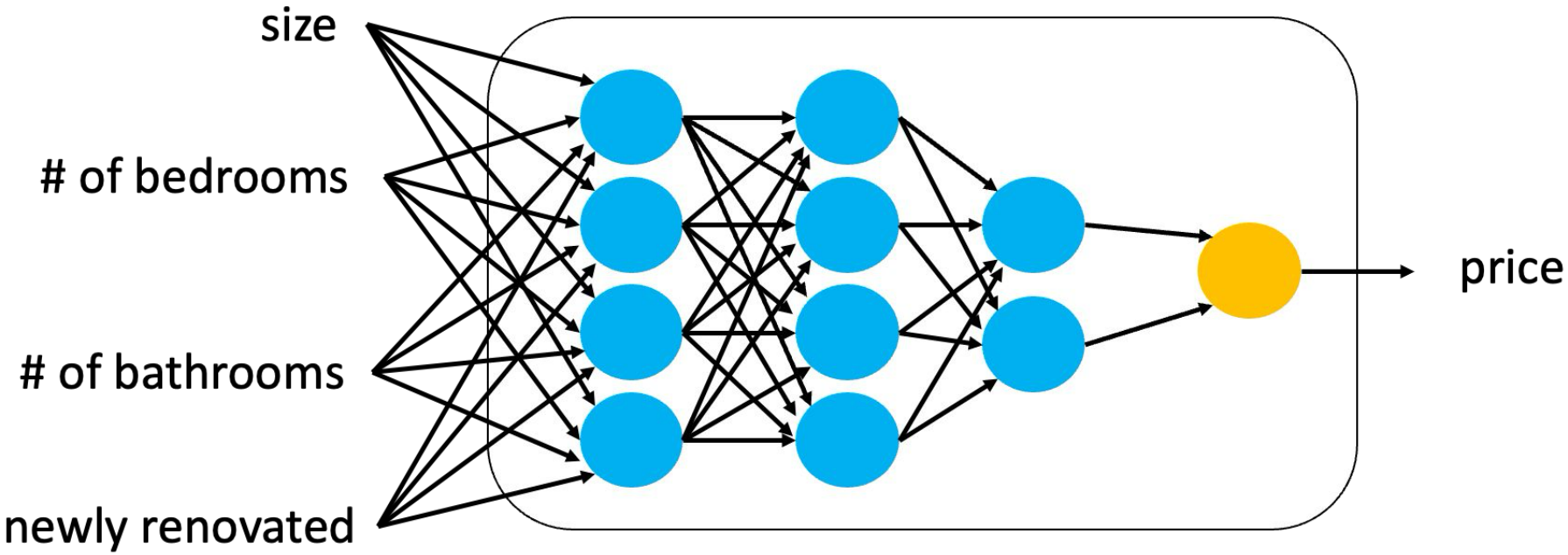
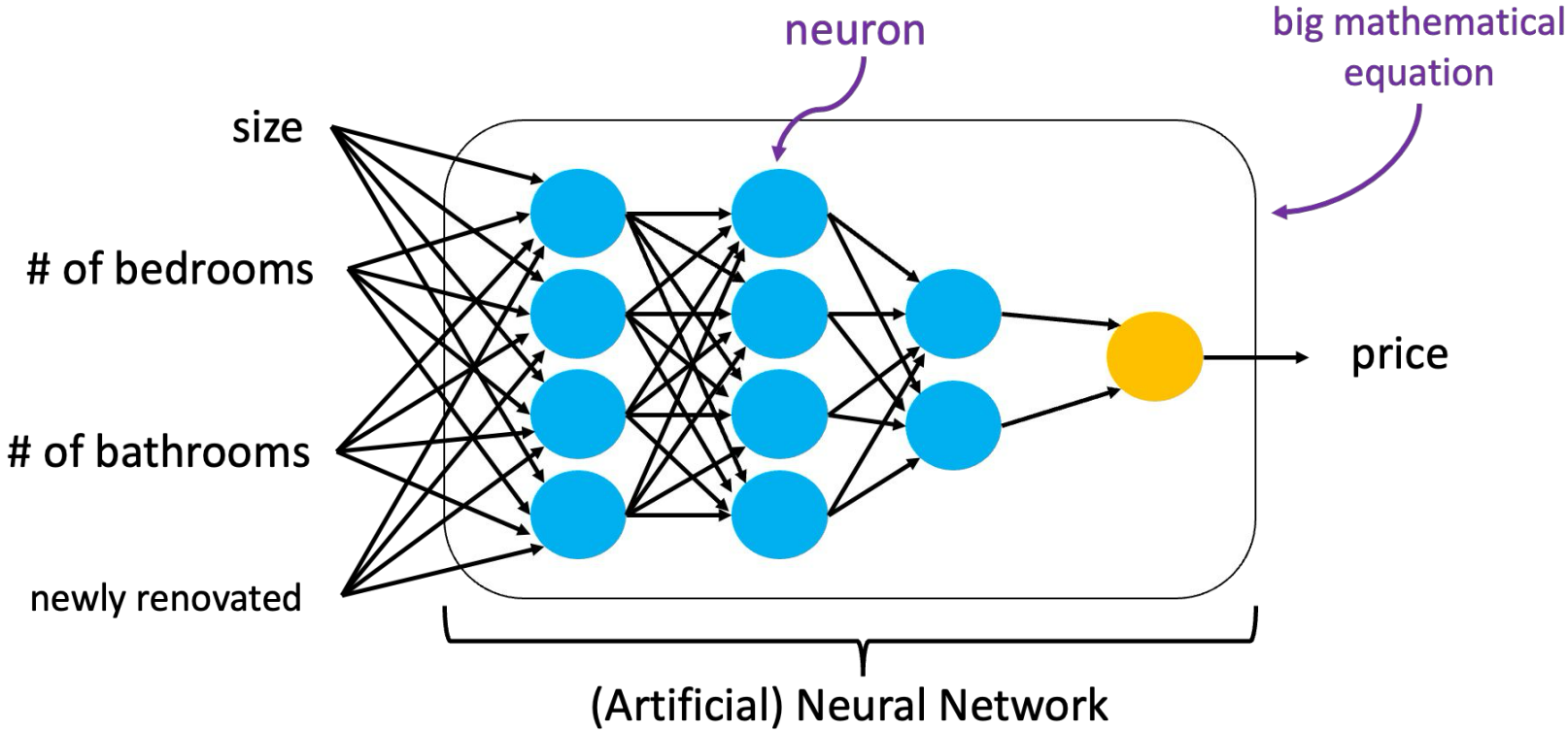
