

Neural Networks

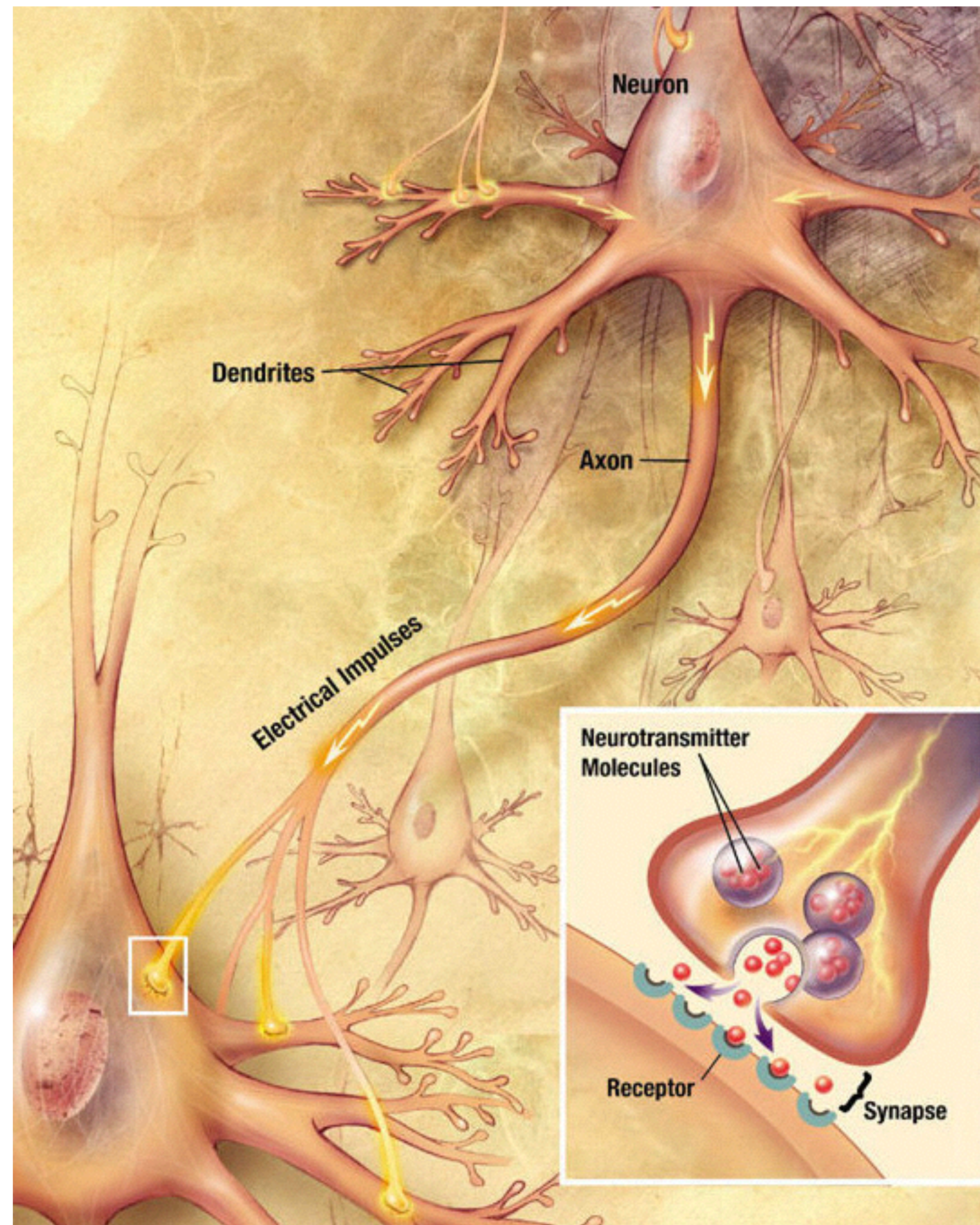
Machine Learning
CSx824/ECEEx242
Bert Huang
Virginia Tech

