Politecnico di Milano
Gaia Biasolo, Matteo Giuseppe Cigada, Federico D'Agostini

Lab09: Neural Network

21/11/2025

Abstract

This report investigates the use of a Multi-Layer Feedforward Network (MLFFN) to approximate a known two-variable function. Through sensitivity analysis and a fractional factorial design, the study evaluated the impact of training set size, network topology, and training ratios on model accuracy.

The results indicate that the number of neurons and the number of hidden layers is the most influential parameter for reducing global error, while the training ratio has no significant effect.

Contents

1. Introduction

A **multi-layer feedforward network (MLFFN)** is used to model the behavior of a simple system. The goal of this lab is to go through all the steps that need to be performed when using a neural network:

1. Collection of the training data
2. Definition of the morphology of the neural network
3. Definition of the training procedure
4. Usage of the neural network to predict the behavior of the system with inputs different from the ones of the training data

Successively, the effect of changing the settings of the neural network on the performance has been investigated; this has been performed first by fixing all the parameters while changing one at a time, then, through a fractional factorial design plan.

1.1 Description of the problem

The system that needs to be modelled is a two-variable function for which the analytical expression has been given $Z = x_1 * e^{-x_1^2 - x_2^2}$ (Figure 1)

In a "real" case, the training data are the only available information, and the relationship between the input and the output is unknown. In this preliminary example, however, the exact relationship between input and output is also known. This allows to have a deeper understanding of the effect of changing the neural network settings on the results (the output of the neural network can be compared with the exact solution to verify how well it is approximated).

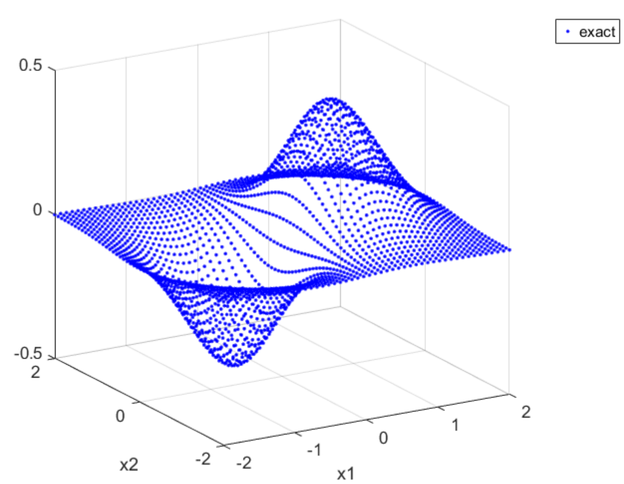

Figure 1

1.2 Basic description of neural network working principle

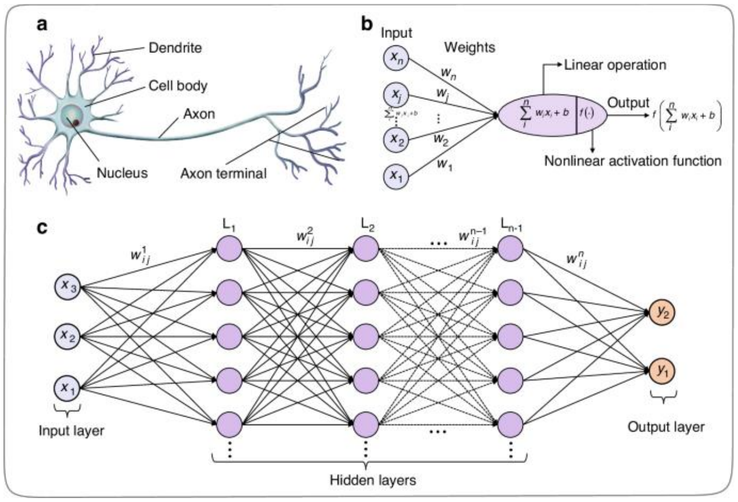

Figure 2

Neural network working principle is based on the human neuron's behavior.

