was therefore excluded from further analysis. All analysed neural networks had biases and sigmoidal activation function, except for the last layer, which included single neurons. The constructed network does not fully meet the assumptions of stress prediction accuracy in the optimisation task. Therefore, in the next stage of the work, numerical studies were carried out in order to improve the selection of the optimal ANN structure and training method. A comparative analysis of the selected methods was carried out. The following were compared:

- Scaled Conjugate Gradient (SCG) method [43];
- Momentum (GDM) method [44];
- Resilient back propagation algorithm (RPROP) [45].

The network was trained from one to four times. The data used to train the ANN and compare the results were the stresses that were generated based on the algorithm in the APDL language. The network was trained until reaching the fiftieth iteration of the training algorithm without improving the weight values.

Network calculations were performed using one to four hidden layers and 25 or 50 neurons in each layer. A total of 150 results were obtained from each network, for which the absolute value of the average and the largest relative error between the stresses determined with the ANN and the APDL algorithm were determined. The results of the analyses are presented in Figures 7–9, where the “out” data are the ANN training database, and the “TEST” database is the newly drawn data, previously unknown to the network.

Conclusions from the research on selection of the network training:

- Almost all networks achieve higher errors for networks with 50 neurons in the layer;
- Multiple training of the network has the greatest impact on the quality of the RPROP algorithm results;
- The RPROP algorithm obtains relatively large discrepancies in the calculation accuracy depending on the number of hidden layers, and especially depending on their parity;
- Errors for the first iteration of the training process for networks with one hidden layer for the SCG and RPROP algorithm are similar;
- SCG and GDM algorithms have a rapid error drop for four-layer networks;
- The GDM algorithm has low sensitivity to the number of hidden layers (except for the four-layer network) and the repetition of the training process.

![Graph](image.png)

Figure 7. Maximum and average relative error for selected structures of neural networks and training algorithms after the first training cycle.

![Graph](image.png)

Figure 8. Maximum and average relative error for selected structures of neural networks and training algorithms after the second training cycle.

The above conclusions provided information that no algorithm stands out in terms of quality. In further research, the use of any of the algorithms was not ruled out. Therefore, an attempt was made to minimise the error through optimisation of the ANN structure. As a result of the analysis, there was no need to repeat the network training process.
Figure 9. Maximum and average relative error for selected structures of neural networks and training algorithms after the third training cycle.

4. Selection of an ANN Structure

In order to determine the optimal ANN structure, two-criteria optimisation was performed with the use of a genetic algorithm. The influence of the number of hidden layers and neurons in a layer on the absolute value of the mean relative error and standard deviation was studied. Optimisation for three training algorithms (RPROP, SCG, and GDM) was performed, as tested and described in the previous chapter. The genetic algorithm performed calculations for 20 generations, simultaneously optimising 50 neural networks with a given number of hidden layers and neurons in each layer were studied. Two cases of training data size were also compared: ±5% and ±10% relative to the variable value in relation to the reference disc. These variables are described in the second paragraph. The results of the optimisation process are presented below. These are the stresses calculated in the seventh node of the first path, calculated using the ANN and the APDL algorithm. For the remaining nodes, a similar correlation to the results obtained for the seventh node was observed. The performed optimisation produced Pareto fronts. The results are presented:

- For the GDM algorithm, in Figures 10 and 11;
- For the SCG algorithm, in Figures 12–15;
- For the RPROP algorithm, in Figures 16–19.

Figure 10. Pareto fronts for all networks trained using the GDM algorithm.
For the GDM algorithm, in Figures 10 and 11;

For the SCG algorithm, in Figures 12–15;

For the RPROP algorithm, in Figures 16–19.

Figure 10. Pareto fronts for all networks trained using the GDM algorithm.

Figure 11. Pareto fronts for all networks trained using the GDM algorithm, in the zone of highest accuracy.

Based on the error distribution and standard deviation, it was found that the GDM algorithm is unable to perform the task regardless of the network structure. The sensitivity of the algorithm was observed only with respect to the span of the database. Figure 11 clearly shows two bands corresponding to the database span of 10% and 5%. During optimisation, a lack of sensitivity to the training accuracy was noticed in relation to the number of hidden layers and the number of neurons in the layer.

Figure 12. Pareto fronts for all networks trained using the SCG algorithm.

Figure 13. Pareto fronts representing the neural networks with the highest accuracy, trained using the SCG algorithm.

Figure 14. Pareto fronts presented on a logarithmic scale for all networks trained using the SCG algorithm.