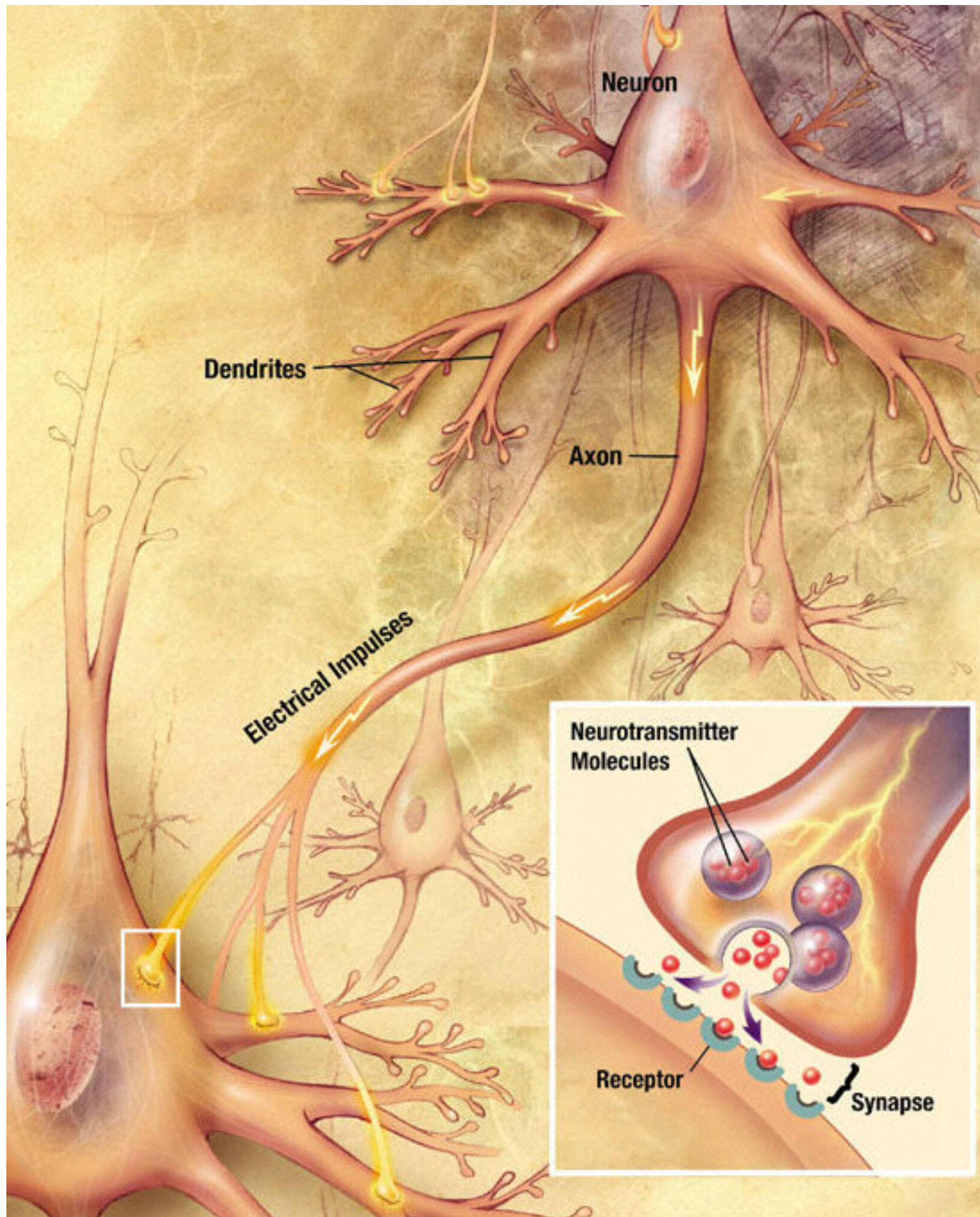
(Feed Forward) Neural Networks

Machine Learning
CSx824/ECEx242
Bert Huang
Virginia Tech

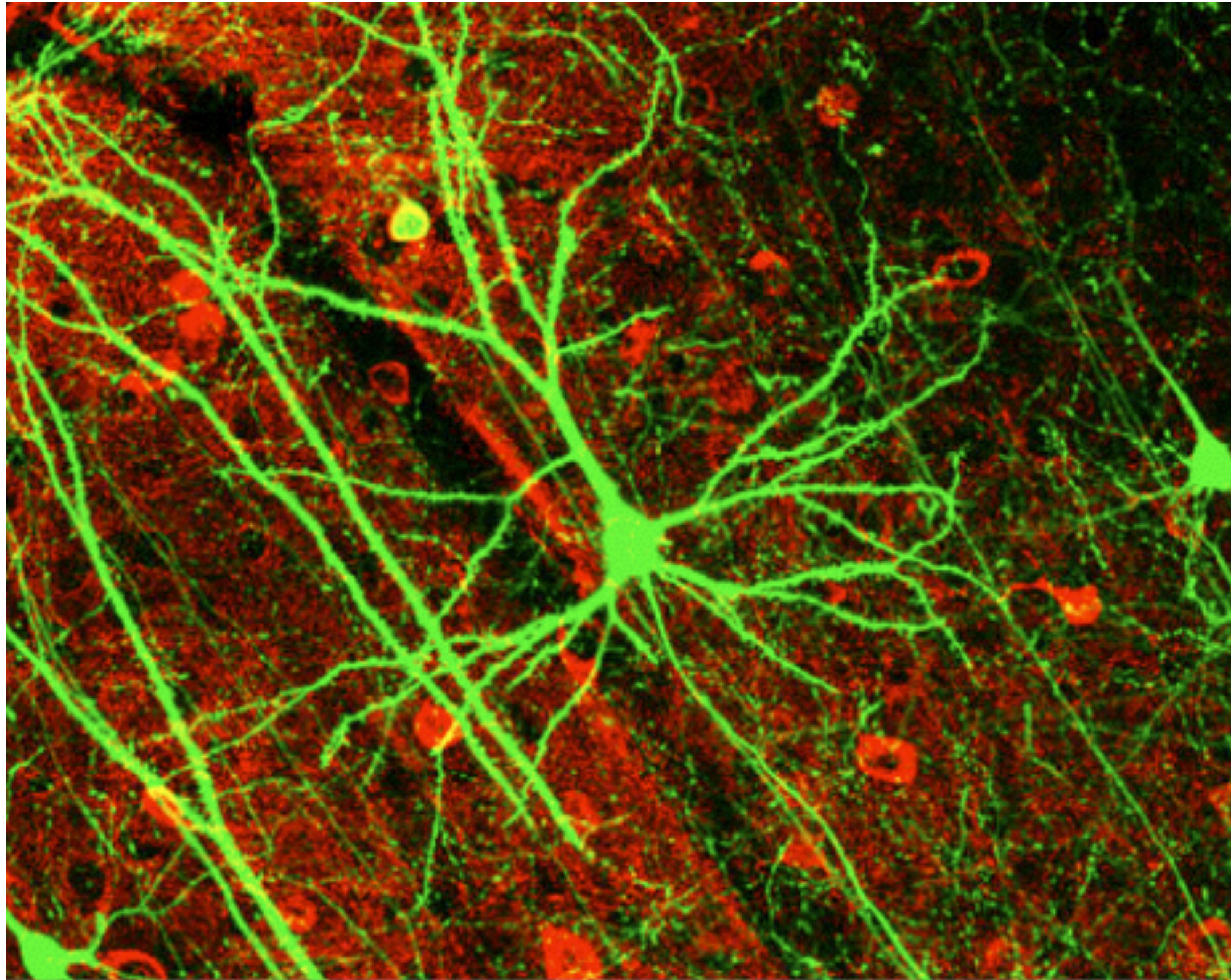
Outline

- Logistic regression and perceptron as neural networks
- Multi-layered perceptron
- Nonlinearity





https://en.wikipedia.org/wiki/Neuron#/media/File:Chemical_synapse_schema_cropped.jpg



