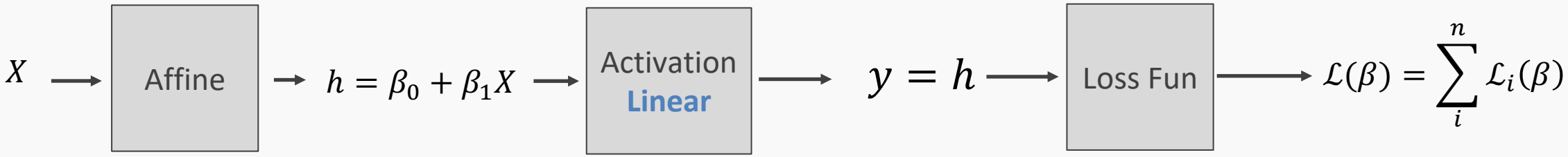
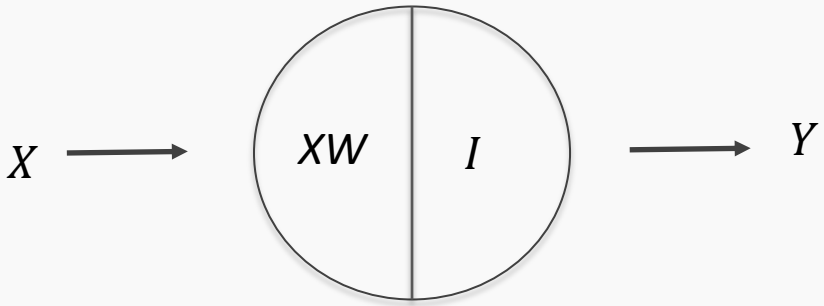
Up to this point we just re-branded logistic regression to look like a neuron.

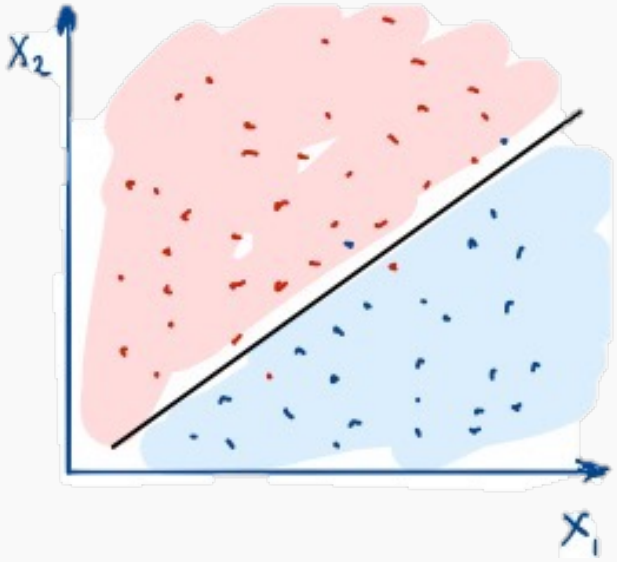
How about linear regression?

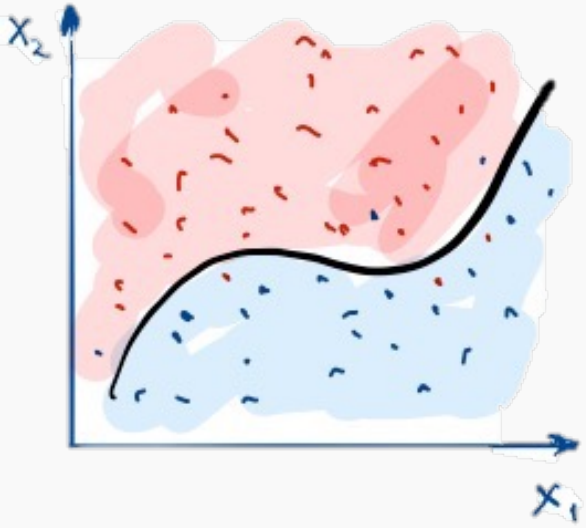


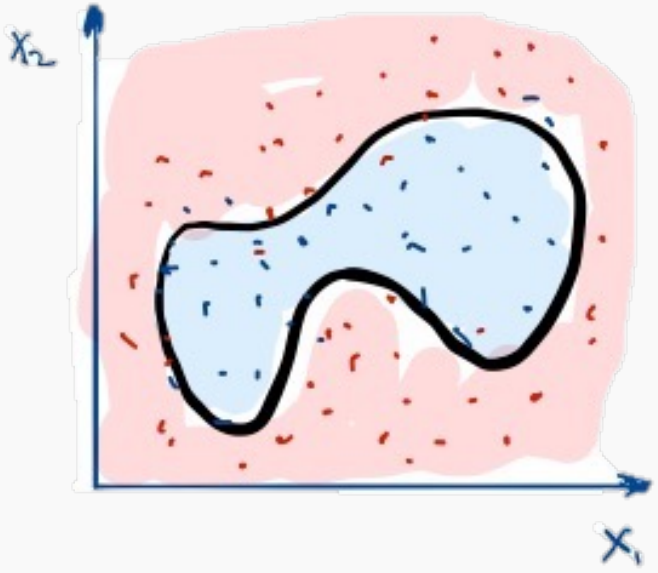


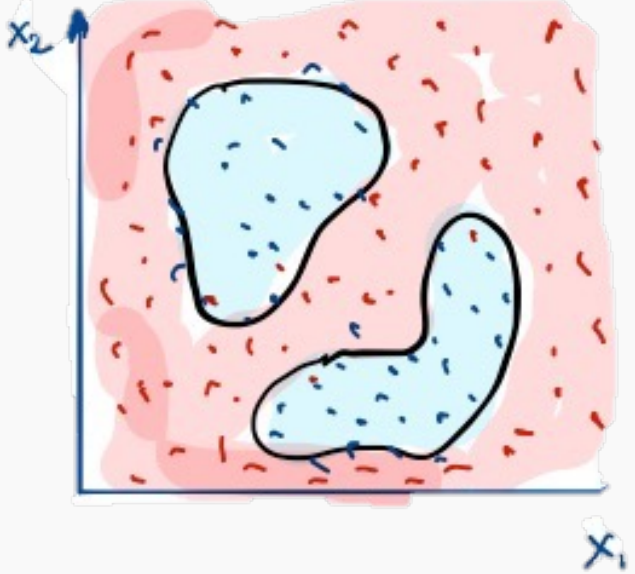
Where I is the identity function

So what's the big deal about Neural Networks?

