Assignment-1

Q-1 Did the simple 3-neuron network approximate the function well?

Sol:

- The small network with only 3 neurons did an *okay job*, but it wasn't great. It was able to follow the general shape of the function, especially the overall upward trend caused by the "+x" part.
- But because the network is so small, it couldn't really capture the wiggles and oscillations caused by the sine term.

So **yes,** it approximated the function to some extent, but it missed a lot of details.

Q-2 How does the choice of activation (ReLU vs tanh) affect performance?

Sol:

- **ReLU (Rectified Linear Unit):** For negative input values, ReLU gives output = 0. This means the network "ignores" the negative side of the input. In our case, since the function has both positive and negative oscillations (because of the sine part), ReLU cannot model it properly. The result is a poor fit.
- **tanh (Hyperbolic tangent):** This activation gives outputs between -1 and +1, smoothly handling both positive and negative values. This makes it much better for capturing **oscillations and wave-like patterns** in the data.

Q-3 Why might deeper/wider networks perform better on complex functions?

Sol:

A very small network with only 3 neurons is too limited to capture all details. A deeper or wider network has more capacity, so different neurons can learn different parts of the function (like the straight-line growth and the oscillations). When combined, this gives a much closer match to the true function. That is why larger networks are used in real-world applications.