Figure 15. Pareto fronts for the two best-matched networks trained using the SCG algorithm.

Figure 16. Pareto front for all studied networks trained using the RPROP algorithm.
The SCG algorithm.

Figure 16. Pareto front for all studied networks trained using the RPROP algorithm.

Figure 17. Pareto fronts representing the neural networks with the highest accuracy, trained using the RPROP algorithm.

Figure 18. Pareto fronts presented on a logarithmic scale for all studied networks trained using the RPROP algorithm.
Based on the error distribution and standard deviation, it was found that the GDM algorithm is unable to perform the task regardless of the network structure. The sensitivity of the algorithm was observed only with respect to the span of the database. Figure 11 clearly shows two bands corresponding to the database span of 10% and 5%. During optimisation, a lack of sensitivity to the training accuracy was noticed in relation to the number of hidden layers and the number of neurons in the layer.

Figures 12–19 present the Pareto fronts for selected neural networks for two variables (absolute value of the average relative error $\delta$ and its standard deviation $\sigma$). As a result of the research, it was noticed that the increase in the number of hidden layers reduces the span impact of the training database on the accuracy of the results. However, the more layers the network has, the lower the accuracy. This may be caused by an insufficiently large sample of learning data; however, for a single-layer network, the learning accuracy is satisfactorily high. For both training methods (SCG and RPROP), the Pareto front points for a neural network with one hidden layer at the data span have the smallest relative error values. In the case of the RPROP algorithm for networks with one and two hidden layers, for data $\pm 5\%$, the Pareto fronts are arranged parallel to each other, creating bands of similar width, with a smaller average error of the network with a single hidden layer. For the SCG algorithm, the front band for a single-layer network looks similar to a network trained with the RPROP algorithm, but the band for a two-layer network is much wider (Figures 12 and 16). Its centre is closer to the single-layer network band, and its minimum values correspond to its values. The reasons for the change in the width and position of the front can be sought in the different number of neurons in the layer, as shown in Tables 1 and 2. The number of neurons in a given network are average values, rounded to integers, obtained in the optimisation process.

| Number of Hidden Layers | 1   | 2   | 3   |
|-------------------------|-----|-----|-----|
| Database span           | 5%  | 10% | 5%  | 10% | 5%  | 10% |
| Average number of neurons in a layer | SCG | 2   | 3   | 5   | 8   | 4   | 7   | 23  | 27  | 8   | 21  | 19  | 5   |
|                         | RPROP | 3   | 3   | 12  | 6   | 16  | 6   | 17  | 22  | 9   | 33  | 23  | 6   |

Table 2 summarises the averaged absolute values of the relative error and its standard deviation for the selected training algorithms and the number of neurons in the layer. The results were obtained for a network with one hidden layer, as this network has the lowest error rate compared to networks with more layers.
Table 2. Absolute value of the mean relative error $\delta$ and its standard deviation $\sigma$ for selected training algorithms and the number of neurons in one hidden layer.

| Number of Neurons | SCG $\delta$ [-] | SCG $\sigma$ [Pa] | RPROP $\delta$ [-] | RPROP $\sigma$ [Pa] |
|-------------------|-------------------|-------------------|-------------------|-------------------|
| 2                 | $1.30 \times 10^{-4}$ | $1.30 \times 10^{-4}$ | $3.07 \times 10^{-4}$ | $2.53 \times 10^{-2}$ |
| 3                 | $1.79 \times 10^{-4}$ | $1.79 \times 10^{-4}$ | $5.34 \times 10^{-4}$ | $5.34 \times 10^{-4}$ |
| 4                 | $3.30 \times 10^{-4}$ | $3.30 \times 10^{-4}$ | -                | -                |

The RPROP algorithm for one hidden layer, excluding the network with two neurons, matches the training database worse than the SCG algorithm. The network trained with the SCG algorithm with two neurons in the hidden layer has the smallest relative error, and the network with two hidden neurons also trained with the SCG algorithm has the smallest standard deviation. Through the Pareto front, distinguished for a particular number of neurons in a layer by the combination of $\delta$ and $\sigma$, on the basis of the trend line, a qualitatively average ANN match to the selected database can be determined, as shown in Figures 20 and 21.

![Figure 20](image-url). Pareto front depending on the number of neurons in the hidden layer with trend lines marked. Network trained with the SCG method.

