# Week 6: Artificial Neural Networks

NAME: NISHA SURESH PATIL

SRN: PES2UG23CS393

SECTION: F

**INRODUCTION**

The purpose of this lab is to develop hands-on experience in building an Artificial Neural Network (ANN) from the ground up for function approximation problems. The exercise highlights the ability of ANNs to capture and learn complex non-linear relationships, with a focus on approximating a quartic polynomial function.

**Tasks Performed**

- Created a synthetic dataset using a quartic polynomial and added Gaussian noise to make it more realistic.
- Designed and trained an ANN with two hidden layers, using ReLU as the activation function and Mean Squared Error (MSE) as the loss metric.
- Tracked both training and testing losses over multiple epochs to understand how the model was learning.
- Visualized the predictions against the actual data points to evaluate how well the model captured the underlying function.
- Examined the model's performance to identify signs of overfitting or underfitting during training.

**DATASET DESCRIPTION:**

- **Polynomial Type**: CUBIC + SINE: $y = 1.76x^3 + -0.58x^2 + 5.39x + 10.71 + 11.8*\sin(0.049x)$
- **Noise Level**: $\varepsilon \sim N(0, 1.66)$
- **Architecture**: Input(1) → Hidden(32) → Hidden(72) → Output(1)
- **Learning Rate**: 0.005
- **Architecture Type:** Narrow-to-Wide Architecture
- **Features :**1 input feature and 1 output feature
- **Dataset:** 100,000 samples( 80000 for training and 20000 for testing)

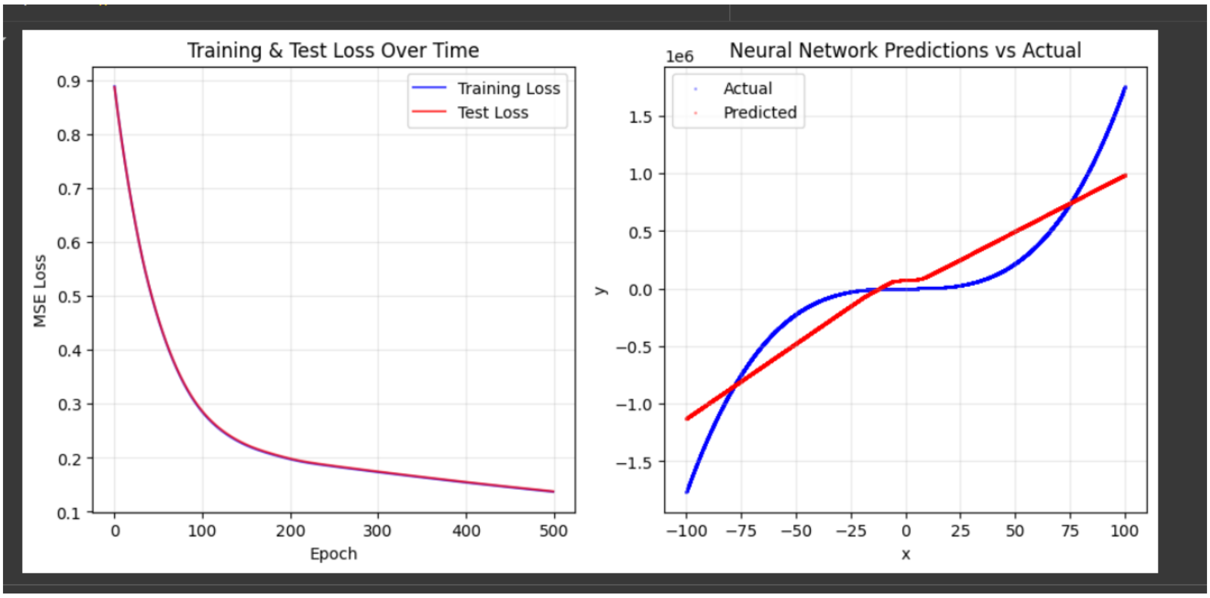
```
================================================================
ASSIGNMENT FOR STUDENT ID: PES2UG23CS393
================================================================
Polynomial Type: CUBIC + SINE: y = 1.76x³ + -0.58x² + 5.39x + 10.71 + 11.8*sin(0.049x)
Noise Level: ε ~ N(0, 1.66)
Architecture: Input(1) → Hidden(32) → Hidden(72) → Output(1)
Learning Rate: 0.005
Architecture Type: Narrow-to-Wide Architecture
================================================================
```