Artificial Neural Networks (ANNs) is an advanced global approximation model, specifically using the Multi-Layer Feedforward Network (MLFFN) architecture where information propagates unidirectionally from input to output layers.

Neuron

Each neuron computes a weighted sum of its input followed by a non-linear activation function, enabling the network to model complex relationships that simpler polynomial regressions cannot capture.

Network

The network is composed of different neurons, organized in subsequent layers. So, in neural networks two important parameters are the **number of layers** and the **number of neurons per layer.**

Training

Through the back-propagation algorithm, the network minimizes the global error by iteratively tuning weights against training data, creating a robust mathematical surrogate for sensitivity analysis.

The global error is also evaluated by means of a specific and different data set, separated from the training set. This allows to reduce the risk of overtraining. A specific parameter for training is the **training ratio**, which is the ratio between the number od data points used for training and the total data points.

2. Sensitivity analysis on neural network parameters

The result shown below should respect the following estimation of the mean square error between estimated network output and a target function.

$$\mathcal{O}\left(\frac{C_f^2}{M}\right) + \mathcal{O}\left(\frac{M * p}{N_t} logN_t\right)$$

Mind that all the results represented are subjected to casual and stochastic distribution of noise. Therefore, minor improvements cannot be considered without further analysis.

2.1 Effect of size of training set

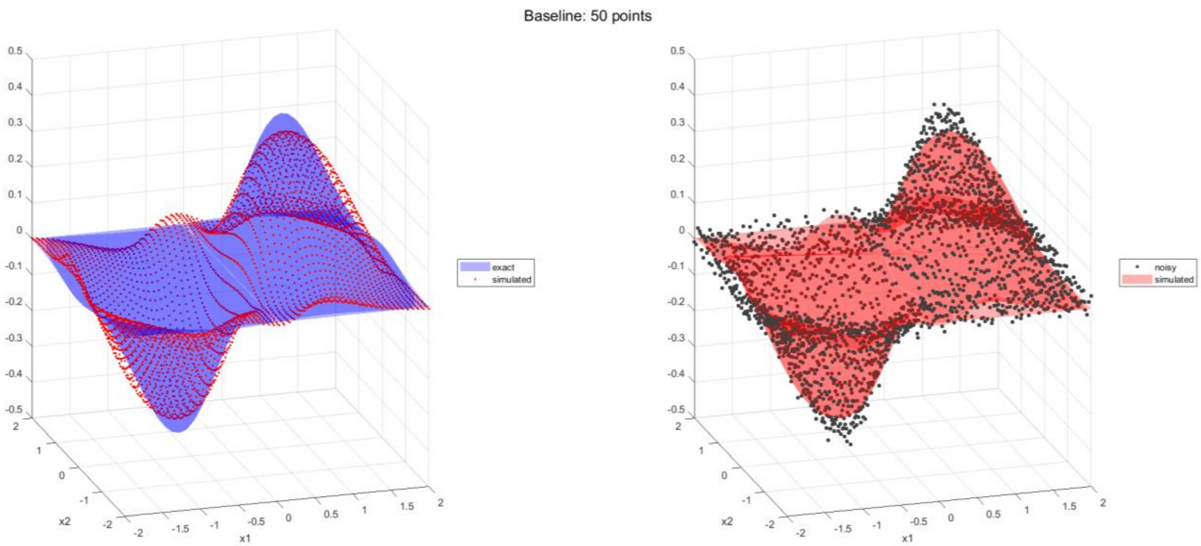

Figure 3

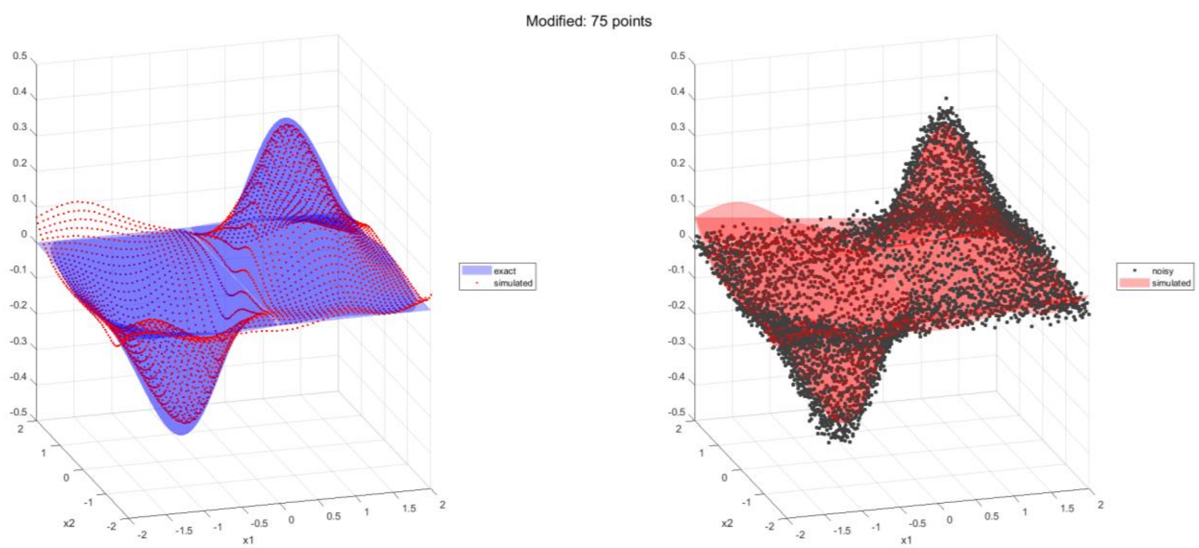

Figure 4

Description of Figure 3 and Figure 4:

- Increasing the population number the approximate result is improved
    - Especially near the peaks
- Overtraining is still visible