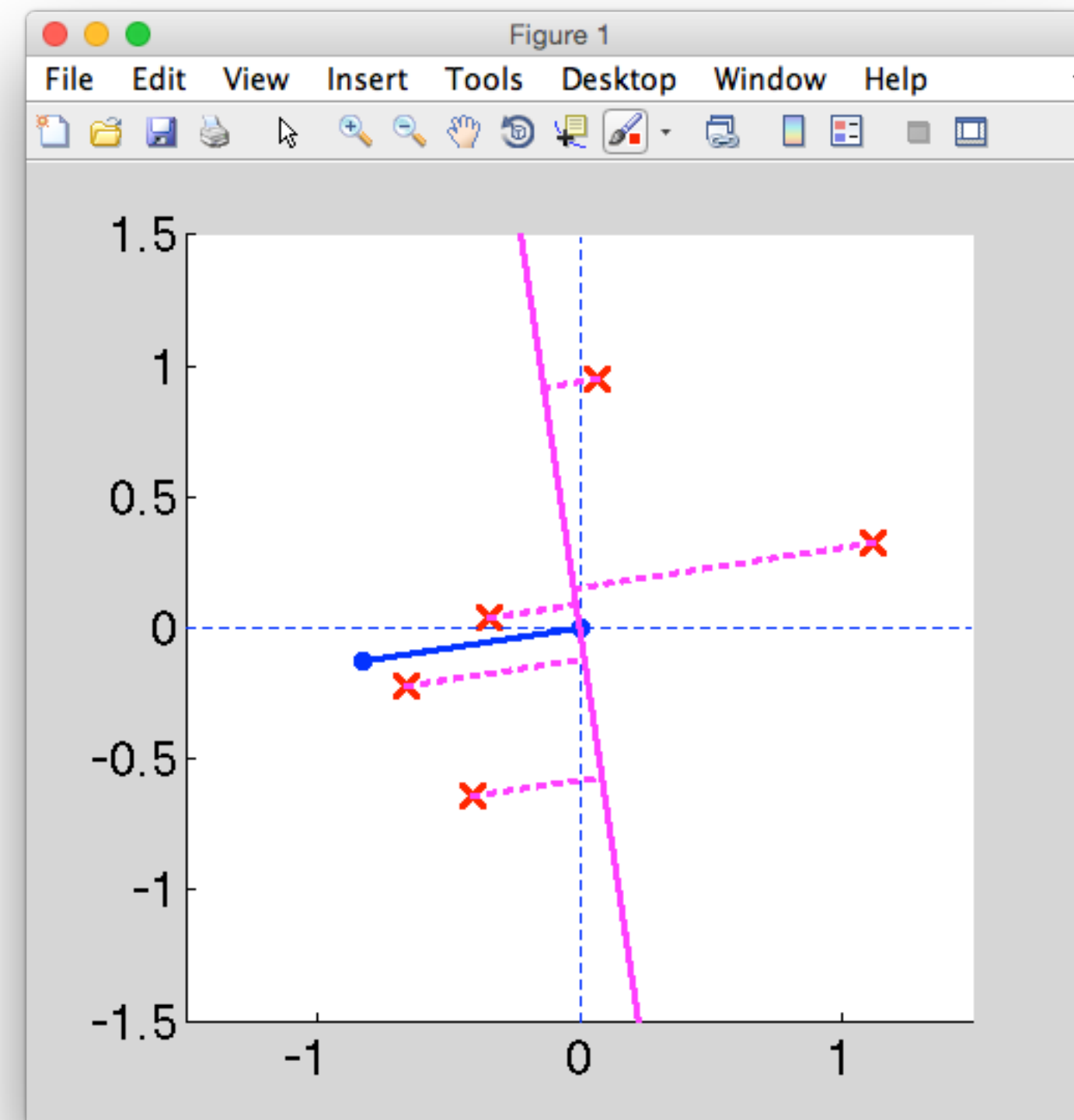
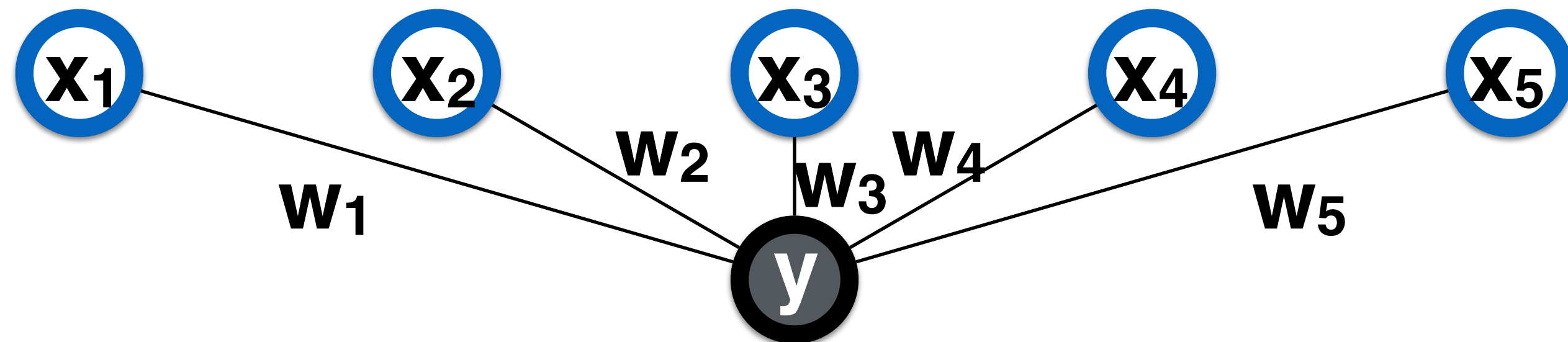
<https://en.wikipedia.org/wiki/Neuron#/media/File:GFPneuron.png>

Parameterizing $p(y|x)$

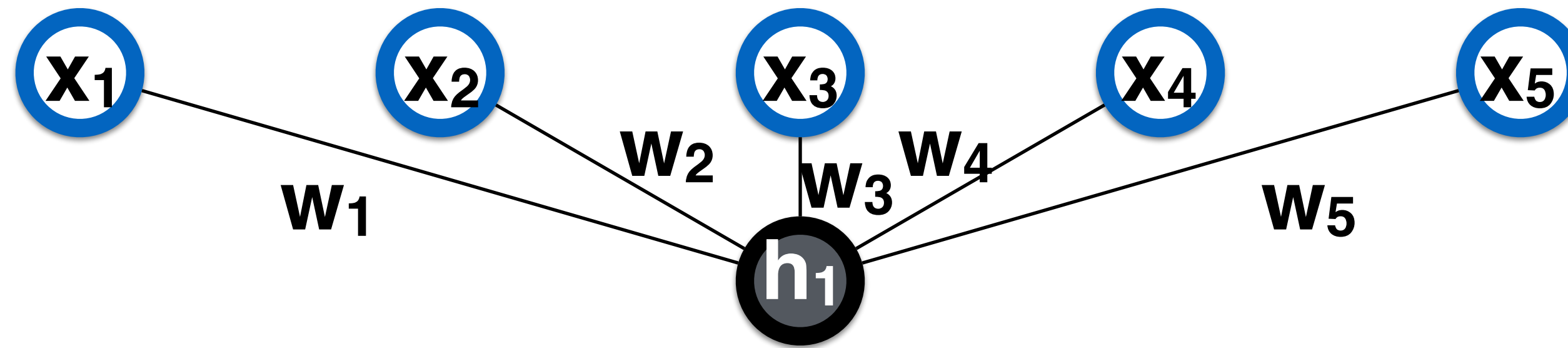
$$p(y|x) := f$$

$$f : \mathbb{R}^d \rightarrow [0, 1]$$

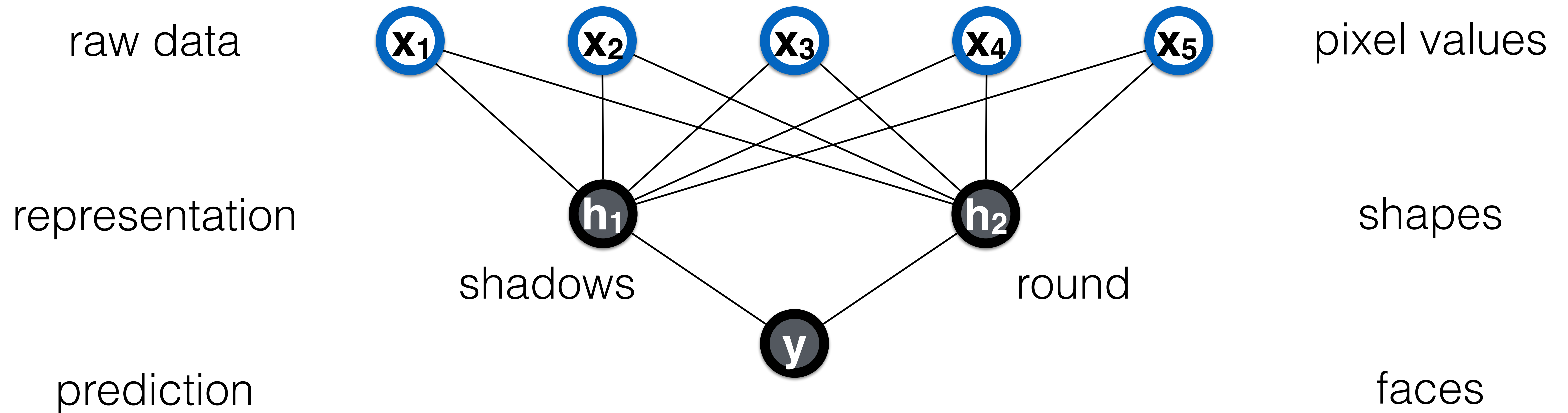
$$f(x) := \frac{1}{1 + \exp(-w^\top x)}$$