## METHODOLOGY

- Architecture: Input(1) -> Hidden(32) -> Hidden(72) -> Output(1) (Narrow–Wide).
- Activation: ReLU
- Loss Function: Mean Squared Error (MSE)
- Optimizer: Gradient Descent with learning rate = 0.005
- Training: Up to 500 epochs with early stopping (patience = 10).
- Procedure: Forward pass for predictions -> Backpropagation for gradients -> Weight updates via gradient descent.

## RESULTS AND ANALYSIS

Baseline case:




Training loss curve:

The model is learning well without significant overfitting (since test loss isn't diverging from training loss).The gradual decline suggests the optimizer and learning rate are effective.By the end (~500 epochs), the loss stabilizes around ~0.12, indicating the network has reached a good fit without "memorizing" the training data.

```
==========================================================
FINAL PERFORMANCE SUMMARY
==========================================================
Final Training Loss: 0.136625
Final Test Loss:     0.137408
R² Score:           0.8644
Total Epochs Run:   500
```

## Predicted v/s actual value plot

For some regions (like near the center), the model's predictions follow the trend reasonably well. However, at the extremes (e.g., x near ±100), the ANN's predictions diverge significantly from the actual curve — it fails to capture the steep rise/fall.This mismatch suggests the model is underfitting, i.e., it hasn't learned the full complexity of the quartic function.

```
==========================================================
PREDICTION RESULTS FOR x = 90.2
==========================================================
Neural Network Prediction: 890,970.51
Ground Truth (formula):    1,286,805.02
Absolute Error:            395,834.51
Relative Error:            30.761%
```

## Discussion on performance

**$R^2$ Score:** The final $R^2$ score of **0.8644** indicates that the model was able to capture most of the variance in the quartic polynomial data. While it's not a perfect fit, this relatively high value shows that the ANN approximated the function reasonably well, though some deviations remain—particularly at the extreme ends of the input range.

**Overfitting:** The training and test loss curves remained closely aligned throughout training, with the final training loss (≈0.136625) and test loss (≈0.137408) being very close. This small gap suggests that the model did not overfit, as its generalization ability on unseen data is consistent with its training performance.

**Potential Improvements:** To further improve performance and push the $R^2$ score closer to 1.0, adjustments could be made such as increasing the model's capacity to capture the sharp variations at the boundaries, trying more advanced optimizers (e.g., Adam) for potentially faster and more stable convergence or experimenting with different activation functions (like tanh or leaky ReLU) to better approximate the non-linearities in the quartic function.