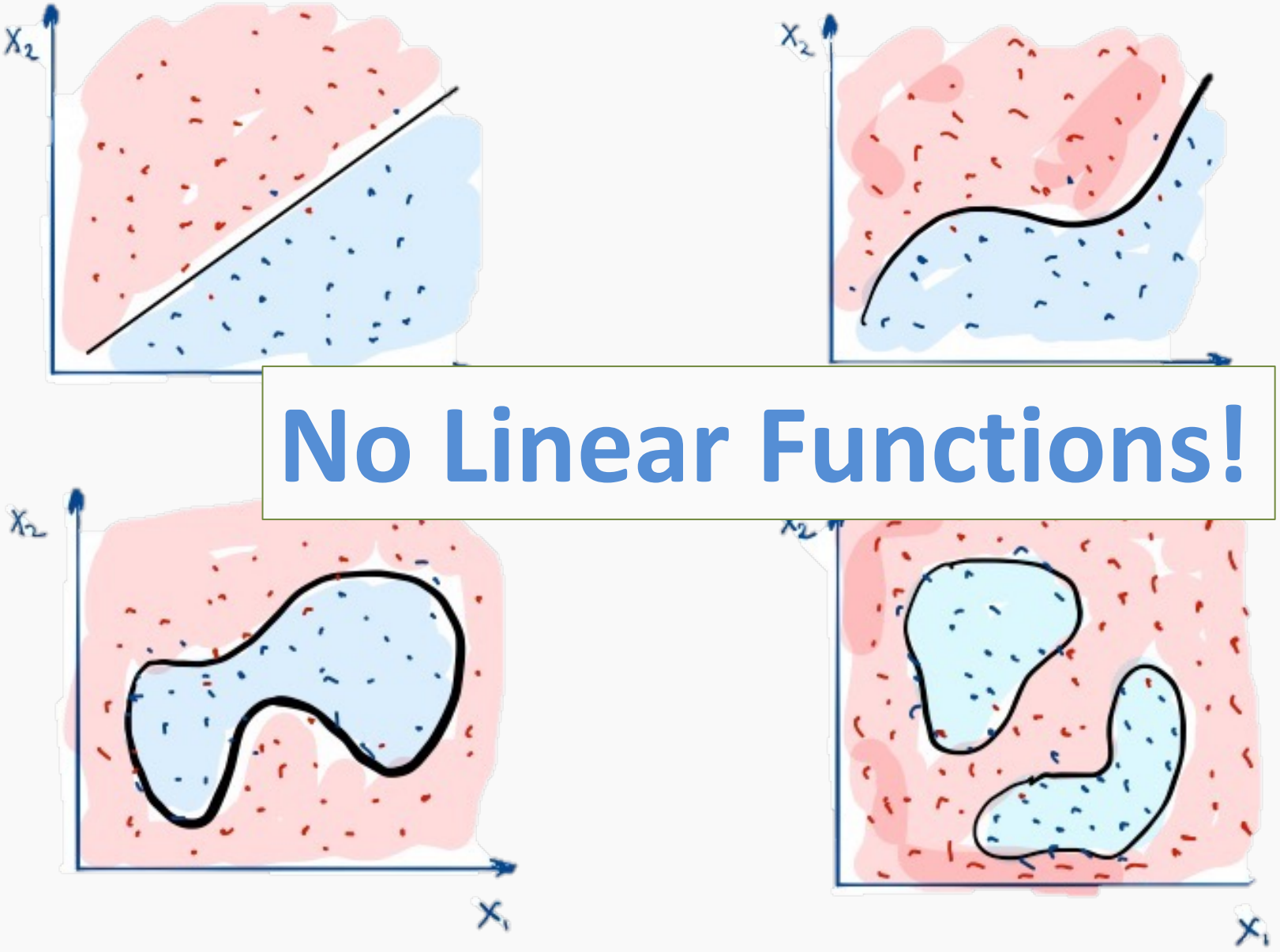






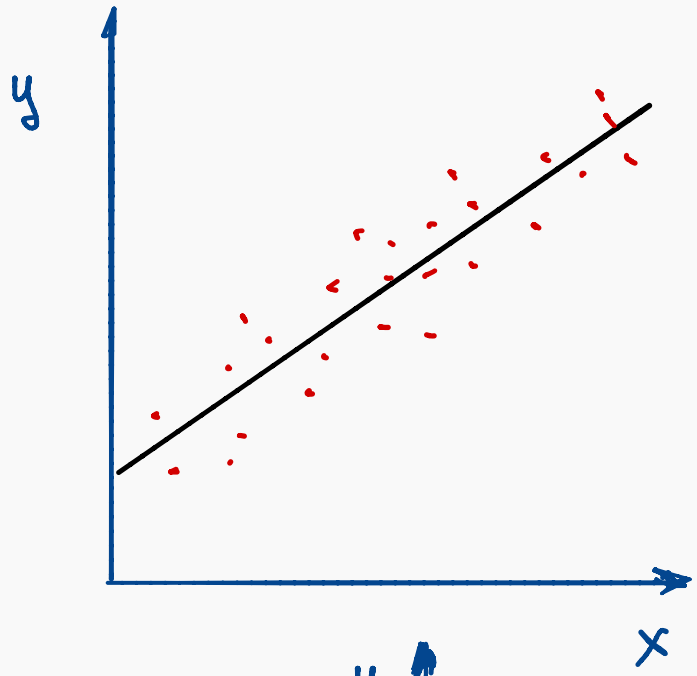


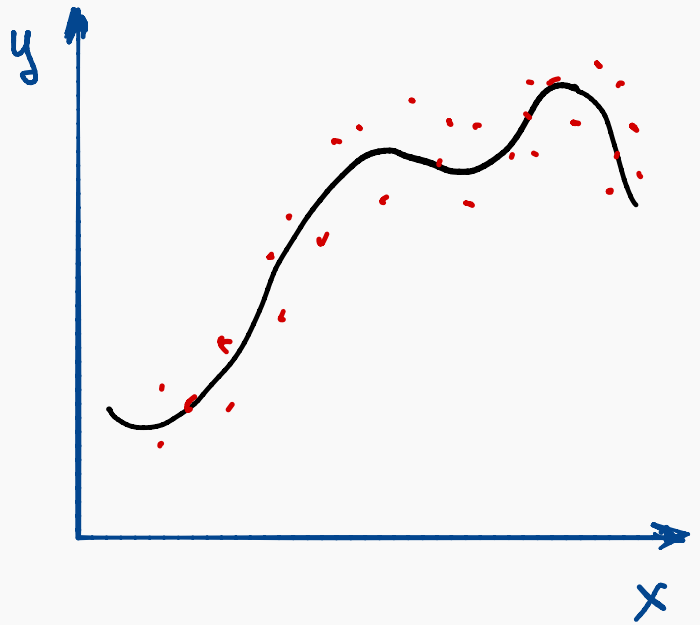
So what's the big deal about Neural Networks?

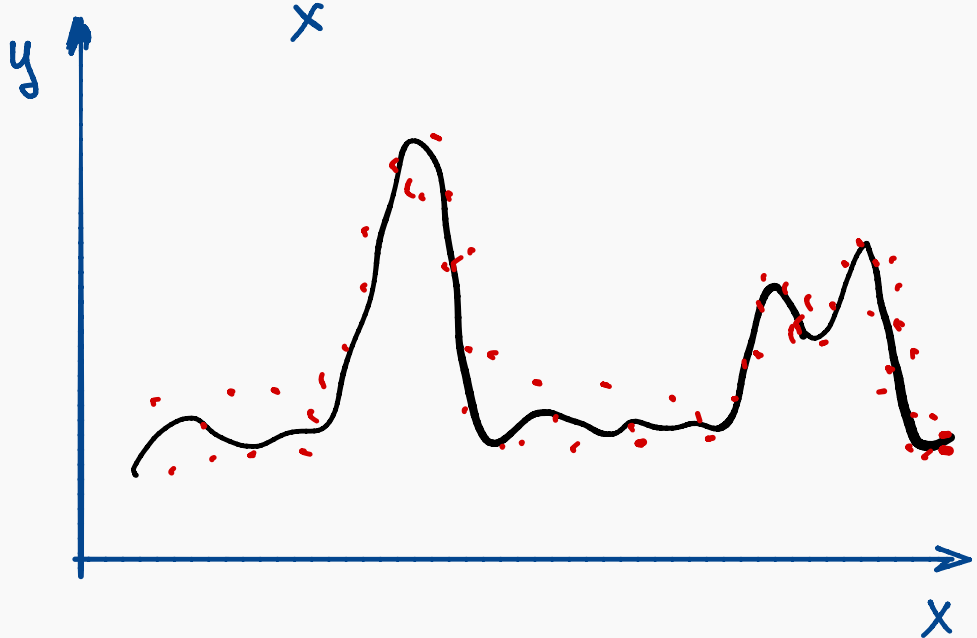


No Linear Functions!

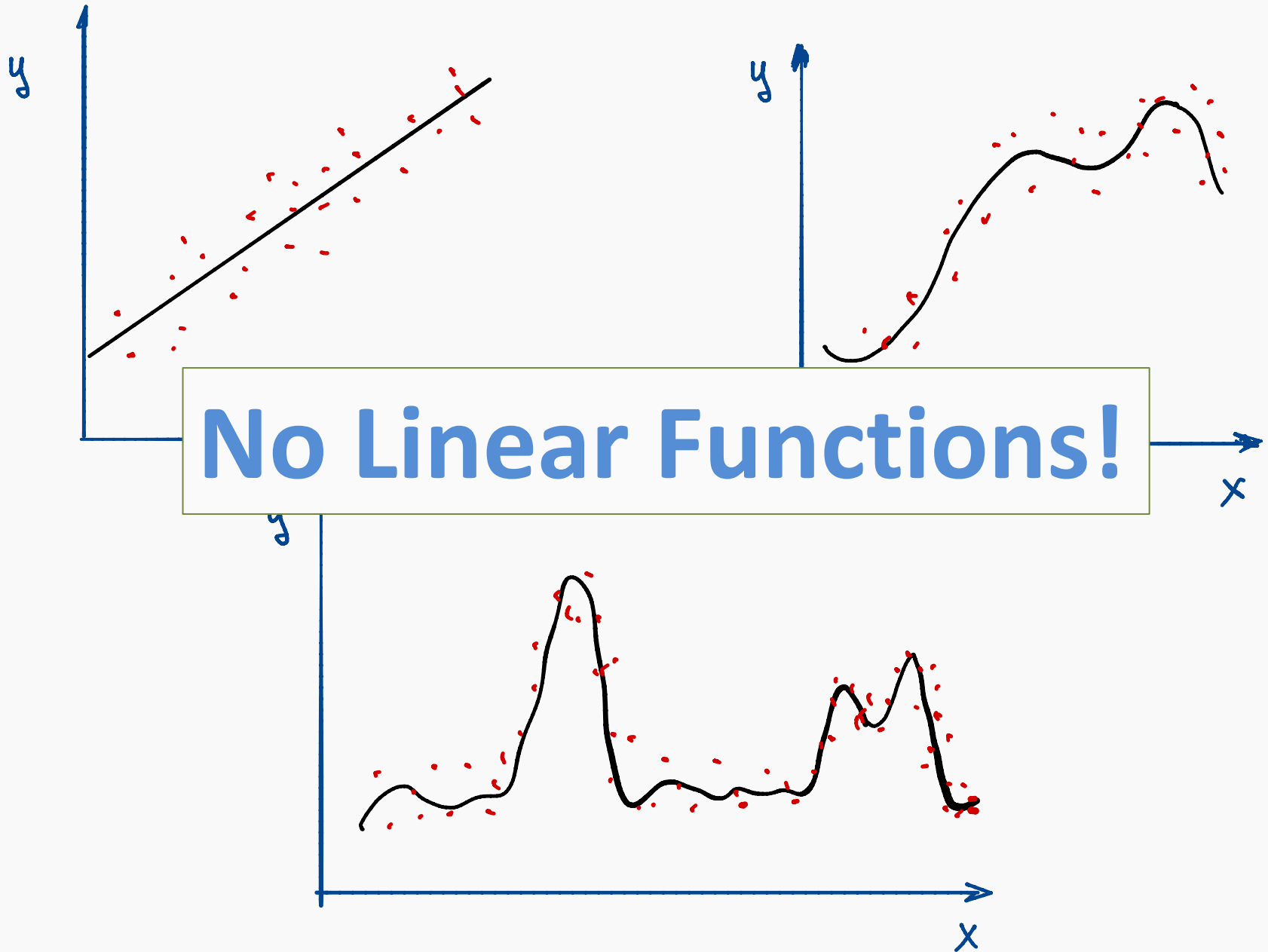
For regression?