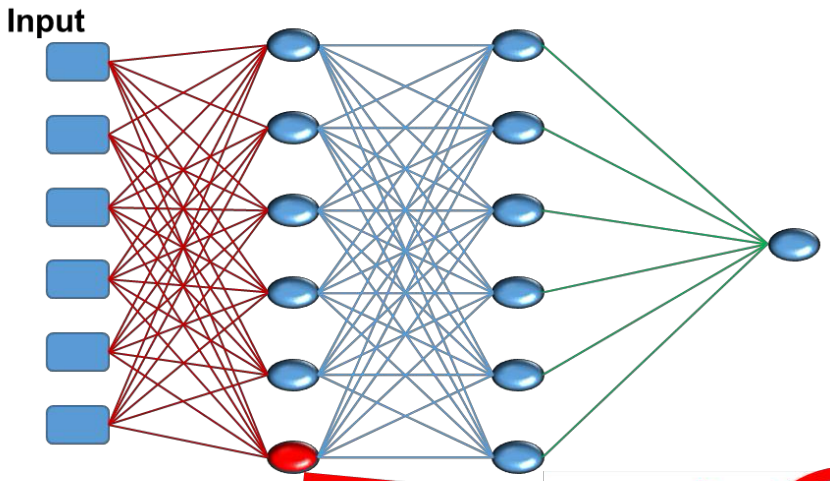
Neural networks