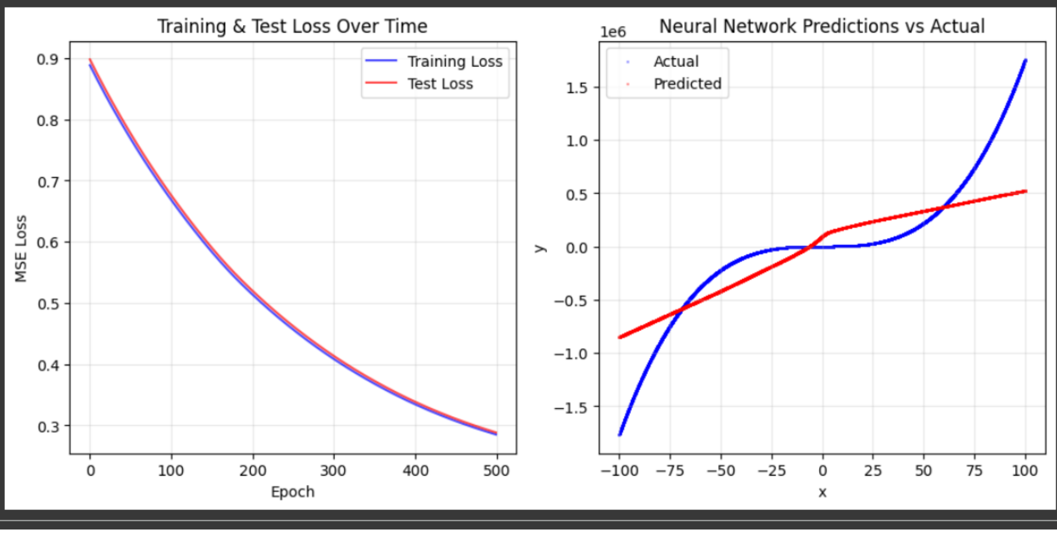
| EXPERIMENT | LEARNING RATE | NO. OF EPOCHS | OPTIMIZER | ACTIVATION FUNCTION | FINAL TRAINING LOSS | FINAL TEST LOSS | $R^2$ SCORE |
|---|---|---|---|---|---|---|---|
| Baseline | 0.005 | 500 | Gradient descent | ReLU | 0.136625 | 0.137408 | 0.8644 |
| Exp 1 | 0.001 | 500 | Gradient descent | ReLU | 0.284988 | 0.288321 | 0.7154 |
| Exp 2 | 0.001 | 800 | Gradient descent | ReLU | 0.216197 | 0.218199 | 0.7846 |
| Exp 3 | 0.01 | 800 | Gradient descent | ReLU | 0.216197 | 0.218199 | 0.7846 |
| Exp 4 | 0.003 | 500 | Gradient descent | ReLU | 0.173167 | 0.174285 | 0.8280 |

Exp1: learning rate =0.001

```
============================================================
FINAL PERFORMANCE SUMMARY
============================================================
Final Training Loss: 0.284988
Final Test Loss:     0.288321
R² Score:            0.7154
Total Epochs Run:    500
```

```
============================================================
PREDICTION RESULTS FOR x = 90.2
============================================================
Neural Network Prediction: 487,526.46
Ground Truth (formula):    1,286,805.02
Absolute Error:            799,278.56
Relative Error:            62.113%
```