For regression?



Outline

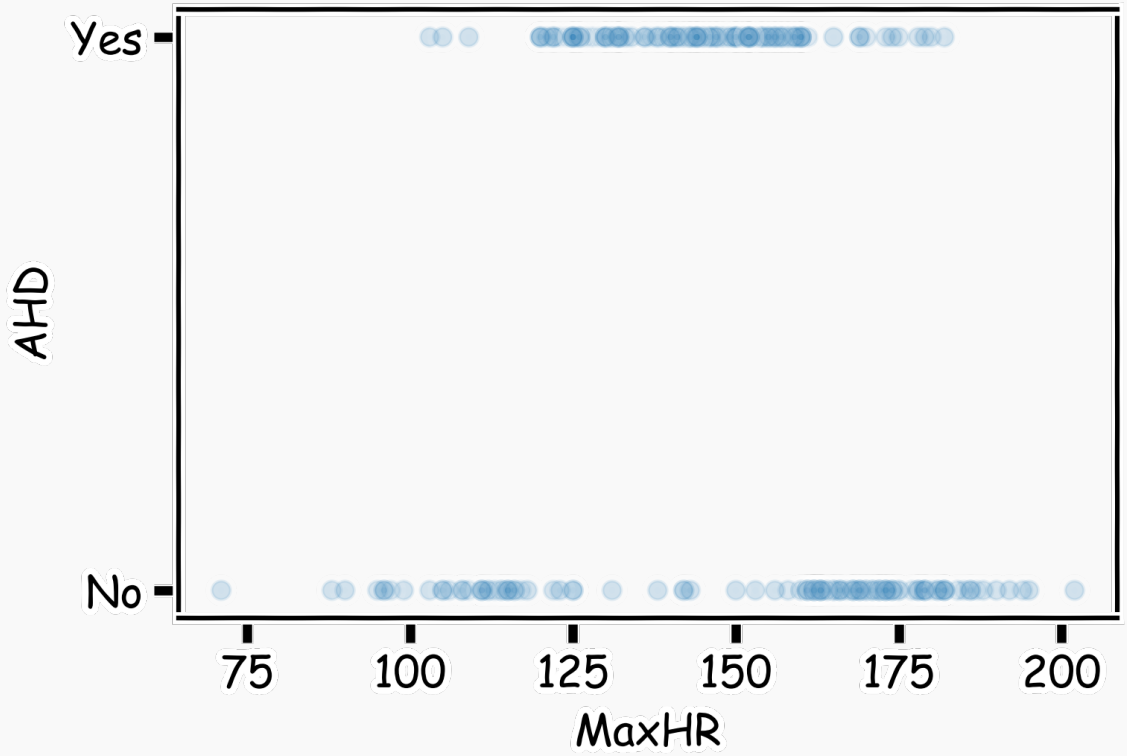
1. Introduction to Artificial Neural Networks
2. Review of Classification and Logistic Regression
3. Single Neuron Network (‘Perceptron’)
4. **Multi-Layer Perceptron (MLP)**

Example Using Heart Data

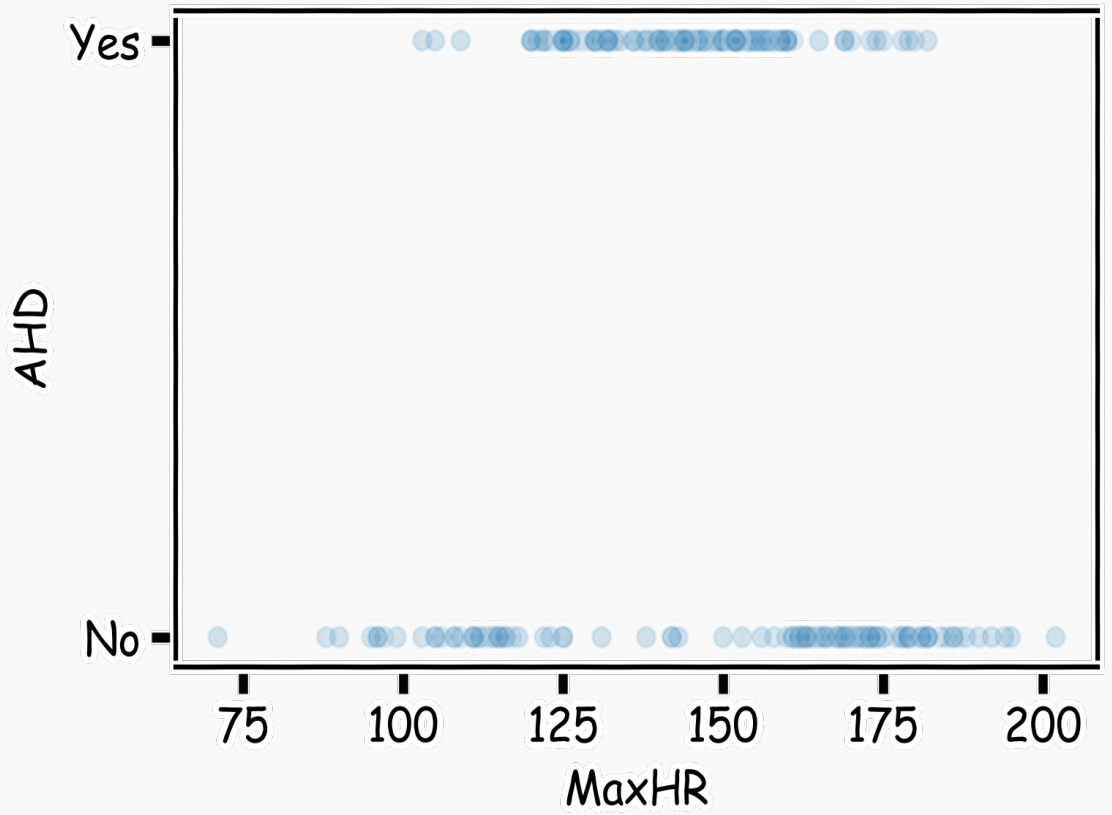
Slightly modified data to illustrate concepts.

Example Using Heart Data

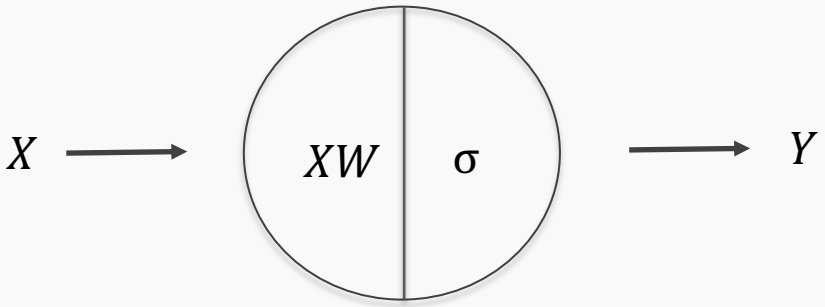
Slightly modified data to illustrate a point.