The determination coefficient ($R^2$) for the SCG and RPROP algorithm with one hidden layer and from 2 to 4 neurons for the database ±5% was also compared. Figures 22–24 present the training accuracy of selected algorithms for the best-matched neural networks. It is, respectively, $R^2 = 0.9978$ and $R^2 = 0.9725$ for the SCG (NN 20-2-1) and RPROP (NN 20-3-1) algorithms. The accuracy of the optimal ANN structure was calculated on the basis of a newly drawn dataset which was not used in the network training process.
The determination coefficient ($R^2$) for the SCG and RPROP algorithm with one hidden layer and from 2 to 4 neurons for the database $\pm 5\%$ was also compared. Figures 22–24 present the training accuracy of selected algorithms for the best-matched neural networks. It is, respectively, $R^2 = 0.9978$ and $R^2 = 0.9725$ for the SCG (NN 20-2-1) and RPROP (NN 20-3-1) algorithms. The accuracy of the optimal ANN structure was calculated on the basis of a newly drawn dataset which was not used in the network training process.

Figure 21. Pareto front depending on the number of neurons in the hidden layer with trend lines marked. Network trained with the RPROP method.

Figure 22. Matching the training data for the SCG algorithm (NN 20-2-1) on the left and RPROP algorithm (NN 20-3-1) on the right.
Figure 23. Comparison of data matching used for training and validation for the SCG algorithm (NN 20-2-1).

Figure 24. Comparison of data matching used for training and validation for the RPROP algorithm (NN 20-3-1).
Based on the above analysis, according to the Pareto approach, a neural network with one hidden layer and two neurons was selected, trained with the SCG algorithm using the database span of ±5% as optimal for a given task. The networks of three and four neurons were rejected because they have few points significantly deviating from the mean, which increases the standard deviation.

The presented algorithm allows the determination of von Mises stress values which can be part of the penalty function for further mass optimization of the structure. The model makes it possible to determine the stress distribution in a drum–disc structure if the mass-geometric properties of the blades are known. The biggest advantage of the algorithm is its short running time, at the cost of an error within the error limit of the FEM method. The algorithm has been tested on the example of a stress-distribution analysis of a real structure of a 14-stage jet engine compressor (Figure 4) and had an error comparable to error for a single stage of disc-drum structure (Figure 1). The algorithm is not able to determine the stress distribution in a design where the disc is not symmetrical or where the disc is connected directly to the shaft. Moreover, with variable disc, hub, or drum thickness, the algorithm fails. An additional limitation is that the effect of temperature in the structure is neglected. However, it is possible to easily modify the APDL algorithm and add one more input parameter to the ANN to include the effect of temperature on the stress distribution.

5. Conclusions

The analysis of the literature and the results of calculations showed that it is possible to use artificial neural networks to solve scientific problems in the area of rotating rotor machine elements, such as an aircraft turbine engine. Although papers [23,46,47] deal with the problems of implementing Artificial Neural Networks to solve engineering problems, the analyses are limited to flows or to the problem of frictional contact with elements loaded with uniform forces. There are no attempts to use SSNs for strength calculations of rotating machine elements. In our paper, we applied ANN to predict stress distribution due to reduce time of analysis and potentially optimization. Three ANN training algorithms (GDM, SCG, and RPROP) were tested. The analysis of the results showed that the algorithm with the momentum method (GDM) had the worst match to both the training and test data. Repeating the network training process did not significantly affect the qualitative results. The calculations of the compressor unit showed that the conjugate gradient (SCG) algorithm and the resilient backpropagation (RPROP) algorithm have the highest and comparable accuracy of calculations.

The concept of using the combination of APDL and the MATLAB language as a stress-based penalty function for compressor unit optimisation has been presented. The optimisation of network structures with the use of a genetic algorithm enabled the determination of the Pareto fronts for optimal network structures. These fronts were grouped depending on the training method and the number of neurons in the layer. This made it possible, according to the Pareto approach, to select a network with an optimal structure and training method. Thanks to the optimisation of the network structure using a genetic algorithm, it is possible to obtain a high regression coefficient with a small number of neurons and hidden layers in artificial neural networks.

The combination of multi-criteria optimisation in the Pareto approach used in the study for the prediction of stresses in order to minimise the mass of the aircraft engine subassembly structure can be considered as intentional. Pareto multi-criteria optimisation is a useful tool to determine the optimal structure of the Artificial Neural Network, while minimising both its mean error and standard deviation.

Further research should concentrate on improving the learning method of Artificial Neural networks with newer algorithms. For example, in papers [48,49] based on the Lyapunov–Krasovskii functional, reciprocally convex technique, Jensen’s inequality reduced the communication load in the network. An event-triggering scheme filtering has
been proposed for delayed neural networks with sampled data and some suitable sufficient conditions of the considered neural networks were achieved.