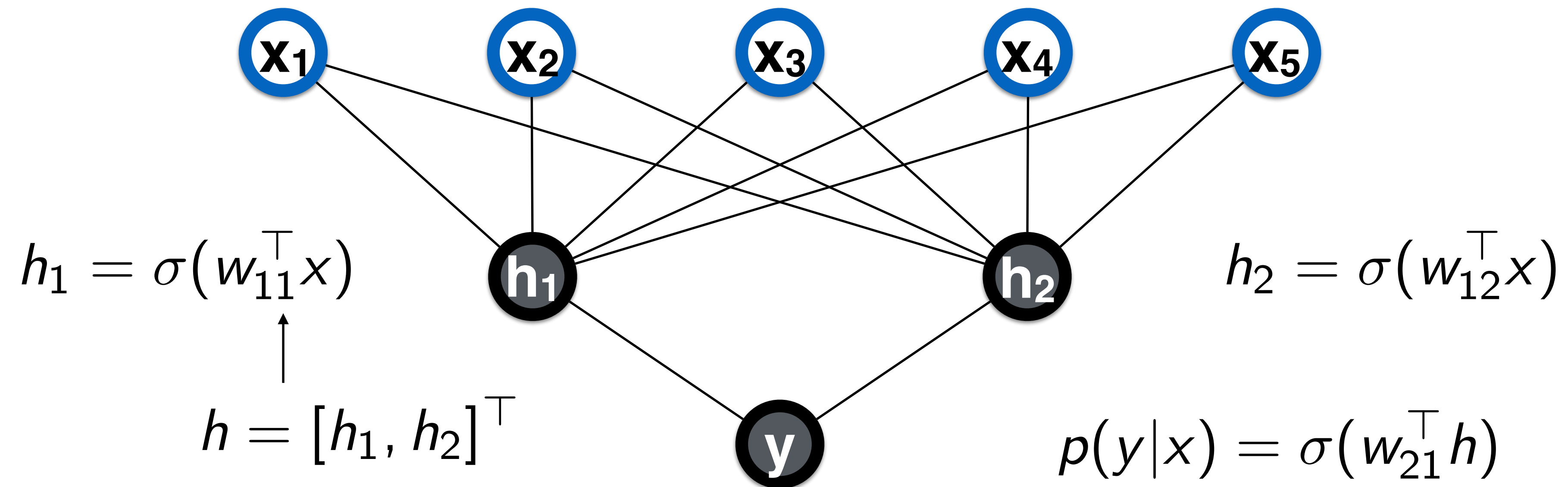
Multi-Layered Perceptron



Multi-Layered Perceptron