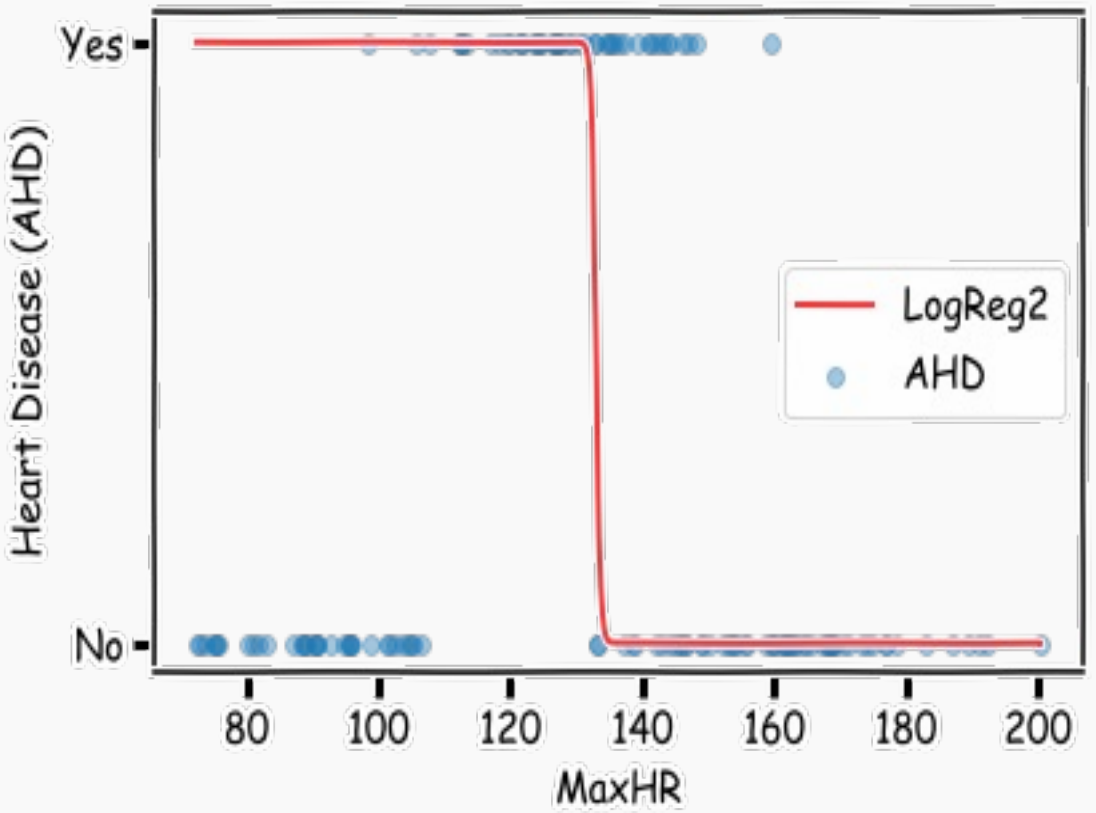
Example Using Heart Data



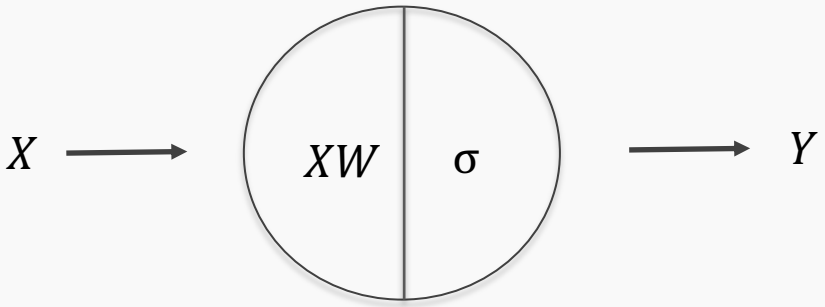
Example Using Heart Data



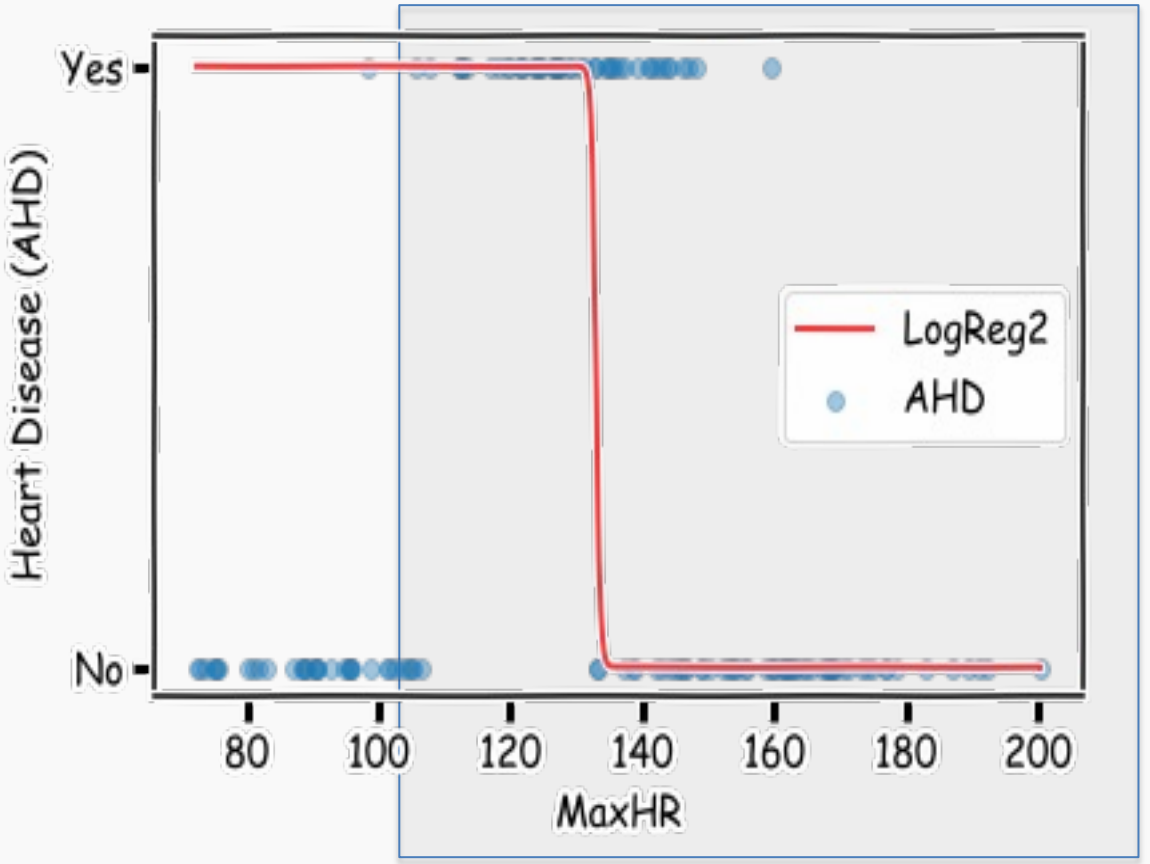
Choose W such as



Example Using Heart Data

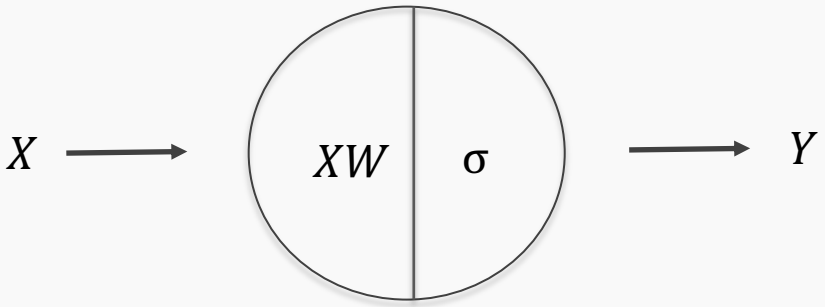


Choose W such as

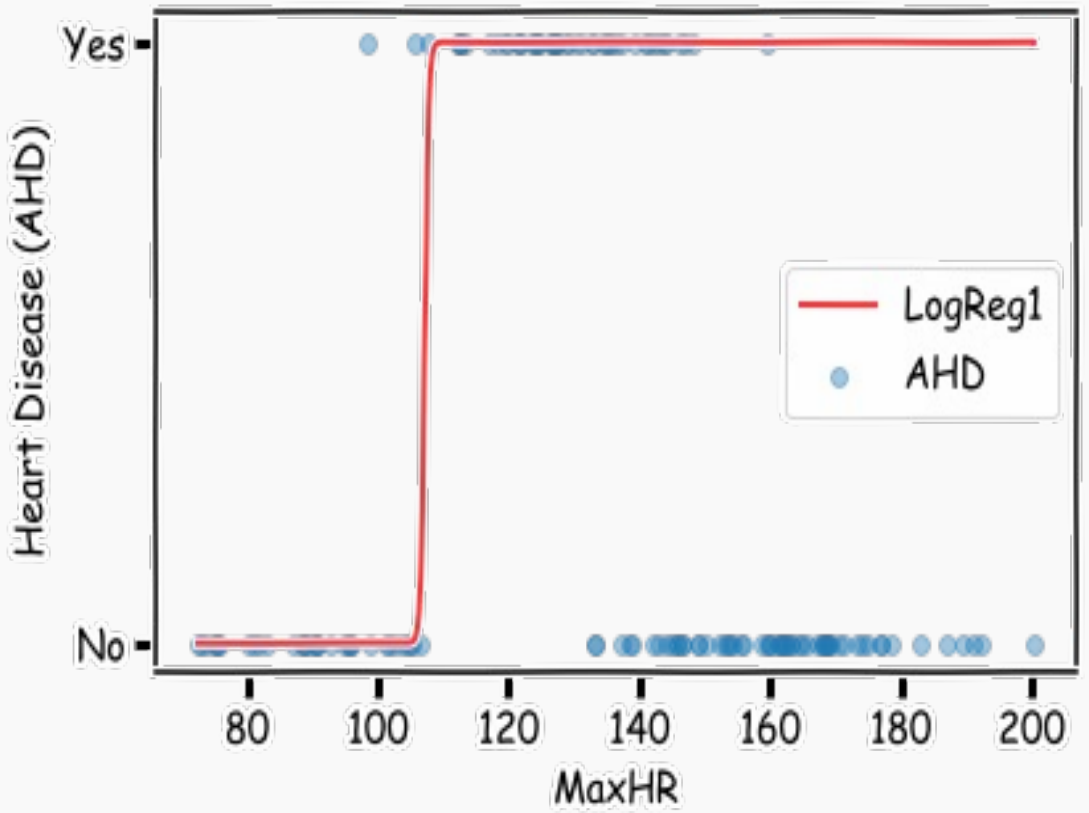


Right part of data is fitted well