2.2 Effect of number of hidden layers

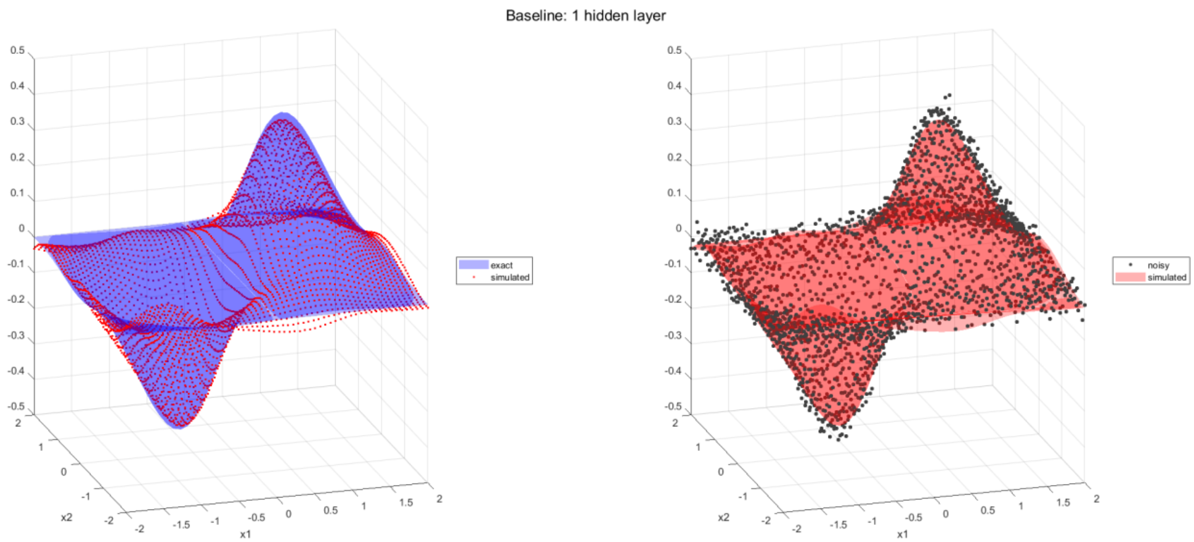

Figure 5

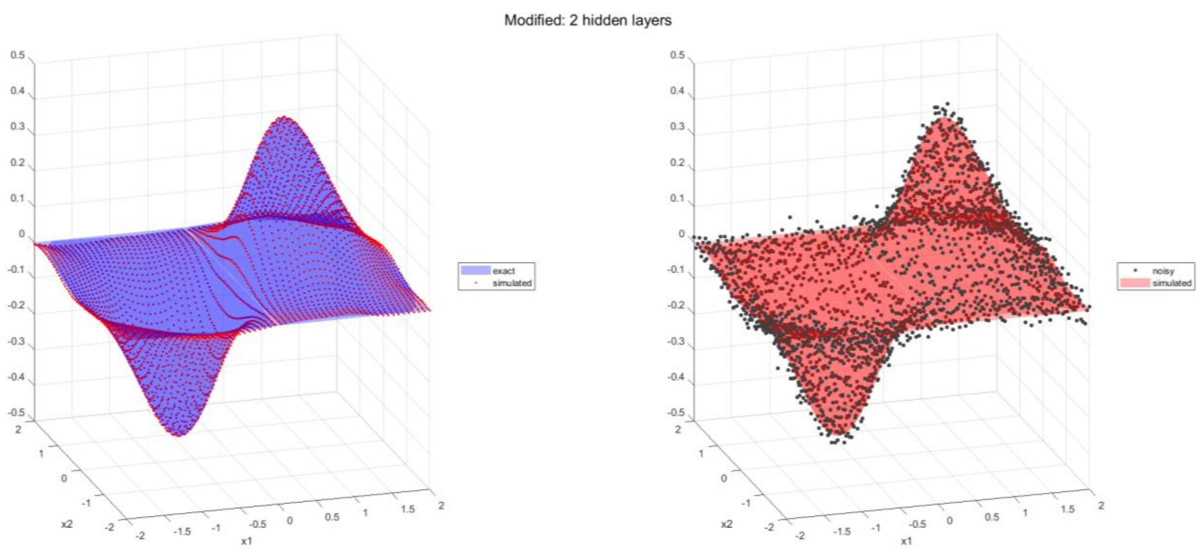

Figure 6

Description of Figure 5 and Figure 6:

- With two hidden layers the results are considerably improved
  - The peaks are very well reproduced