**Author Contributions:** Conceptualization A.K. and R.K.; methodology, R.K.; software, R.K.; validation, A.K. and R.K.; formal analysis, A.K. and R.K.; data curation, A.K. and R.K.; writing—original draft preparation and R.K.; writing—review and editing, A.K.; visualization, R.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Military University of Technology, Warsaw, Poland, as part of the projects UGB 781/2022.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the study’s design; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

**References**

1. McCulloch, W.S.; Pitts, W.H. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* **1943**, *5*, 115–133. [CrossRef]

2. Karlik, B.; Olgac, A.V. Performance analysis of various activation functions in generalized MLP architectures of neural networks. *Int. J. Artif. Intell. Expert Syst.* **2011**, *1*, 111–122. Available online: [https://www.csejournals.org/manuscript/Journals/IJAE/Volume1/Issue4/IJAE-26.pdf](https://www.csejournals.org/manuscript/Journals/IJAE/Volume1/Issue4/IJAE-26.pdf) (accessed on 7 May 2022).

3. Hebb, D.O. *The Organization of Behavior: A Neuropsychological Theory*; Psychology Press: London, UK, 2005.

4. Baturynska, I.; Semeniuta, O.; Martinsen, K. Optimization of process parameters for powder bed fusion additive manufacturing by combination of machine learning and finite element method: A conceptual framework. *Procedia CIRP* **2018**, *67*, 227–232. [CrossRef]

5. Karban, P.; Pánek, D.; Orosz, T.; Petrášová, I.; Doležel, I. FEM based robust design optimization with Agros and Artap. *Comput. Math. Appl.* **2021**, *81*, 618–633. [CrossRef]

6. Jeong, K.; Lee, H.; Kwon, O.M.; Jung, J.; Kwon, D.; Han, H.N. Prediction of uniaxial tensile flow using finite element-based indentation and optimized artificial neural networks. *Mater. Des.* **2020**, *196*, 109104. [CrossRef]

7. Seventekidis, P.; Giagopoulos, D.; Arailopoulos, A.; Markogiannaki, O. Structural Health Monitoring using deep learning with optimal finite element model generated data. *Mech. Syst. Signal Process.* **2020**, *145*, 106972. [CrossRef]

8. Maheshwari, A.; Davendra Lingam, N.; DeLaurentis, D.A. A comparative study of machine learning techniques for aviation applications. In *Proceedings of the 2018 Aviation Technology, Integration, and Operations Conference*, Atlanta, GA, USA, 25–29 June 2018; p. 3980. [CrossRef]

9. Kozakiewicz, A.; Kieszek, R. Application of artificial neural networks for stress calculations of rotors. *Przegląd Mechan. LXXVII* **2018**, *1*, 28–30. [CrossRef]

10. Kieszek, R.; Kozakiewicz, A.; Rogólski, R. Optimization of a Jet Engine Compressor Disc with Application of Artificial Neural Networks for Calculations Related to Time and Mass. *Adv. Sci. Technol. Res. J.* **2021**, *15*, 208–218. [CrossRef]

11. Kumar, A.; Srivastava, A.; Banerjee, A.; Goel, A. Performance based anomaly detection analysis of a gas turbine engine by artificial neural network approach. In *Proceedings of the Annual Conference of the PHM Society*, Minneapolis, MN, USA, 23–27 September 2012; Volume 4. [CrossRef]

12. DePold, H.R.; Gass, F.D. The application of expert systems and neural networks to gas turbine prognostics and diagnostics. *J. Eng. Gas Turbines Power* **1999**, *121*, 607–612. [CrossRef]

13. Nascimento, R.G.; Viana, F.A. Cumulative damage modeling with recurrent neural networks. *AIAA J.* **2020**, *58*, 5459–5471. [CrossRef]

14. Nyulászi, L.; Andoga, R.; Butka, P.; Főző, L.; Kovacs, R.; Moravec, T. Fault detection and isolation of an aircraft turbojet engine using a multi-sensor network and multiple model approach. *Acta Polytech. Hung.* **2018**, *15*, 189–209. Available online: [http://acta.uni-obuda.hu/Nyulaszi_Andoga_Butka_Fozo_Kovacs_Moravec_81.pdf](http://acta.uni-obuda.hu/Nyulaszi_Andoga_Butka_Fozo_Kovacs_Moravec_81.pdf) (accessed on 29 April 2022).