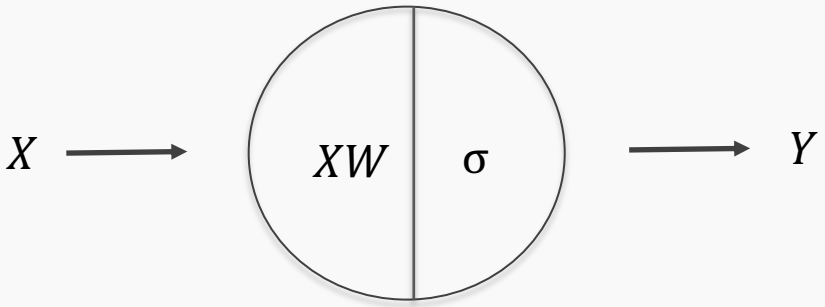
Example Using Heart Data



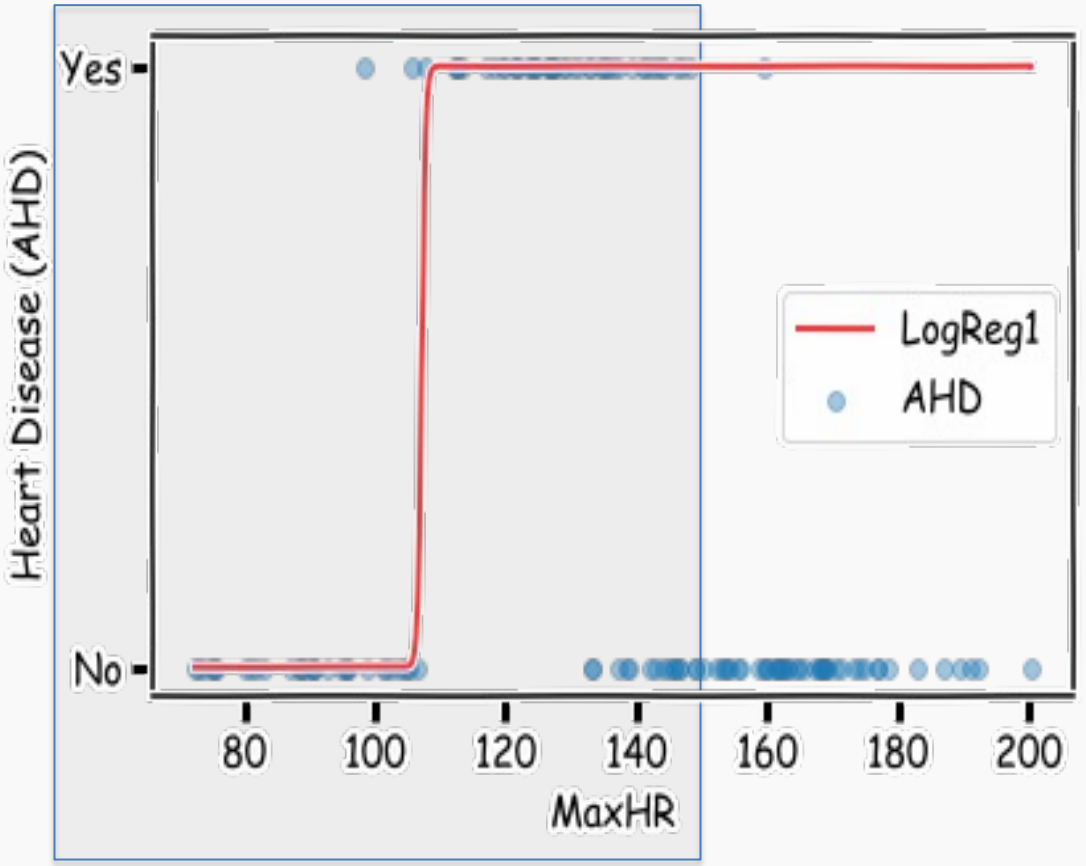
Choose W such as



Example Using Heart Data



Choose W such as

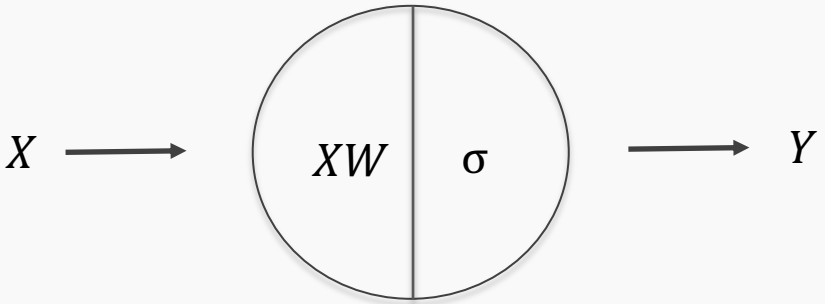


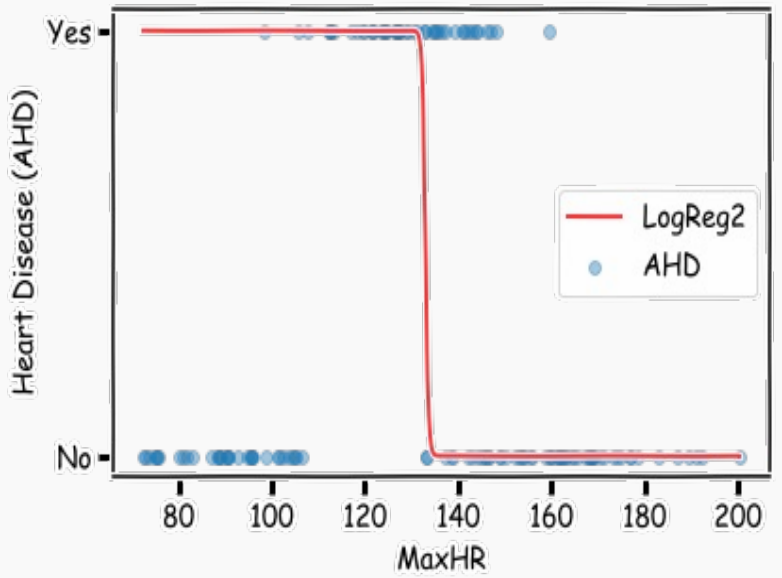
Left part of data is fitted well

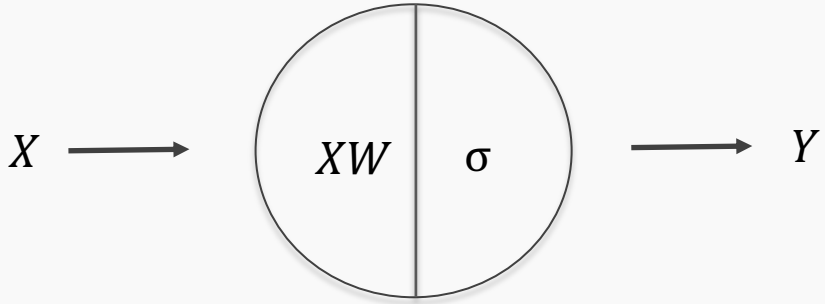
Example Using Heart Data

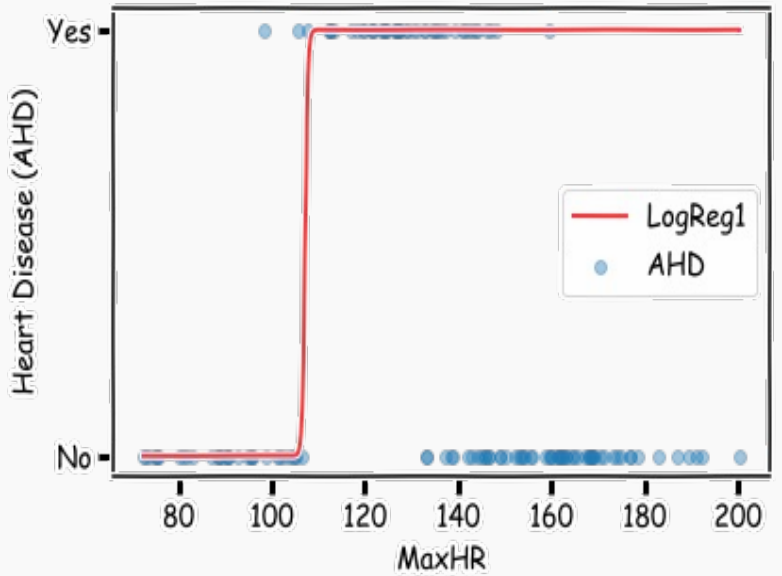
?