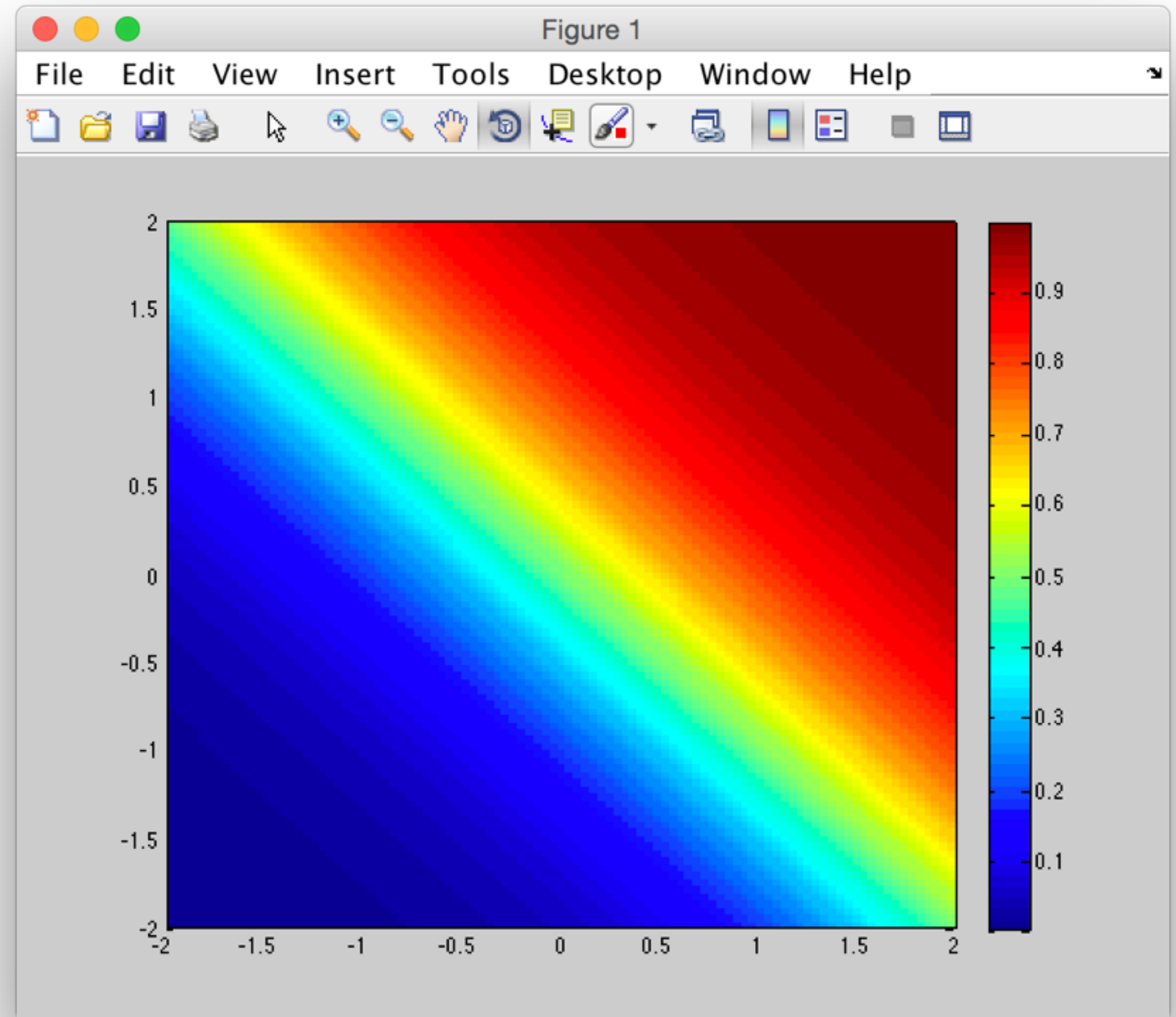
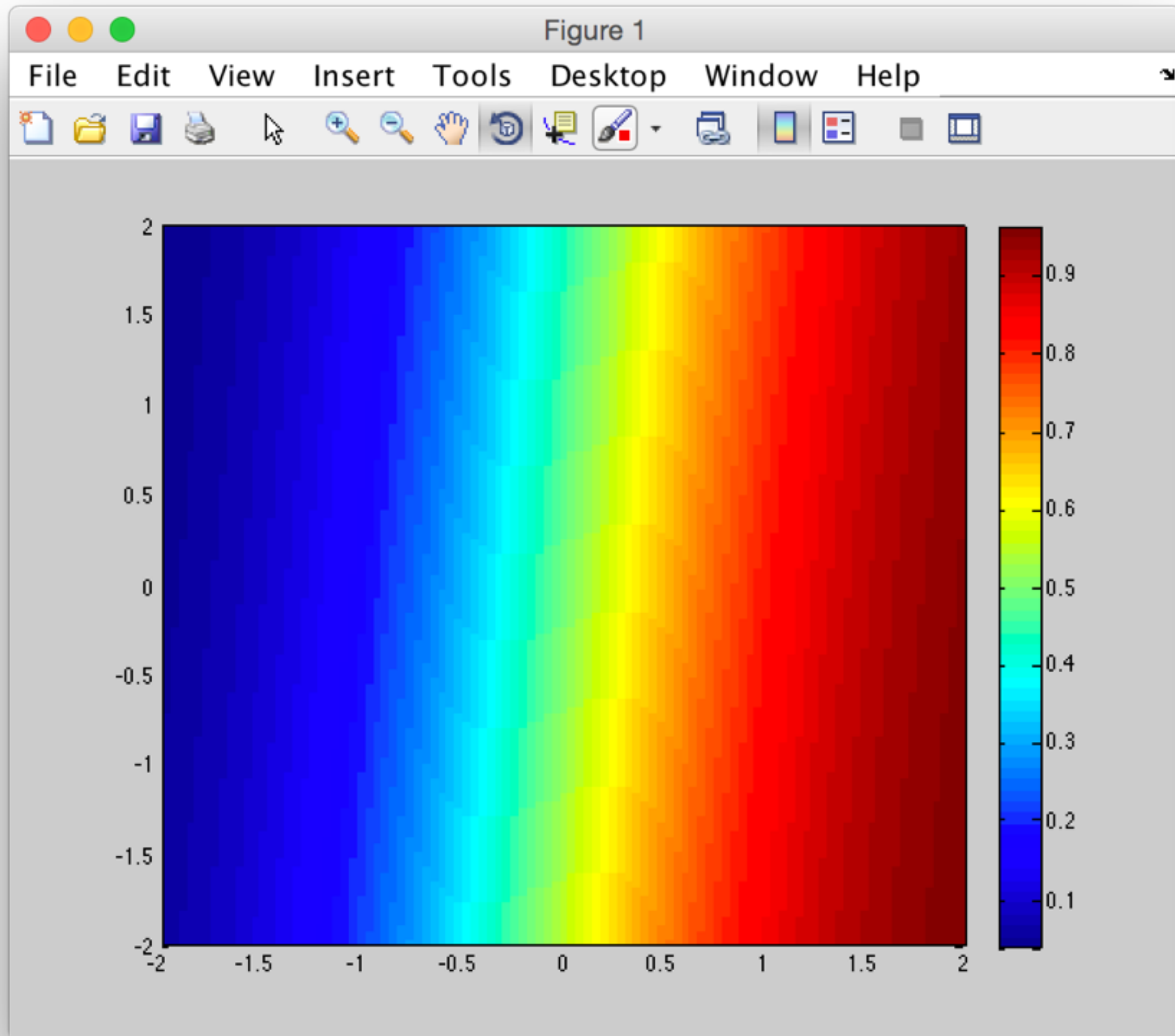