2.3 Effect of number of neurons in each hidden layer

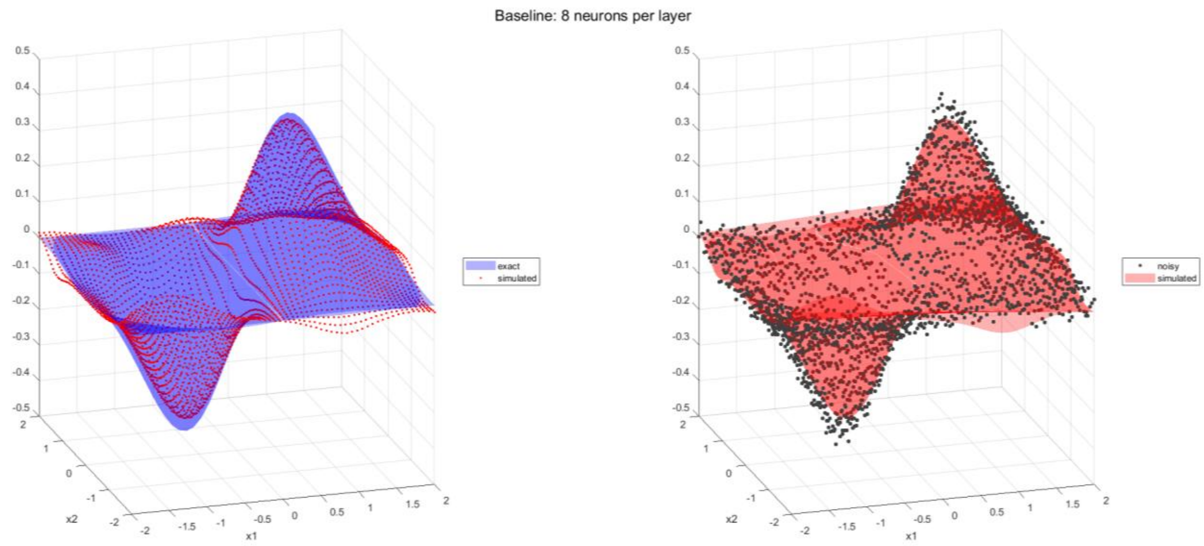

Figure 7

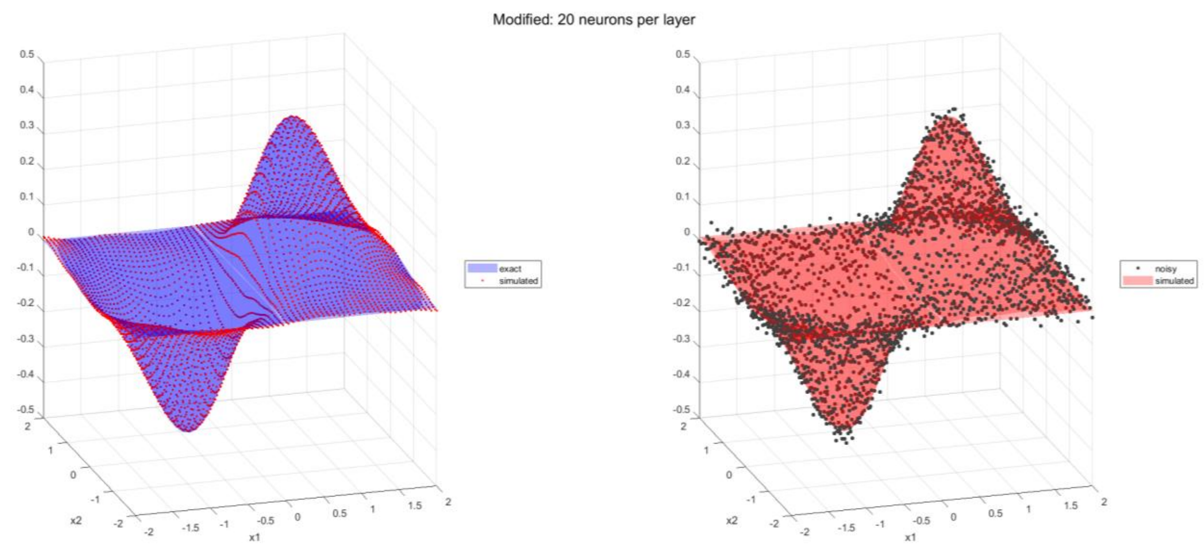

Figure 8

Description of Figure 7 and Figure 8:

- The number of neurons per layer seems to be the most influential factor
    - With more than double the number of initial neurons, the neural network is capable of almost perfectly reproducing the function analyzed

