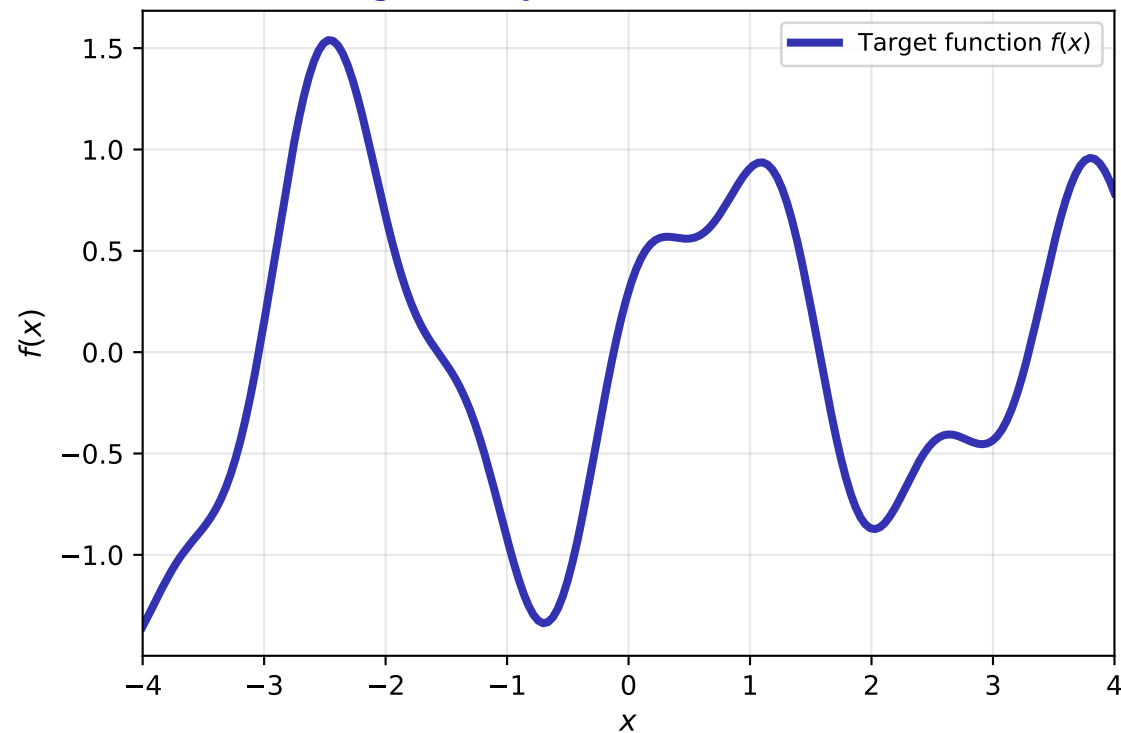
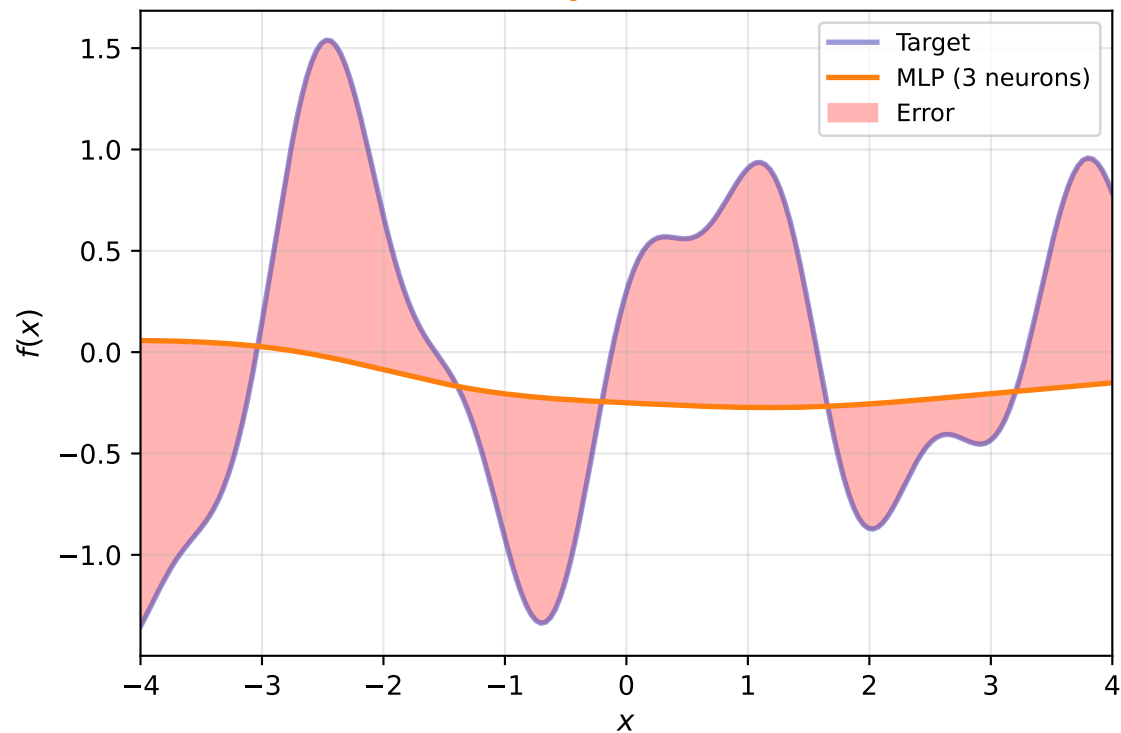