Multi-Layered Perceptron

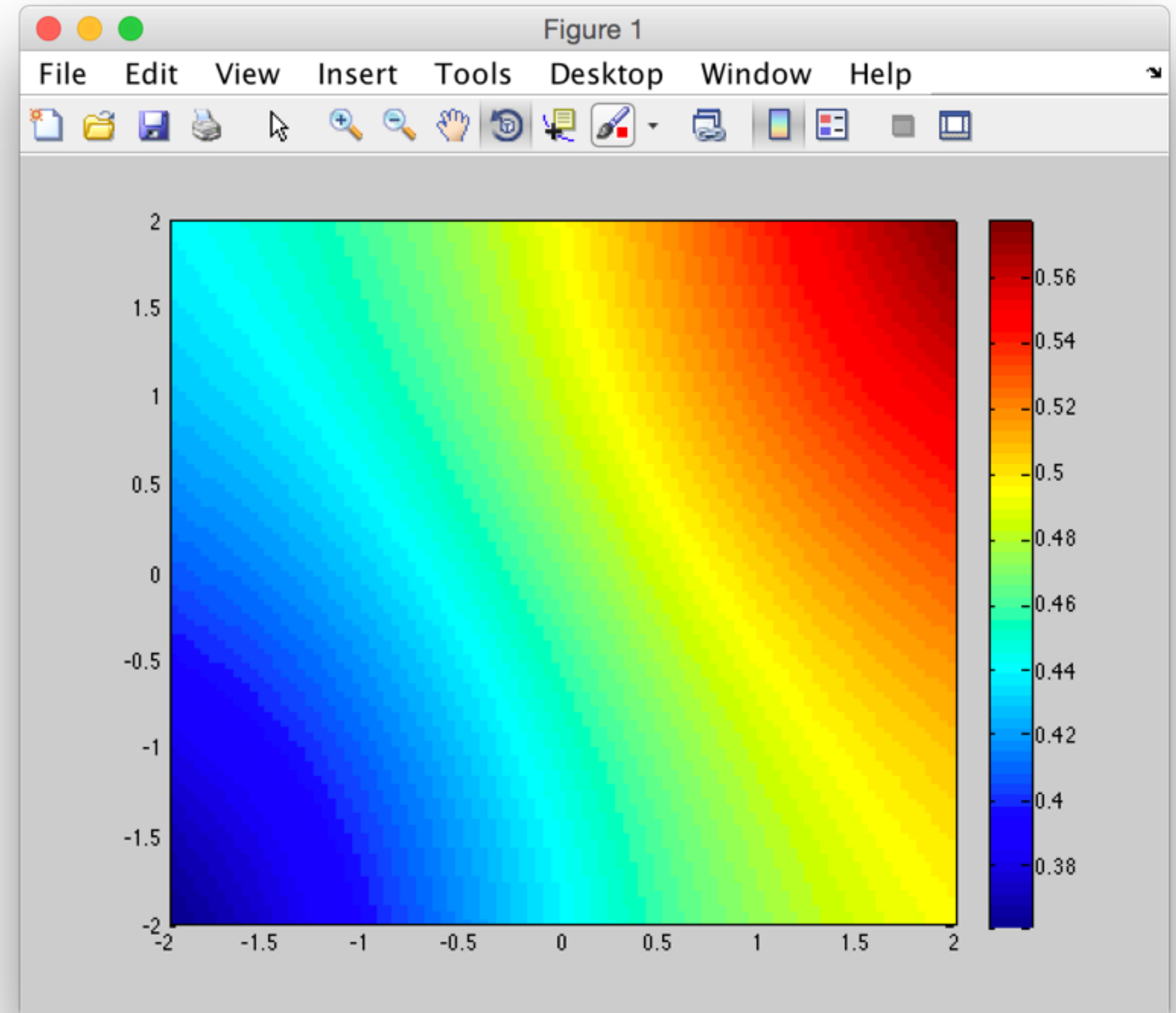
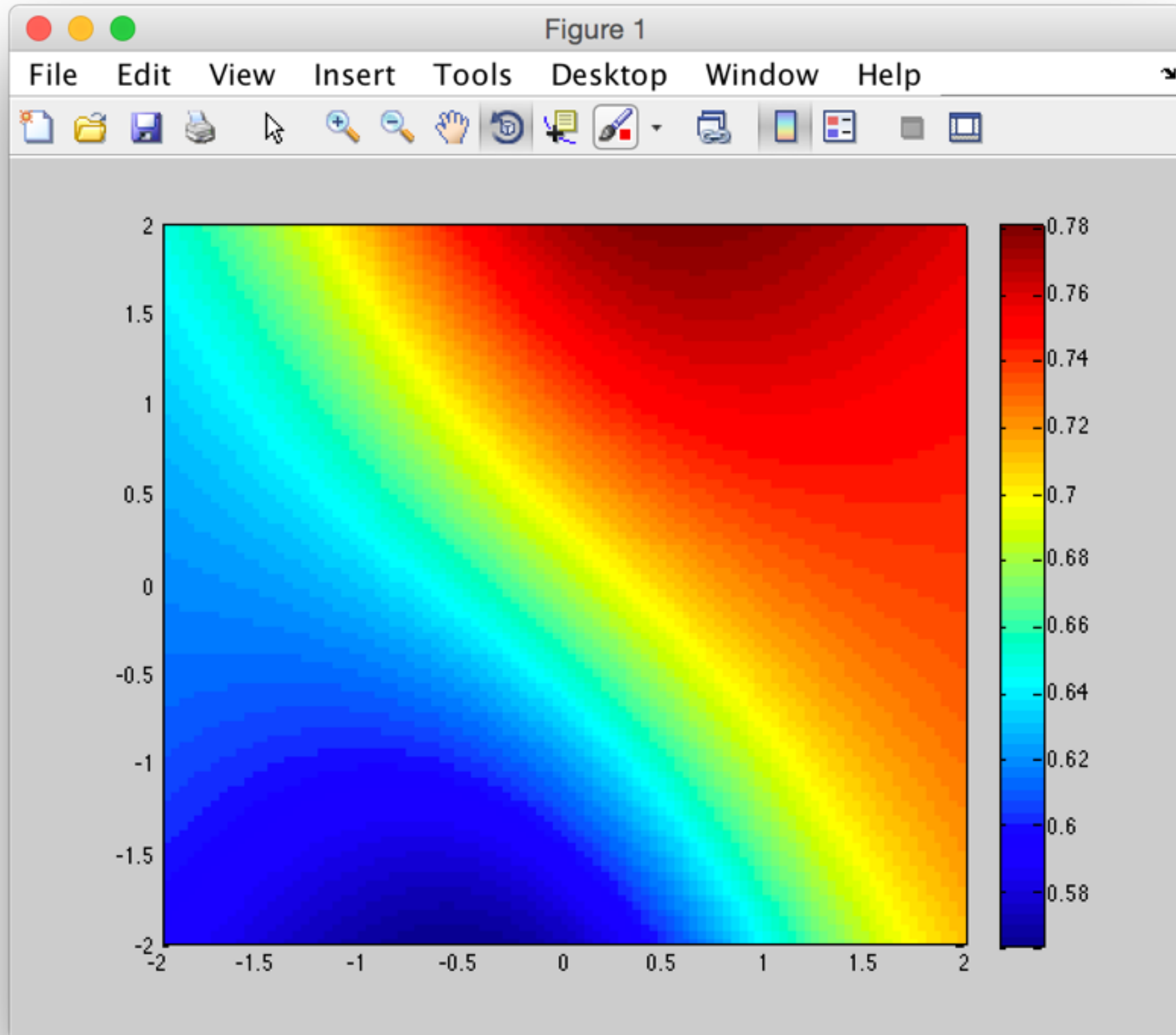


$$p(y|x) = \sigma \left(w_{21}^\top \left[\sigma(w_{11}^\top x), \sigma(w_{12}^\top x) \right]^\top \right)$$