Neural networks

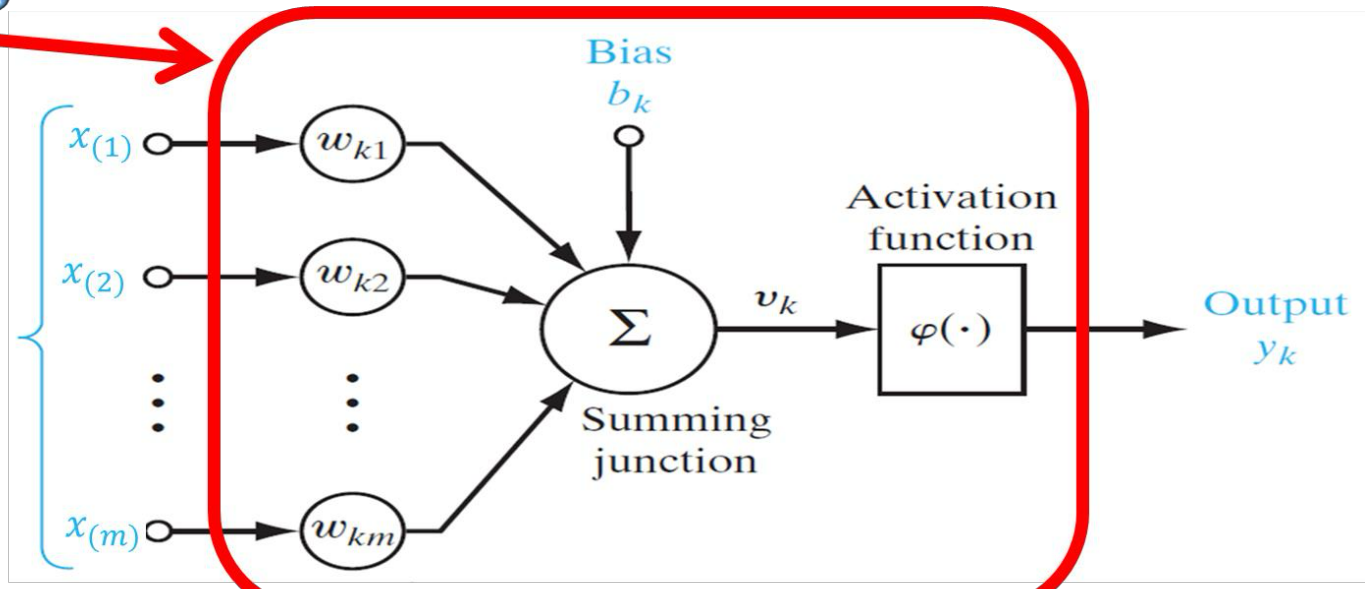


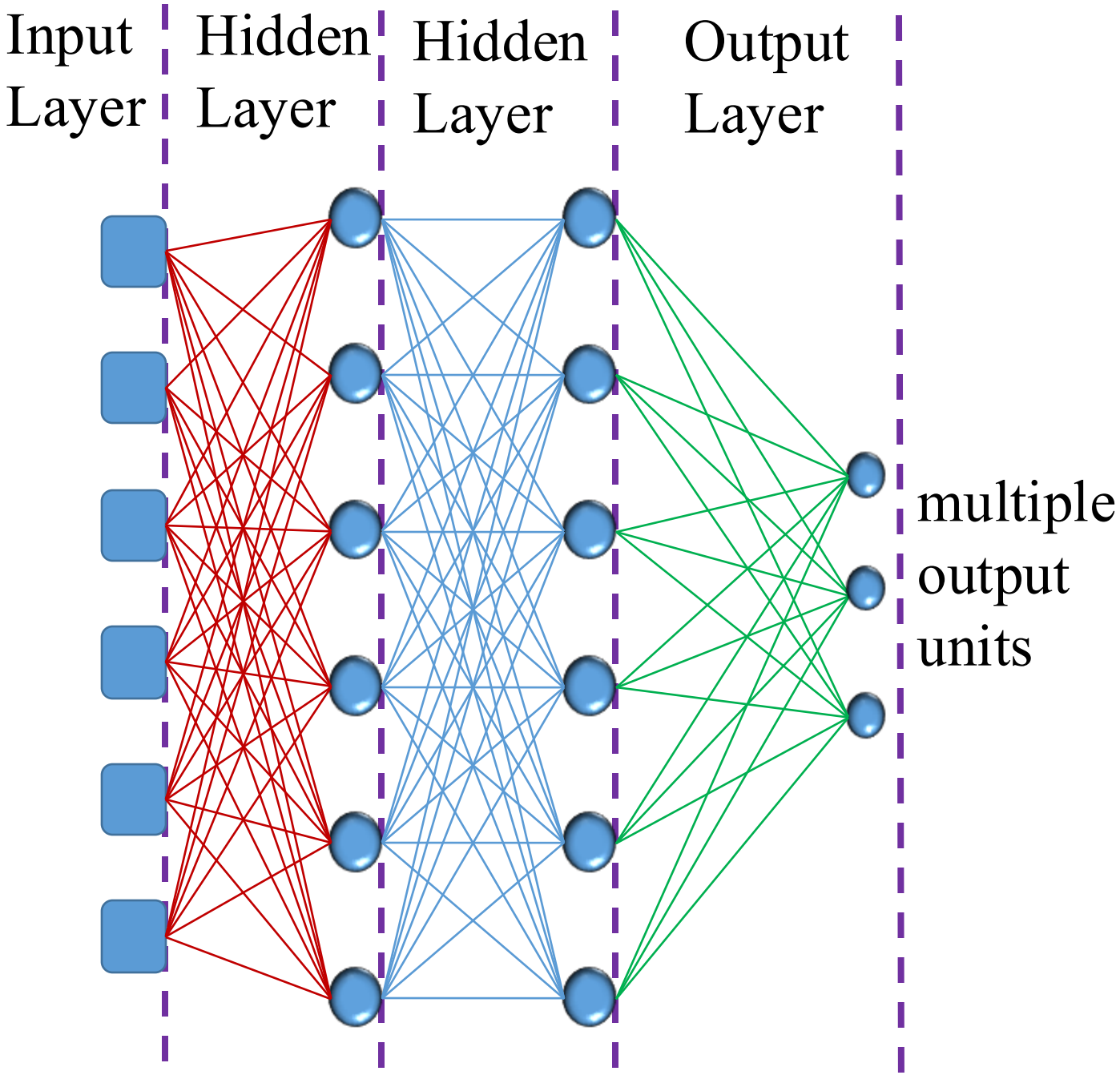
A Neural Network is a nonlinear model if the activation functions are nonlinear



This network is also called **Multi-Layer Perceptron (MLP)**

One Unit (Perceptron)





In general,
a neural network (MLP) could have many hidden layers, and the output could be a vector.