# Universal Approximation: MLPs Can Learn Any Function

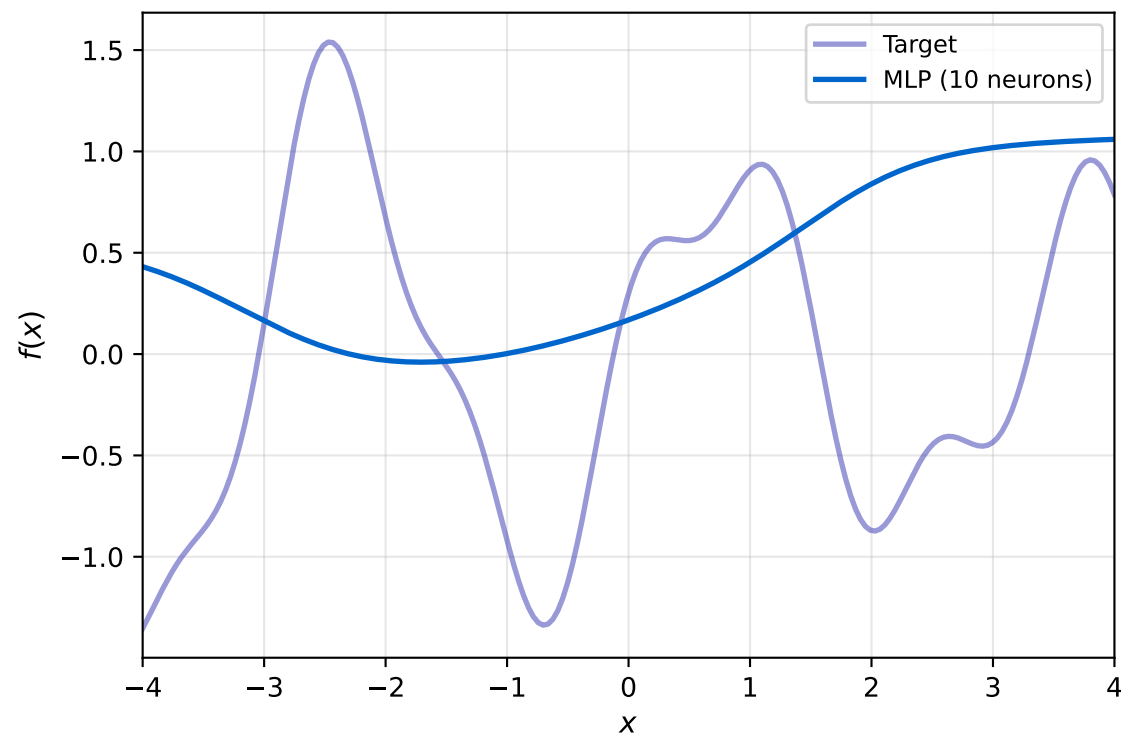
## Target: Complex Non-Linear Function



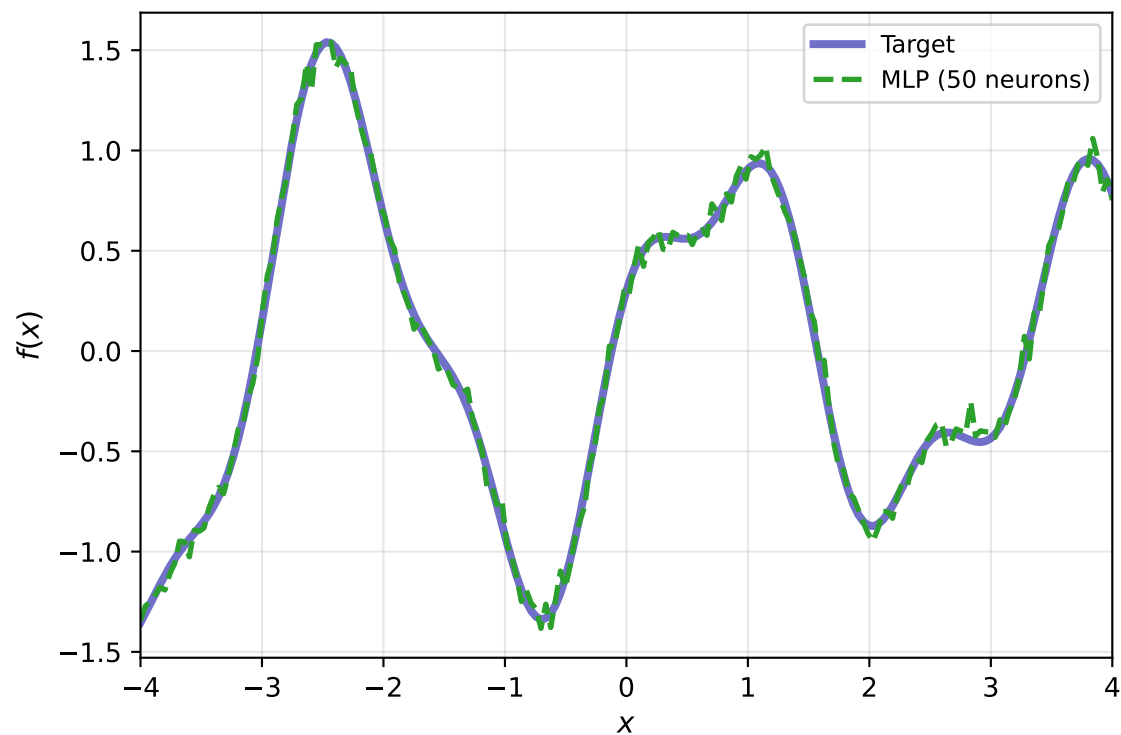
## Poor Fit: Only 3 Hidden Neurons



## Better: 10 Hidden Neurons



## Excellent: 50 Hidden Neurons



Universal Approximation Theorem (Cybenko, 1989): A feedforward network with a single hidden layer containing a finite number of neurons can approximate any continuous function on compact subsets of  $\mathbb{R}^n$