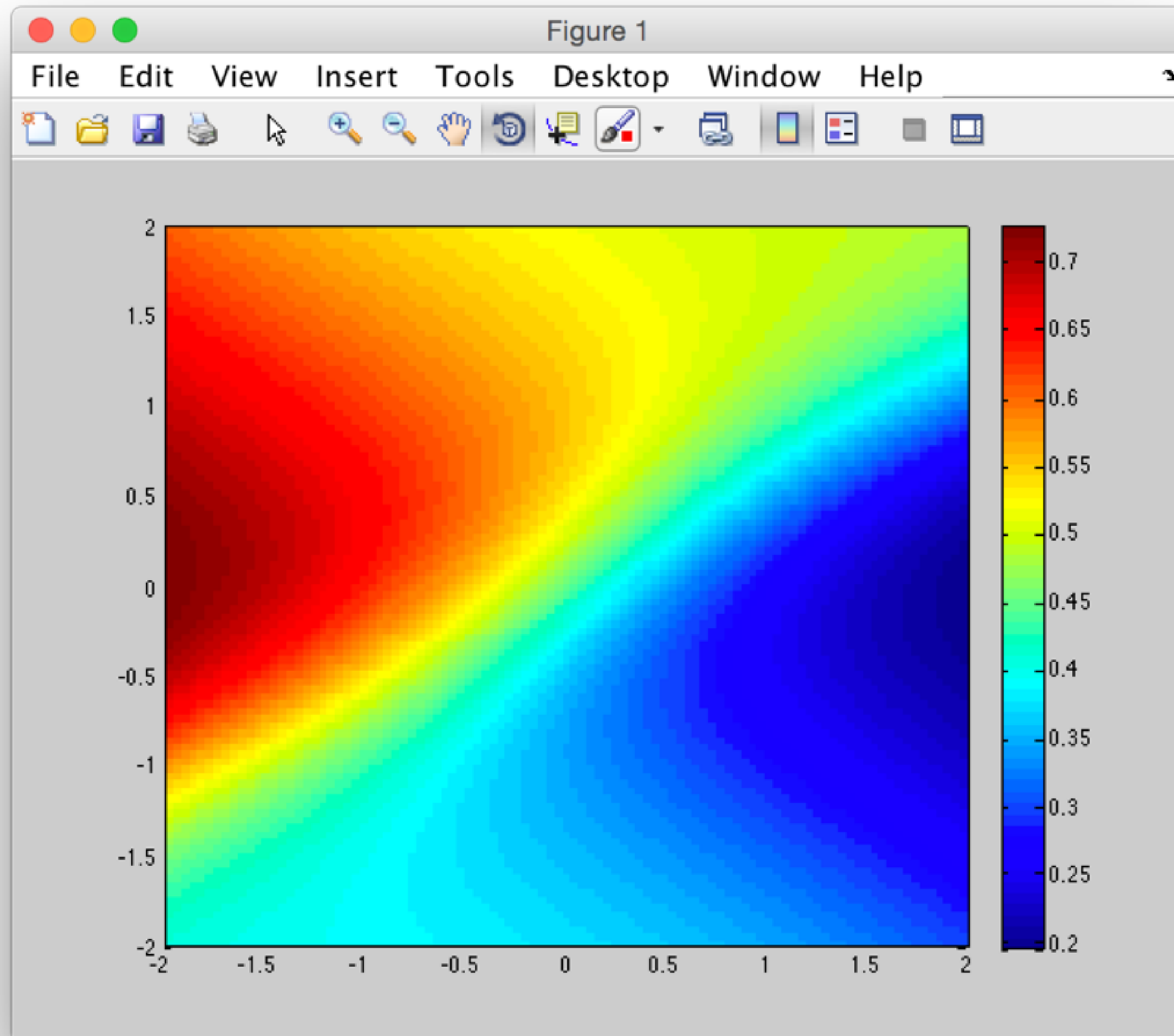
Decision Surface: Logistic Regression



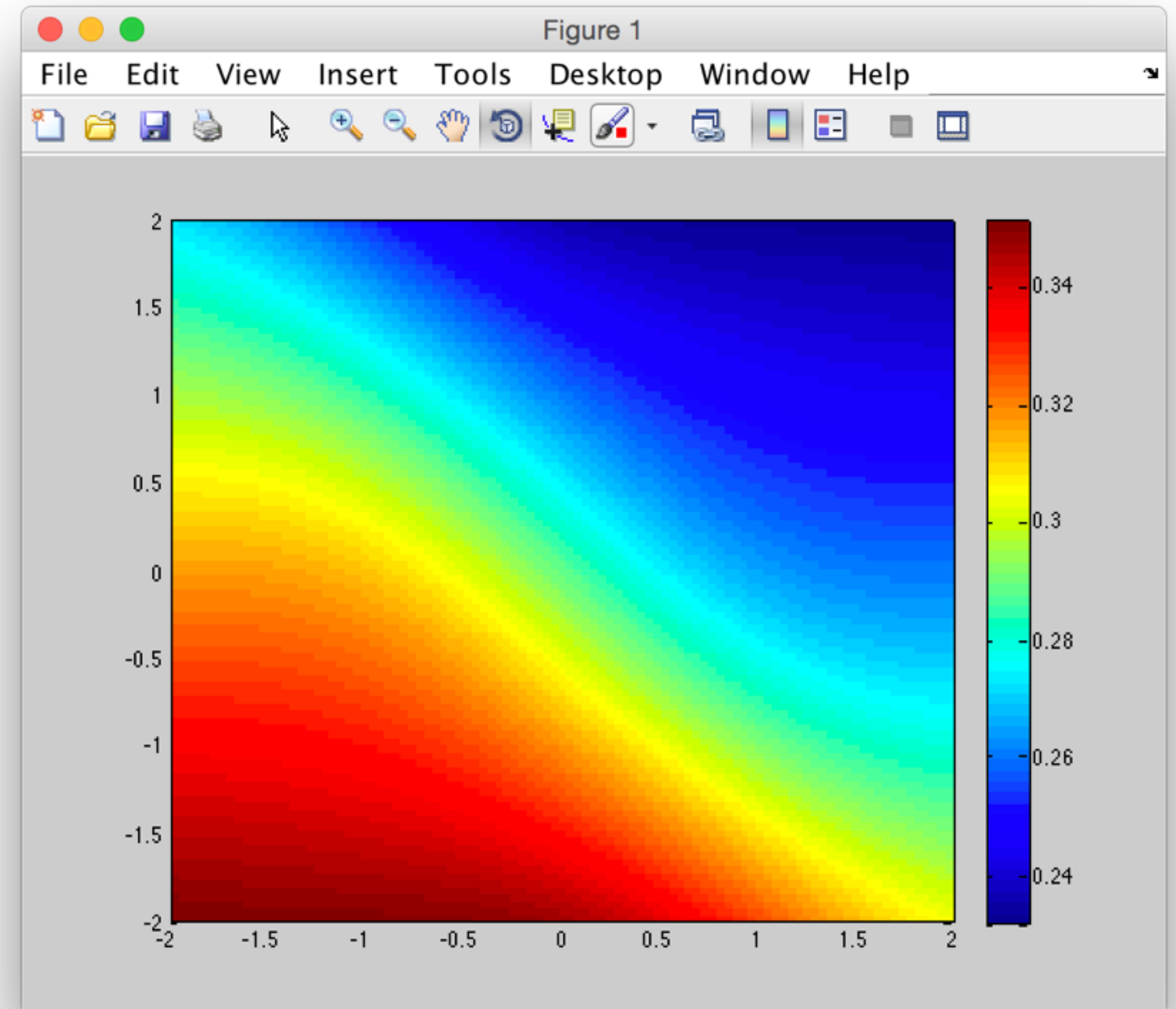
Decision Surface: 2-Layer, 2 Hidden Units