Training a Neural Network using Gradient Descent: Forward Pass, Backward Pass, Parameter Update

- Forward Pass to perform inference

  compute the output of each unit/layer

- Backward Pass to perform learning

  compute the derivatives of the loss

- Parameter Update to adjust the parameters using derivatives

  Update: $w \leftarrow w - \eta \frac{\partial L}{\partial w}$

  η is called learning rate

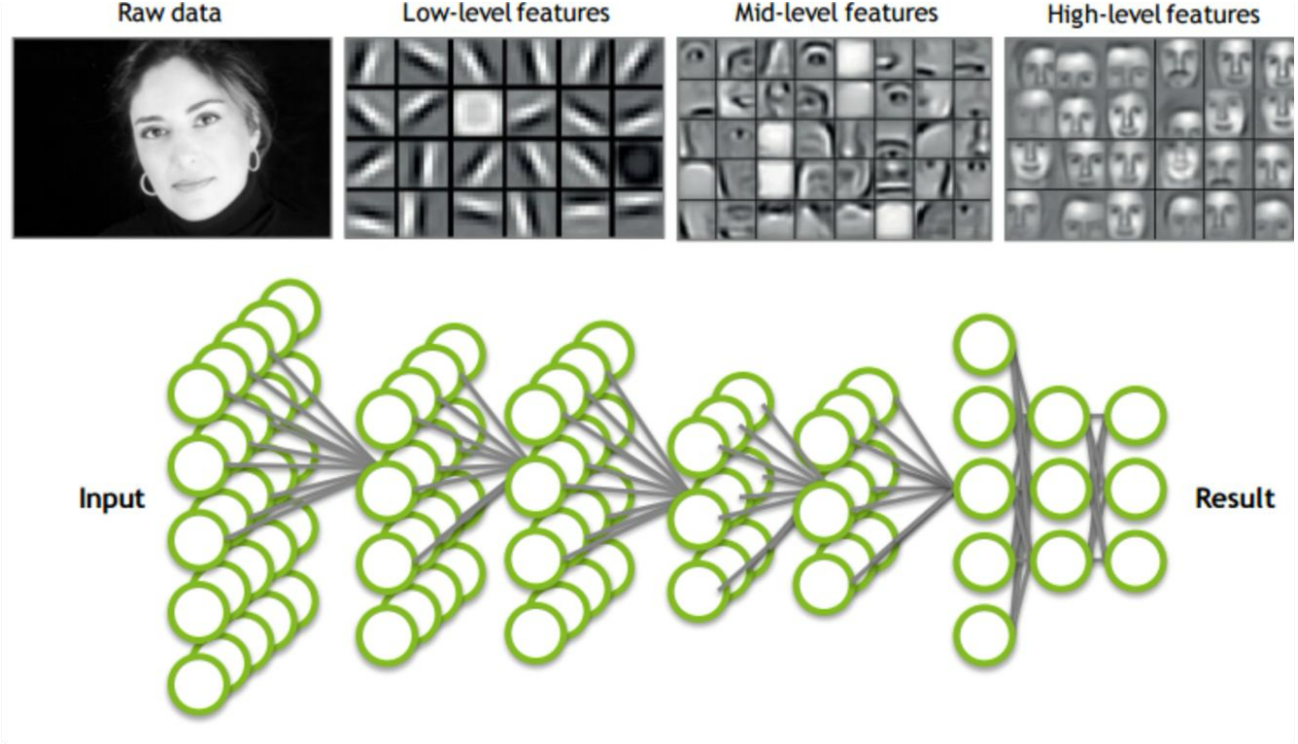
A Multi-layer (≥1 hidden layer) Neural Network is a Universal Function Approximator

- Universal Approximation Theorem
  - Every bounded continuous function can be approximated with arbitrary small error, by a network with one hidden layer
  - Activation functions need to be locally bounded and piecewise continuous

- Deep network vs shallow network

A continuous function can be approximated by a deep network or a shallow network. The deep network usually uses less number of units.

Image Recognition



[Image credit: NVIDIA]