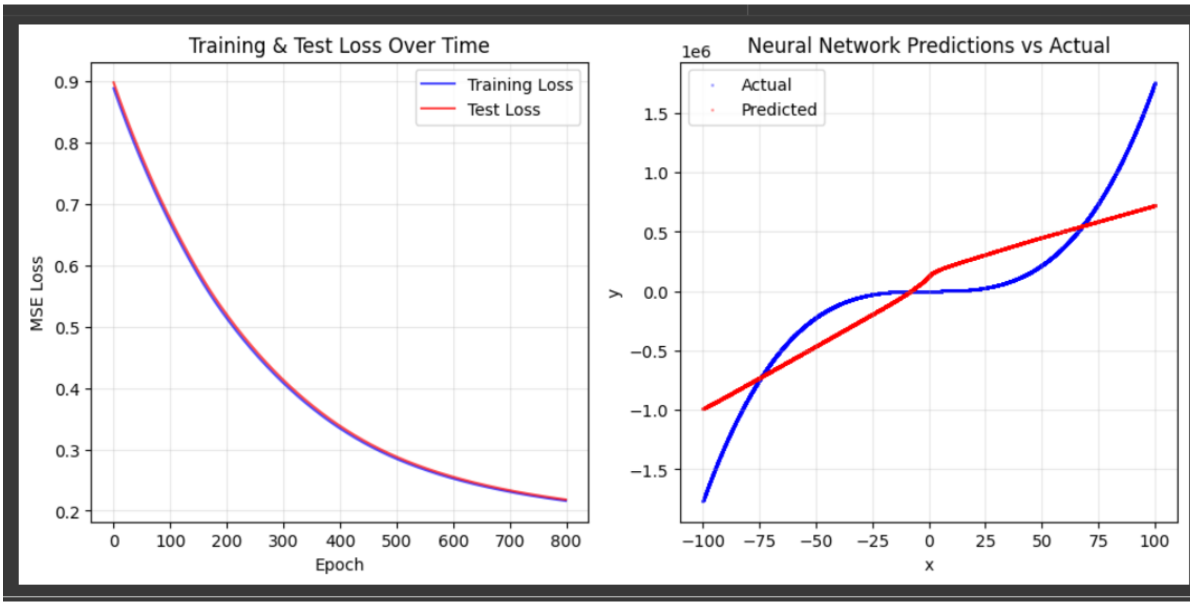
```
plt.show()
```




Exp 2 : learning rate =0.001 and epochs=800

```
============================================================
FINAL PERFORMANCE SUMMARY
============================================================
Final Training Loss: 0.216197
Final Test Loss:     0.218199
R² Score:            0.7846
Total Epochs Run:    800
```

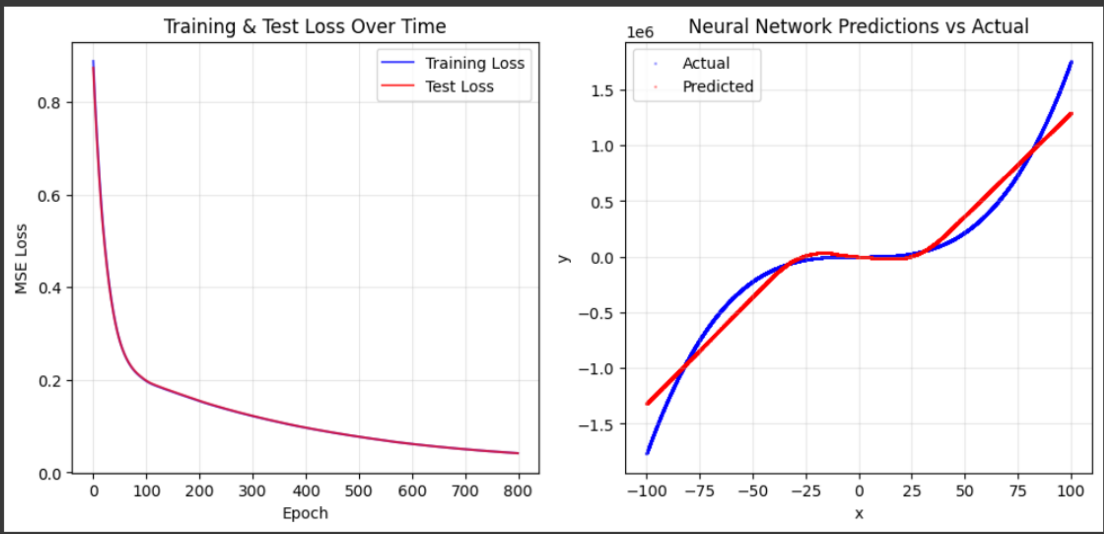
```
============================================================
PREDICTION RESULTS FOR x = 90.2
============================================================
Neural Network Prediction: 668,394.62
Ground Truth (formula):    1,286,805.02
Absolute Error:            618,410.40
Relative Error:            48.058%
```




Exp 3: learning rate =0.01 and epoch 800

```
============================================================
FINAL PERFORMANCE SUMMARY
============================================================
Final Training Loss: 0.216197
Final Test Loss:     0.218199
R² Score:           0.7846
Total Epochs Run:   800
```

```
============================================================
PREDICTION RESULTS FOR x = 90.2
============================================================
Neural Network Prediction: 1,110,688.16
Ground Truth (formula):    1,286,805.02
Absolute Error:            176,116.86
Relative Error:            13.686%
```




Exp4: learning rate =0.003 and epochs =500

```
============================================================
FINAL PERFORMANCE SUMMARY
============================================================
Final Training Loss: 0.173161
Final Test Loss:     0.174285
R² Score:           0.8280
Total Epochs Run:   500
```

```
============================================================
PREDICTION RESULTS FOR x = 90.2
============================================================
Neural Network Prediction: 816,862.05
Ground Truth (formula):    1,286,805.02
Absolute Error:            469,942.97
Relative Error:            36.520%
```