Two regions, two nodes

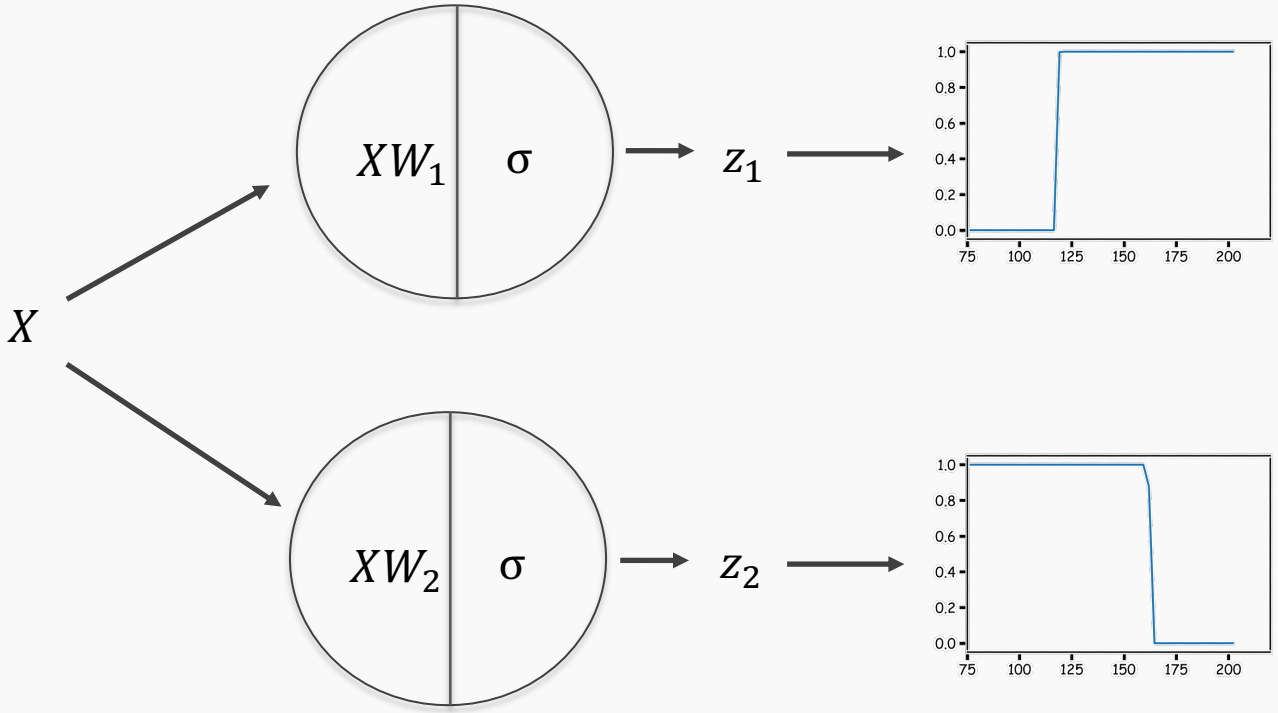




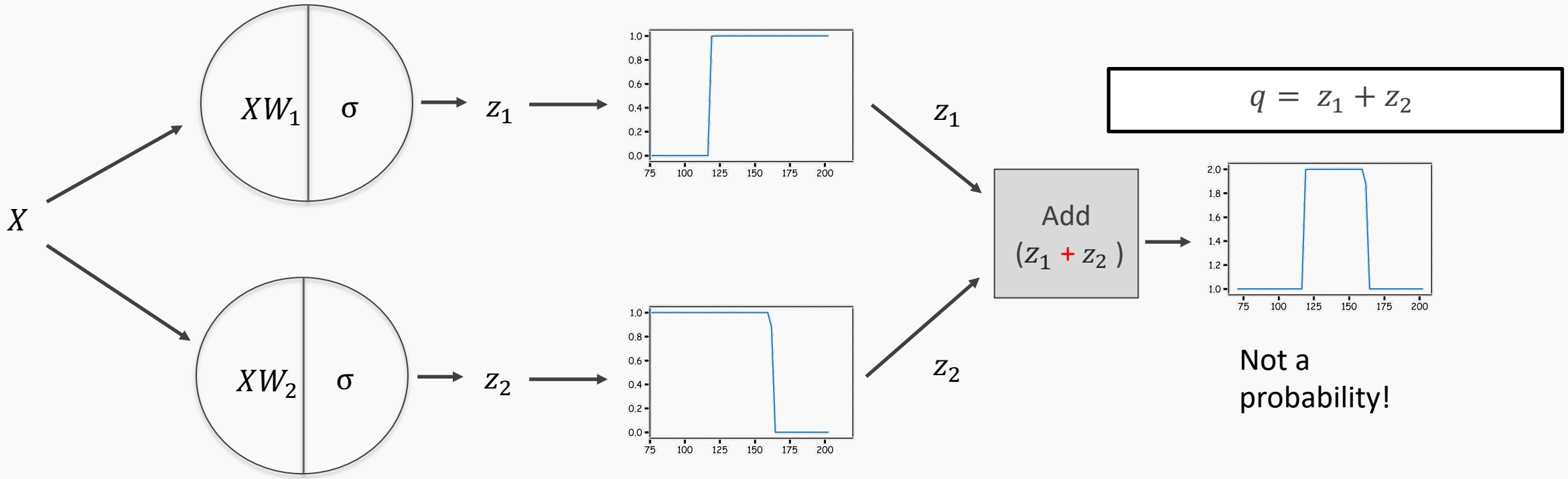




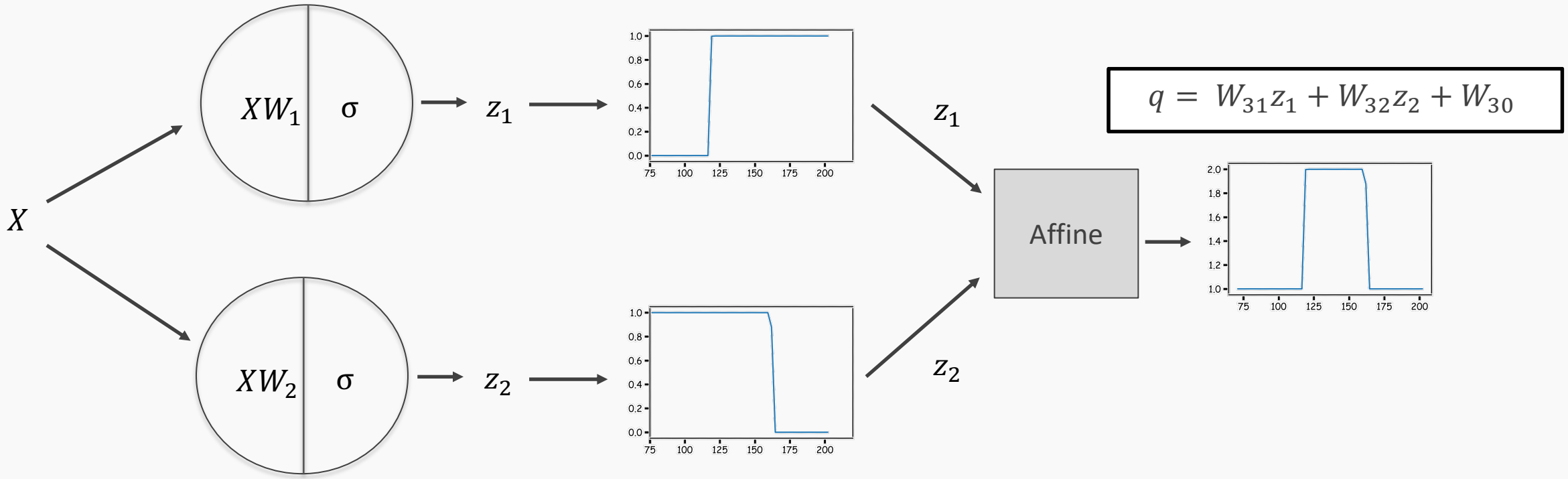
Combining Neurons



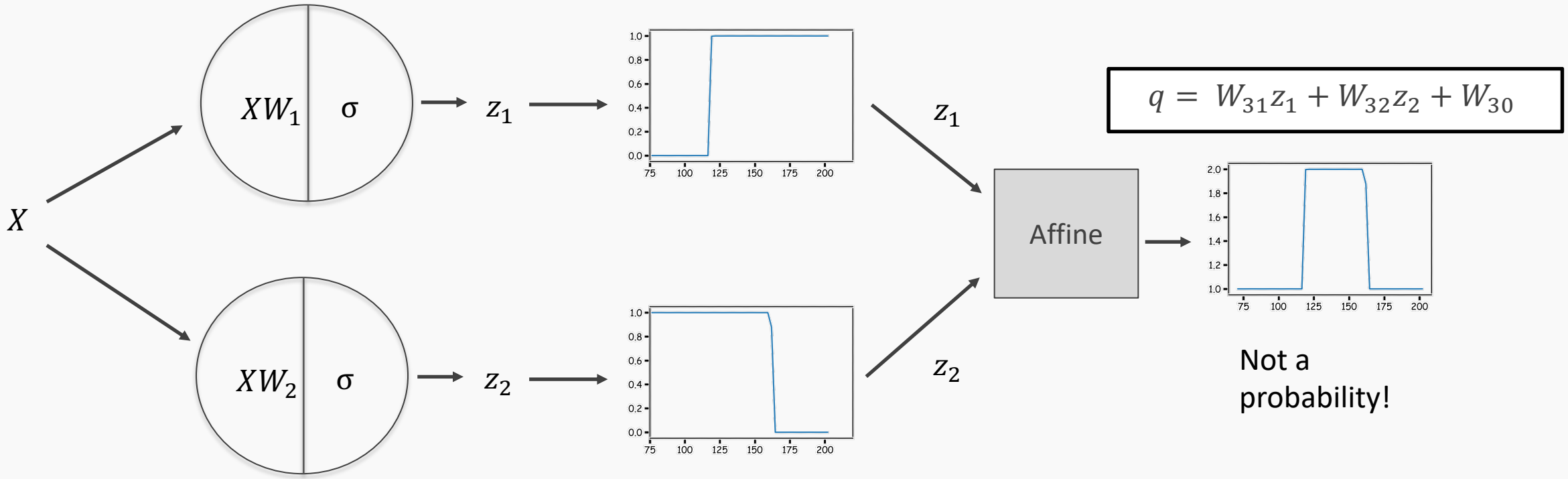
Combining Neurons



