

IE4497 Homework 2 - Question 6

Full Name: Nguyen Tuan Anh

Matric Number: U2120070G

Question 6: Write and run a Python program to train a 1-N-1 neural network to approximate

$$g(p) = 1 + \sin\left(\frac{6\pi}{4}p\right)$$

. Search for an appropriate N. Show your figures.

Submit your source code (Write your answers to this Homework in a PDF file without your source code. Put your PDF and Jupyter Notebook of this question into a zip file. Submit your zip file to NTULearn).

Answer:

Best number of neurons (N) : 7

Best MSE: 2.076750479321255e-05

The output of source is as below:

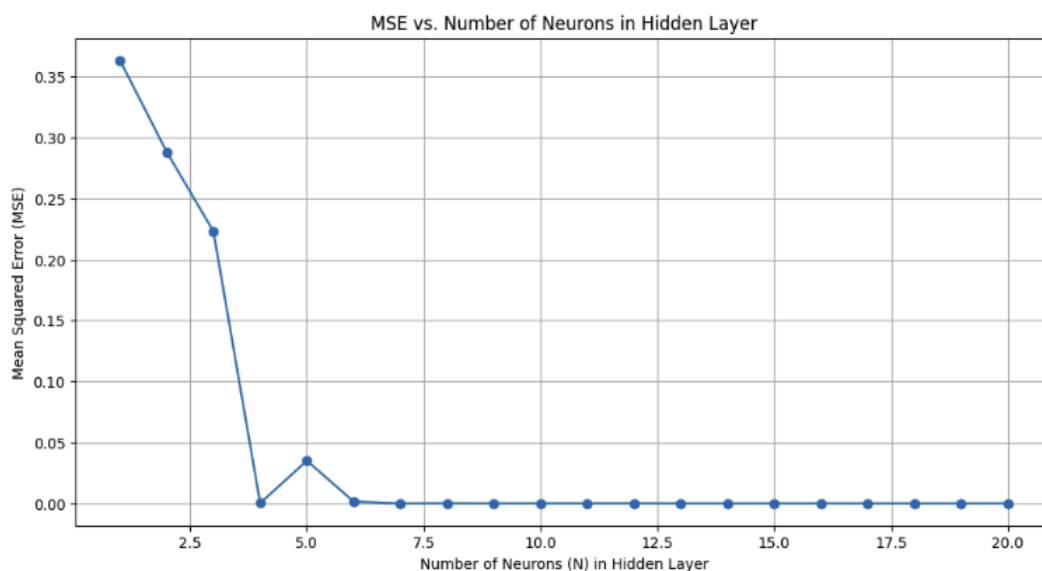


Figure 1: MSE vs. Number of Neurons in Hidden Layer

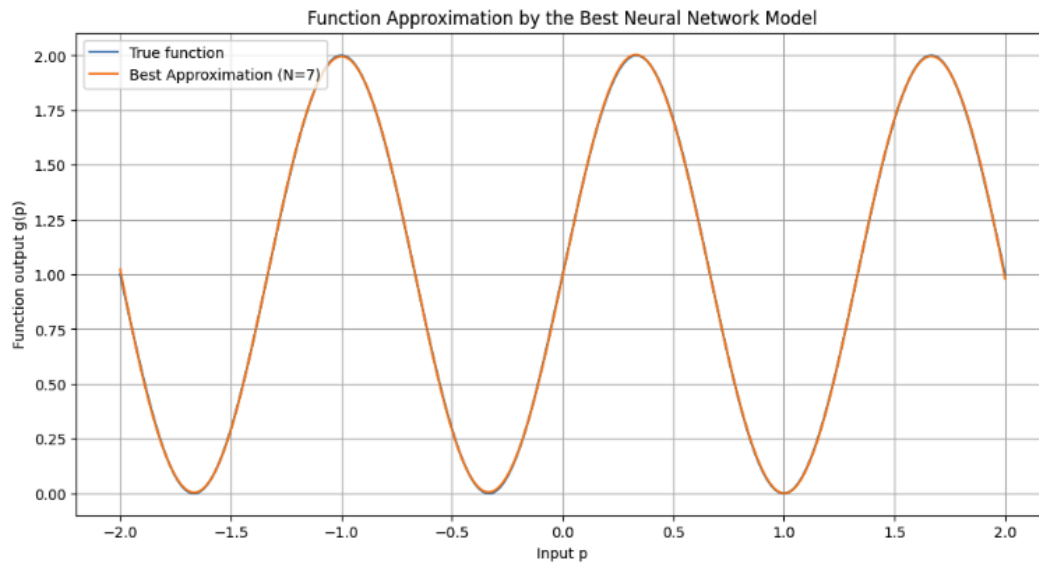


Figure 2: Function Approximation by the Best Neural Network Model

Insights from output submitted:

- Optimal Number of Neurons:

The neural network that yielded the lowest mean squared error (MSE) in approximating the given function had 7 neurons in its hidden layer. This suggests that for this specific problem, a moderately sized hidden layer provides a good balance between complexity and the ability to capture the underlying pattern of the function.

- Low Error:

The best MSE achieved was approximately $2.076750479321255e-05$, which indicates a very close approximation to the true function by the neural network. This low error implies that the network has learned the mapping from input to output with high accuracy.

- Model Complexity vs. Performance:

As the MSE plot would show, initially, as the number of neurons increases, the model's ability to approximate the function improves (reflected by a lower MSE). Beyond the optimal number of neurons, increasing the complexity further might not lead to significant improvements and could potentially lead to overfitting.

- Generalization Capability:

The model has been trained on a subset of the data and tested on unseen data (due to the train-test split). The low MSE on the test set suggests that the model has generalized well rather than memorizing the training data.

- Solver Efficiency:

The '**lbfgs**' solver used here effectively trained the neural network to a high level of precision. This choice seems to be appropriate for the problem's size and complexity.

- Stability of the Solution:

Given that the same result was achieved upon repeated runs (assuming no variability in N), we can infer that the solution is stable and the optimization process is reliably converging to a good solution.

- Visualization as a Diagnostic Tool:

The plots serve as a diagnostic tool to visualize the performance of the model across different complexities (number of neurons) and to see how well the model's predictions match the true function. This visual evidence is very useful when evaluating performance and can be used to make decisions about potential improvements or adjustments.

- Goodness of Fit:

The second plot, which compares the true function with the neural network's approximation, would visually indicate how well the neural network has learned the underlying structure of the data. A close match between the true and approximated function suggests that the network structure (with neurons) is appropriate for modeling this particular function.