2.4 Effect of training and validation ratio

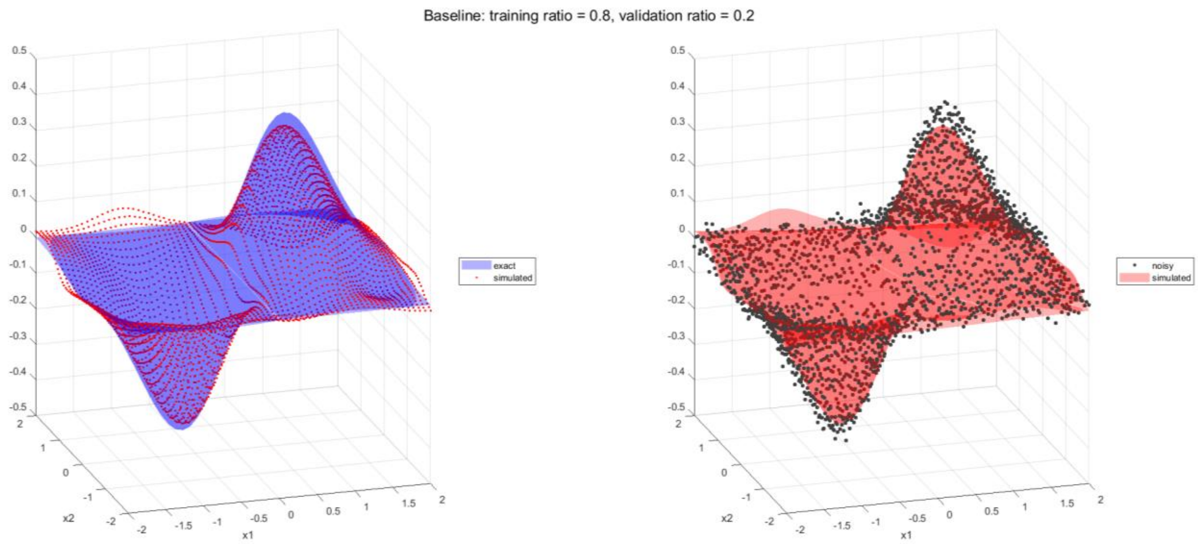

Figure 9

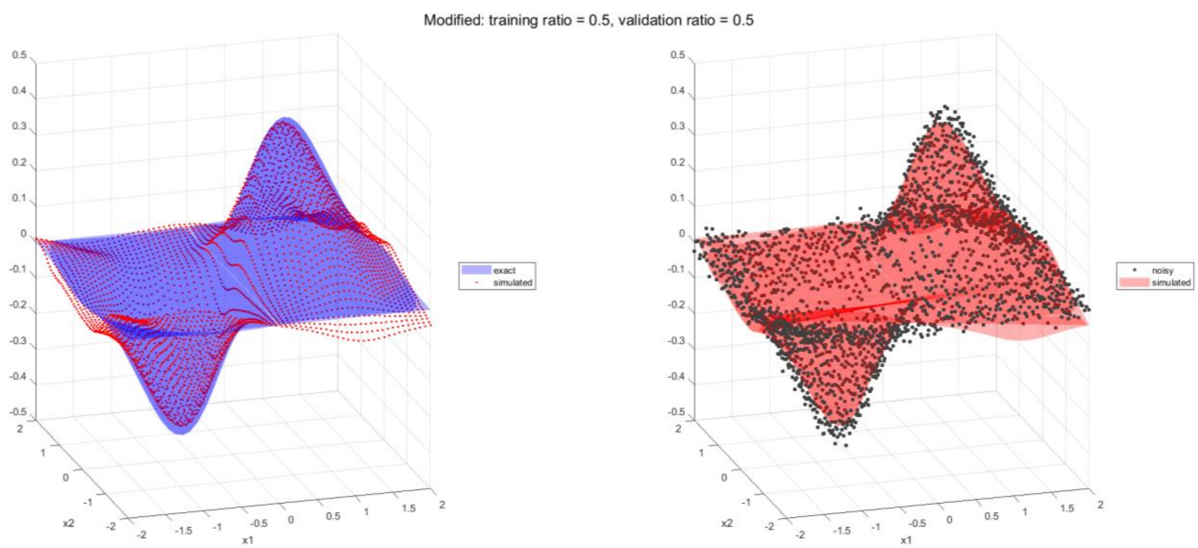

Figure 10

Description of Figure 9 and Figure 10:

- In the lower training ratio, a slight improvement in the results is shown. However, it's not possible to conclude anything, given the stochastic nature of the neural network. The training ratio is not a key factor in neural network training.

2.5 Design of experiments

A design of experiments was carried out to study in deeper detail the combined influence of the different parameters of the neural network on the accuracy of the fitting (specifically, the errors were considered and compared).

In detail, the training set dimension, the number of neurons, the number of layers and the training ratio, were varied on **three levels** and a **fractional factorial** plan (central composite design - faced) was implemented and run. CCD-Faced was used as 3 level design schemes instead of Box-Behnken because the goal is not to test “average” combinations but also “extreme” ones.