15. De Giorgi, M.G.; Ficarella, A.; De Carlo, L. Jet engine degradation prognostic using artificial neural networks. *Aerér. Eng. Aerosp. Technol.* **2019**, *92*, 296–303. [CrossRef]

16. Chen, H.S.; Lan, T.S.; Lai, Y.M. Prediction Model of Working Hours of Cooling Turbine of Jet Engine with Back-propagation Neural Network. *Sens. Mater.* **2021**, *33*, 843–858. [CrossRef]

17. Morinaga, M.; Morig, J.; Matsu, T.; Kawase, Y.; Hanaka, K. Identification of jet aircraft model based on frequency characteristics of noise by convolutional neural network. *Acoust. Sci. Technol.* **2019**, *40*, 391–398. [CrossRef]
18. De Giorgi, M.G.; Quarta, M. Hybrid multigene genetic programming-artificial neural networks approach for dynamic performance prediction of an aeroengine. *Aerosp. Sci. Technol.* **2020**, *103*, 105902. [CrossRef]
19. Kumarin, A.; Kuznetsov, A.; Makaryants, G. Hardware-in-the-loop neuro-based simulation for testing gas turbine engine control system. In Proceedings of the 2018 Global Fluid Power Society PhD Symposium (GFPS), Samara, Russia, 18–20 July 2018; IEEE: Manhattan, NY, USA; pp. 1–5. [CrossRef]
20. Yu, Y.; Chen, L.; Sun, F.; Wu, C. Neural-network based analysis and prediction of a compressor’s characteristic performance map. *Appl. Energy* **2007**, *84*, 48–55. [CrossRef]
21. Ghorbaniyan, K.; Cholamrezai, M. An artificial neural network approach to compressor performance prediction. *Appl. Energy* **2009**, *86*, 1210–1221. [CrossRef]
22. Cortés, O.; Urquiza, G.; Hernández, J.A. Optimization of operating conditions for compressor performance by means of neural network inverse. *Appl. Energy* **2009**, *86*, 2487–2493. [CrossRef]
23. Grzymkowski, A. Neural model as an alternative to the numerical model of subsonic flow through the turbine palisade. *Mechanik* **2014**, *87*, 217–224. Available online: https://www.inforond.pl/resource/bwmeta1.element.baztech-11e13efa-bc1b-4ac4-831d-9b64a5d1ca1 (accessed on 8 May 2022).
24. Grzymkowski, A.; Szewczuk, N. Modeling of steam flow through subsonic turbine stage with the use of artificial neural networks. *Mechanik* **2016**, *89*, 698–699.
25. Gluch, J.; Butterweck, A. Application of artificial neural networks for thermal-flow diagnostics of rings of steam turbines. *Mechanik* **2014**, *87*, 173–180. Available online: https://www.mechanik.media.pl/pliki/do_pobrania/artykuly/10/173-180.pdf (accessed on 29 April 2022).
26. Pierret, S.; Van den Braembussche, R.A. Turbomachinery blade design using a Navier-Stokes solver and artificial neural network. *J. Turbomach.* **1999**, *121*, 326–332. [CrossRef]
27. Keshtegar, B.; Bagheri, M.; Fei, C.W.; Taylan, O.; Thai, D.K. Multi-extremum-modified response basis model for nonlinear response prediction of dynamic turbine blisk. *Eng. Comput.* **2021**, *38*, 1243–1254. [CrossRef]
28. Zhang, C.Y.; Wei, J.S.; Wang, Z.; Yuan, Z.S.; Fei, C.W.; Lu, C. Creep-based reliability evaluation of turbine blade-tip clearance with novel neural network regression. *Materials* **2019**, *12*, 3552. [CrossRef]
29. Zhang, C.; Wei, J.; Jing, H.; Fei, C.; Tang, W. Reliability-based low fatigue life analysis of turbine blisk with generalized regression extreme neural network method. *Materials* **2019**, *12*, 1545. [CrossRef]
30. Song, L.K.; Bai, G.C.; Fei, C.W. Dynamic surrogate modeling approach for probabilistic creep-fatigue life evaluation of turbine disks. *Aerosp. Sci. Technol.* **2019**, *95*, 105439. [CrossRef]
31. Spodniak, M.; Semrđ, K.; Draganovi, K. Turbine Blade Temperature Field Prediction Using the Numerical Methods. *Appl. Sci.* **2021**, *11*, 2870. [CrossRef]