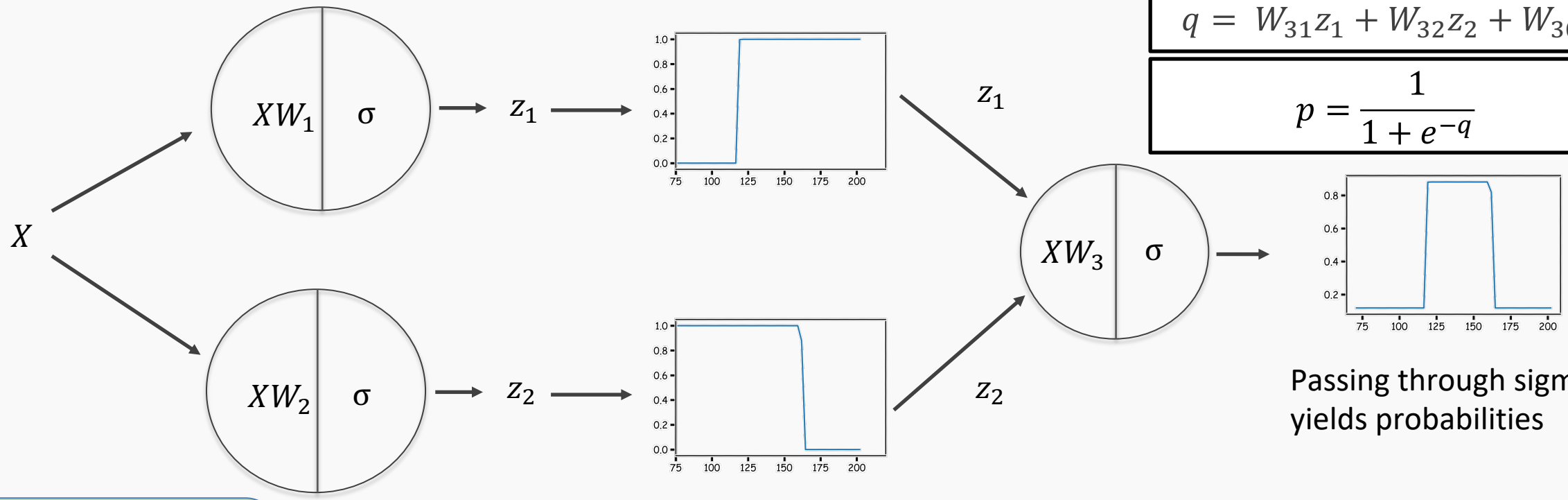
Combining Neurons



Combining Neurons



Combining Neurons



$$q = W_{31}z_1 + W_{32}z_2 + W_{30}$$

$$p = \frac{1}{1 + e^{-q}}$$

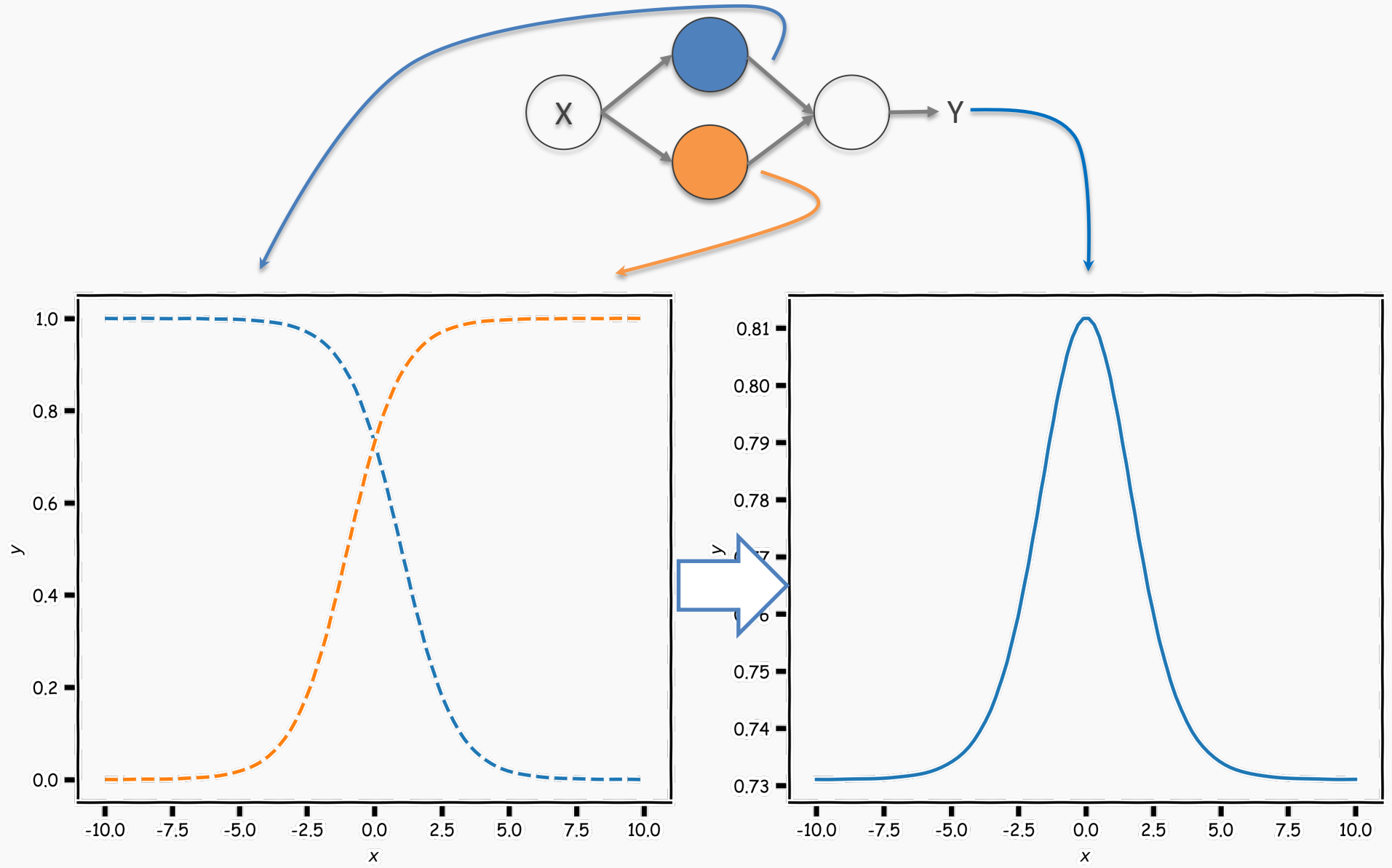
Passing through sigmoid yields probabilities