Decision Surface: 2-Layer, More Hidden Units

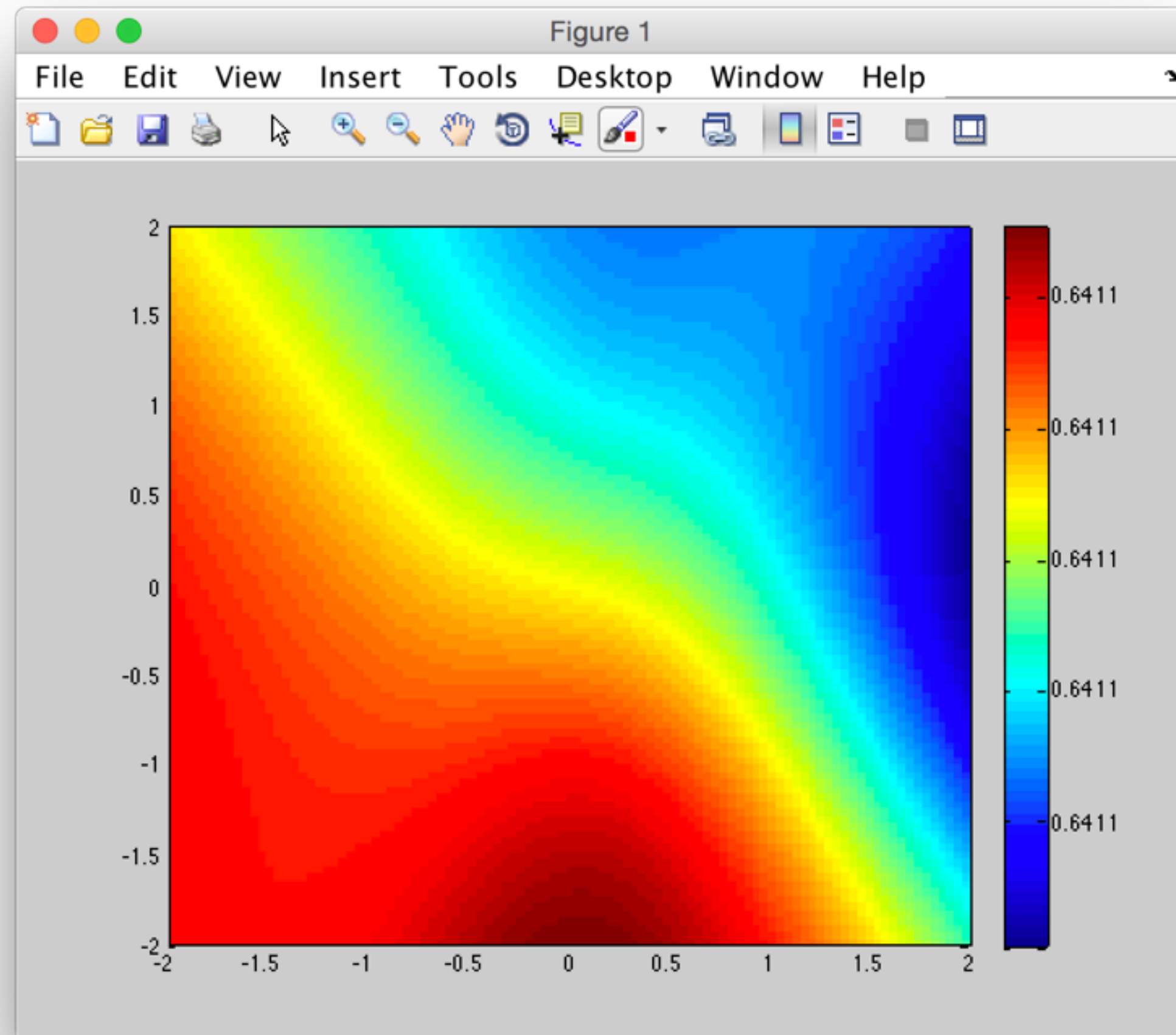


3 hidden units

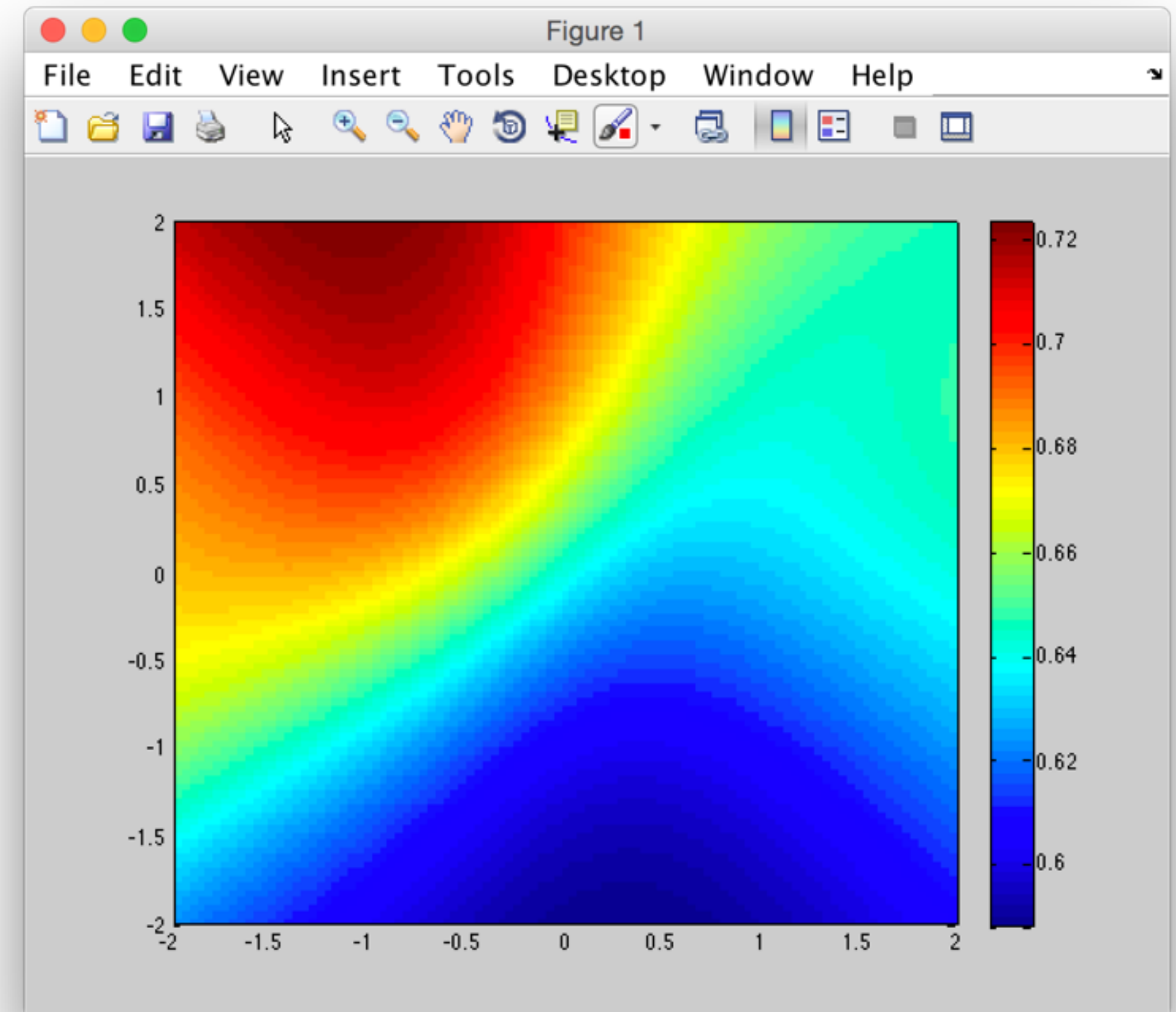


10 hidden units

Decision Surface: More Layers, More Hidden Units



10 layers, 5 hidden units per layer



4 layers, 10 hidden units each layer

Training

- Back propagation:
 - Compute hidden unit activations: **forward propagation**
 - Compute gradient at output layer: error
 - Propagate error back one layer at a time
- Chain rule via dynamic programming