32. Liu, X.; Liu, Z.; Liang, Z.; Zhu, S.P.; Correia, J.A.; De Jesus, A.M. PSO-BP neural network-based strain prediction of wind turbine blades. *Materials* **2019**, *12*, 1889. [CrossRef]
33. Liu, Z.; Liu, X.; Wang, K.; Liang, Z.; Correia, J.A.; De Jesus, A.M. GA-BP neural network-based strain prediction in full-scale static testing of wind turbine blades. *Energies* **2019**, *12*, 1026. [CrossRef]
34. Sessarego, M.; Feng, J.; Ramos-Garcia, N.; Horcas, S.G. Design optimization of a curved wind turbine blade using neural networks and an aeroelastic vortex method under turbulent inflow. *Renew. Energy* **2020**, *146*, 1524–1535. [CrossRef]
35. Albanesi, A.; Roman, N.; Bre, F.; Fachinotti, V. A metamodel-based optimization approach to reduce the weight of composite laminated wind turbine blades. *Compos. Struct.* **2018**, *194*, 345–356. [CrossRef]
36. Dornberger, R.; Büche, D.; Stoll, P. Multidisciplinary optimization in turbine blade design. In Proceedings of the ECCOMAS 2000, Barcelona, Spain, 11–14 September 2000; Available online: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.18.8046&rep=rep1&type=pdf (accessed on 29 April 2022).
37. Kaliszewski, I. Multi-Criteria Decision Making: Soft Computing for Complex Decision Problems; Wydawnictwa Naukowo-Techniczne: Warszawa, Poland, 2008; Available online: http://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.ojs-issn-2391-761X-year-2008-issue-9-article-129/c/129-110.pdf (accessed on 29 April 2022).
38. Kaliszewski, I. Quantitative Pareto Analysis by Cone Separation Technique; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012; Available online: https://link.springer.com/content/pdf/10.1007/978-1-4615-2772-5_5.pdf (accessed on 29 April 2022).
39. Timoshenko, S.; Goodier, J.N. Theory of Elasticity: By S. Timoshenko and JN Goodier; McGraw-Hill: New York, NY, USA, 1951.
40. Dżygadło, Z.; Łyżwiński, M.; Otyś, J.; Szczeciński, S.; Wiater, R. *Turbine Engine Rotor Assemblies*; WKŁ: Warsaw, Poland, 1982.
41. Thaddeus, W. *Fowler, Jet Engines and Propulsion Systems for Engineers*, Training and Education Development and the University of Cincinnati for Human Resource Development; GE Aircraft Engines: Cincinnati, OH, USA, 1989.
42. Hetnarski, R.B.; Eslami, M.R.; Gladwell, G.M.L. *Thermal Stresses—Advanced Theory and Applications*, Solid Mechanics and Its Applications; Springer: Berlin, Germany, 2009; Volume 158, Available online: https://link.springer.com/content/pdf/10.1007/978-3-030-10436-8.pdf (accessed on 21 June 2022).
43. Mellor, M.F. A scaled conjugate gradient algorithm for fast supervised learning. *Neural Netw.* **1993**, *6*, 525–533. [CrossRef]
44. Hagan, M.T.; Demuth, H.B.; Beale, M.H. *Neural Network Design*; PWS Publishing: Boston, MA, USA, 1996.
45. Riedmiller, M.; Braun, H. A direct adaptive method for faster backpropagation learning: The RPROP algorithm. In Proceedings of the IEEE International Conference on Neural Networks, San Francisco, CA, USA, 28 March–1 April 1993; pp. 586–591. [CrossRef]
46. Głuch, J.; Butterweck, A. Application of artificial neural networks to heat-flow diagnostics of steam turbine bladings. *Mechanik 2014, XVIII, 7.*

47. Grzymkowska, A.; Szewczuk, N. Modelling of steam flow through transonic turbine rows using artificial neural networks. *Mechanik 2016, 7, 698–699.* [CrossRef]

48. Arslan, E.; Vadivel, R.; Ali, M.S.; Arik, S. Event-triggered $H_{\infty}$ filtering for delayed neural networks via sampled-data. *Neural Netw. 2017, 91, 11–21.* [CrossRef] [PubMed]

49. Hymavathi, M.; Muhiuddin, G.; Syed Ali, M.; Al-Amri, J.F.; Gunasekaran, N.; Vadivel, R. Global exponential stability of fractional order complex-valued neural networks with leakage delay and mixed time varying delays. *Fractal Fract. 2022, 6, 140.* [CrossRef]