Parameters	“Low” Level	“Medium” Level	“High” Level
Training set size	80	160	200
Number of hidden layers	1	2	2
Number of neurons per layer	2	4	8
Training ratio	0.5	0.7	0.9

- The size of the training set was chosen to be compliant with the total number of connections and neurons in the network
- The number of hidden layers was limited to 2 because it is good practice to keep the number of layers low
- The lowest value for training ratio considered is 0.5 because it doesn’t make sense to use more data for the validation than for the actual training

The 36 simulations design plan was run 5 times, and the results were averaged to have consistent results since neural networks work statistically. The results are shown in Figure 11.

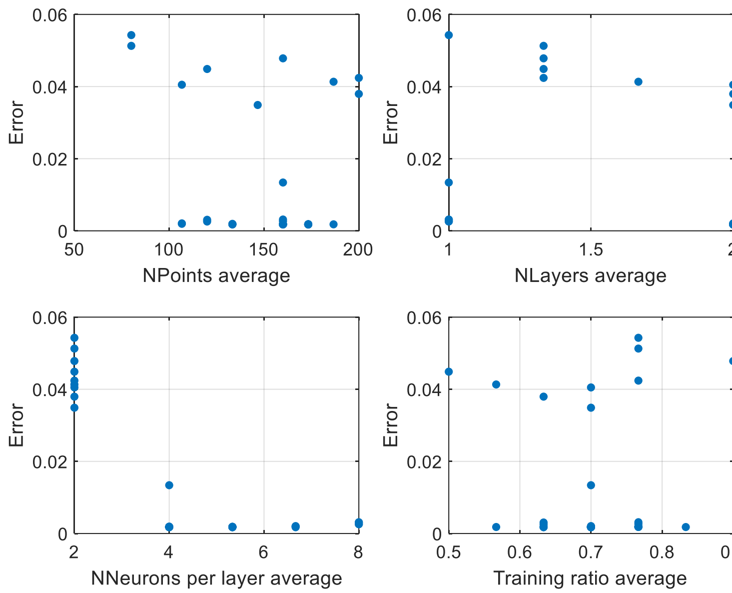

Figure 11

Observing the plots in Figure 11, the general trend shows a decrease in error by increasing the number of layers, the size of the training set and the number of neurons per layer. Instead, varying the training ratio average does not show a clear trend.

To quantify these dependencies, the Spearman correlation coefficient ρ has been computed between the columns of the matrix where all the data have been stored. The results are reported in Figure 12**Error! Reference source not found.**. The output matrix ρ is defined such that $\rho_{i,j}$ is the spearman correlation coefficient between the variables in column i and column j of the input matrix; **the last column** contains the correlation coefficients between the error and the different parameters.

$$\rho = \begin{bmatrix} 1 & 0.2224 & -0.0260 & -0.0321 & \mathbf{-0.3483} \\ 0.2224 & 1 & 0.0160 & -0.1122 & \mathbf{-0.6711} \\ -0.0260 & 0.0160 & 1 & 0.0326 & \mathbf{-0.4133} \\ -0.0321 & -0.1122 & 0.0326 & 1 & \mathbf{-0.1943} \\ -0.3483 & -0.6711 & -0.4133 & -0.1943 & \mathbf{1} \end{bmatrix}$$

Figure 12

3. Conclusion

In this assignment, a feedforward neural network was implemented to fit a two-variable function of which the analytical expression was known. After having generated some training data (analytical function + casual noise), the network was trained several times to study the influence of different parameters, both in the definition of the neural network and for phenomenon of overtraining. In particular, the size of the training set, the number of neurons, the number of hidden layers, and the training ratios were varied, the results were also analyzed by means of a design of experiments. Although the parameter that resulted to be more influential was the number of hidden layers, it should be remembered that the artificial neural network approach depends on statistics, so this could not be generally true for every problem considered.