$$L = -y\ln(\mathrm{p}) - (1 - \mathrm{y})\ln(1 - \mathrm{p})$$

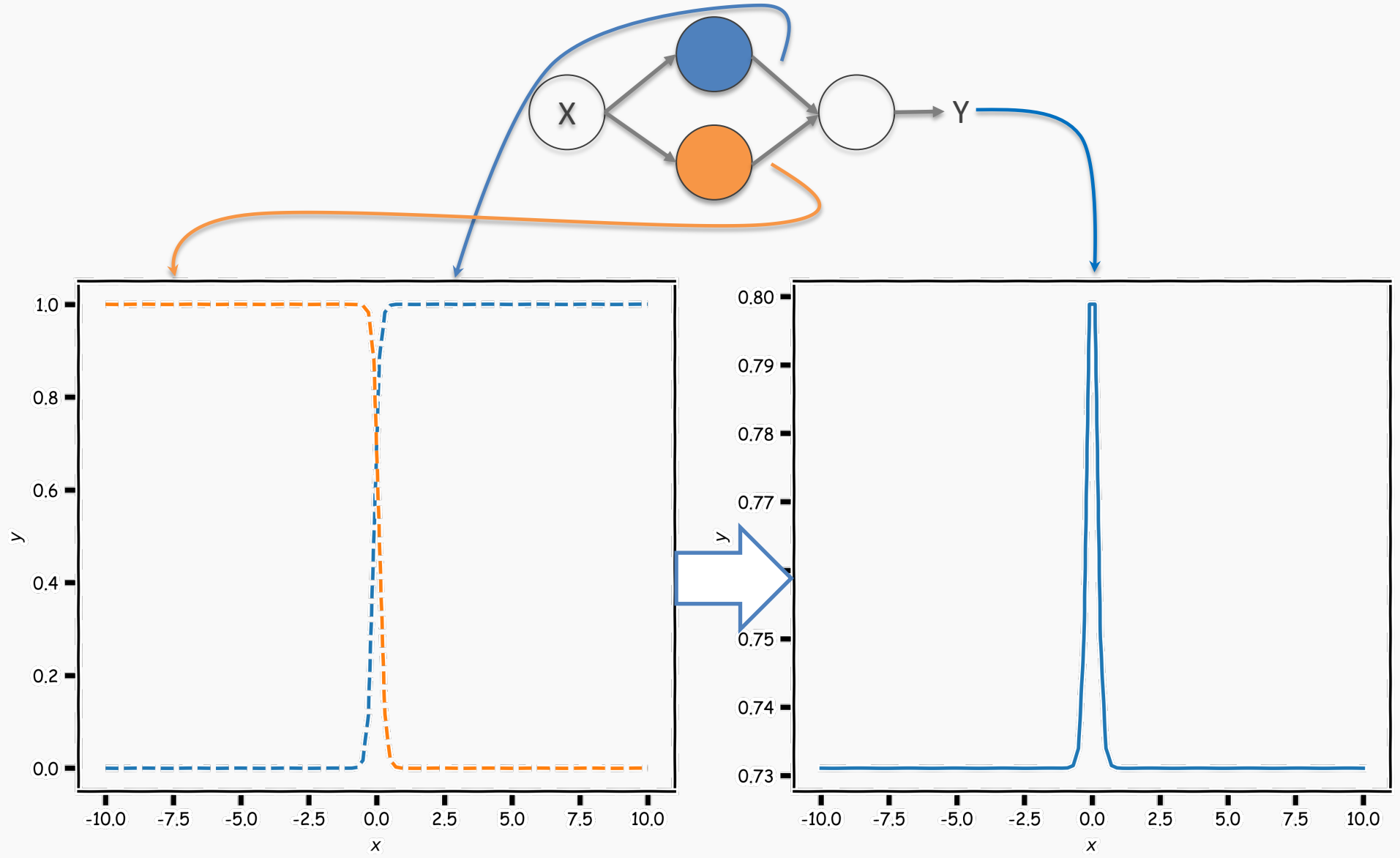
We will talk about about this in the coming lectures

Need to learn W_1, W_2 and W_3

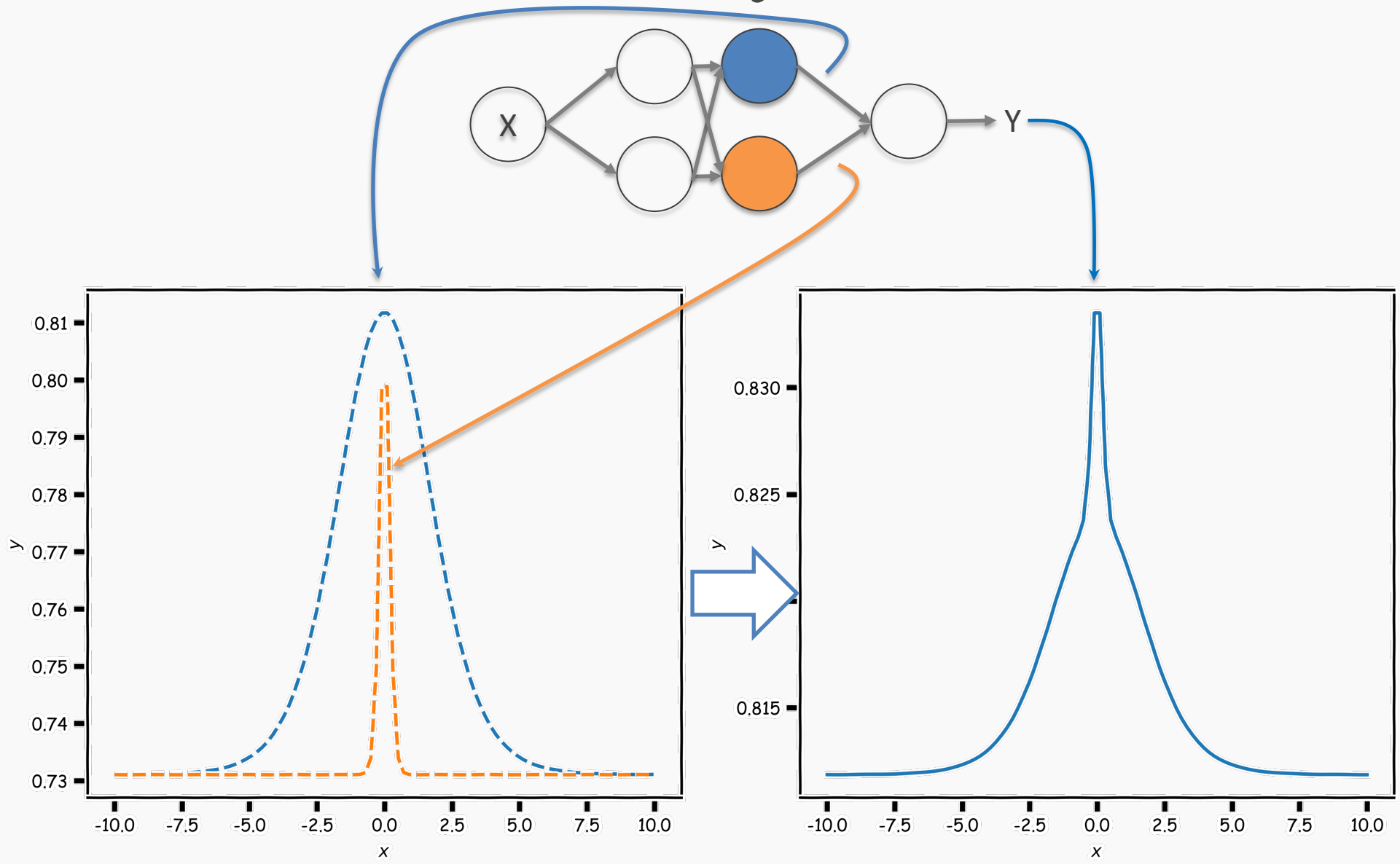
Combining neurons allows us to model interesting functions



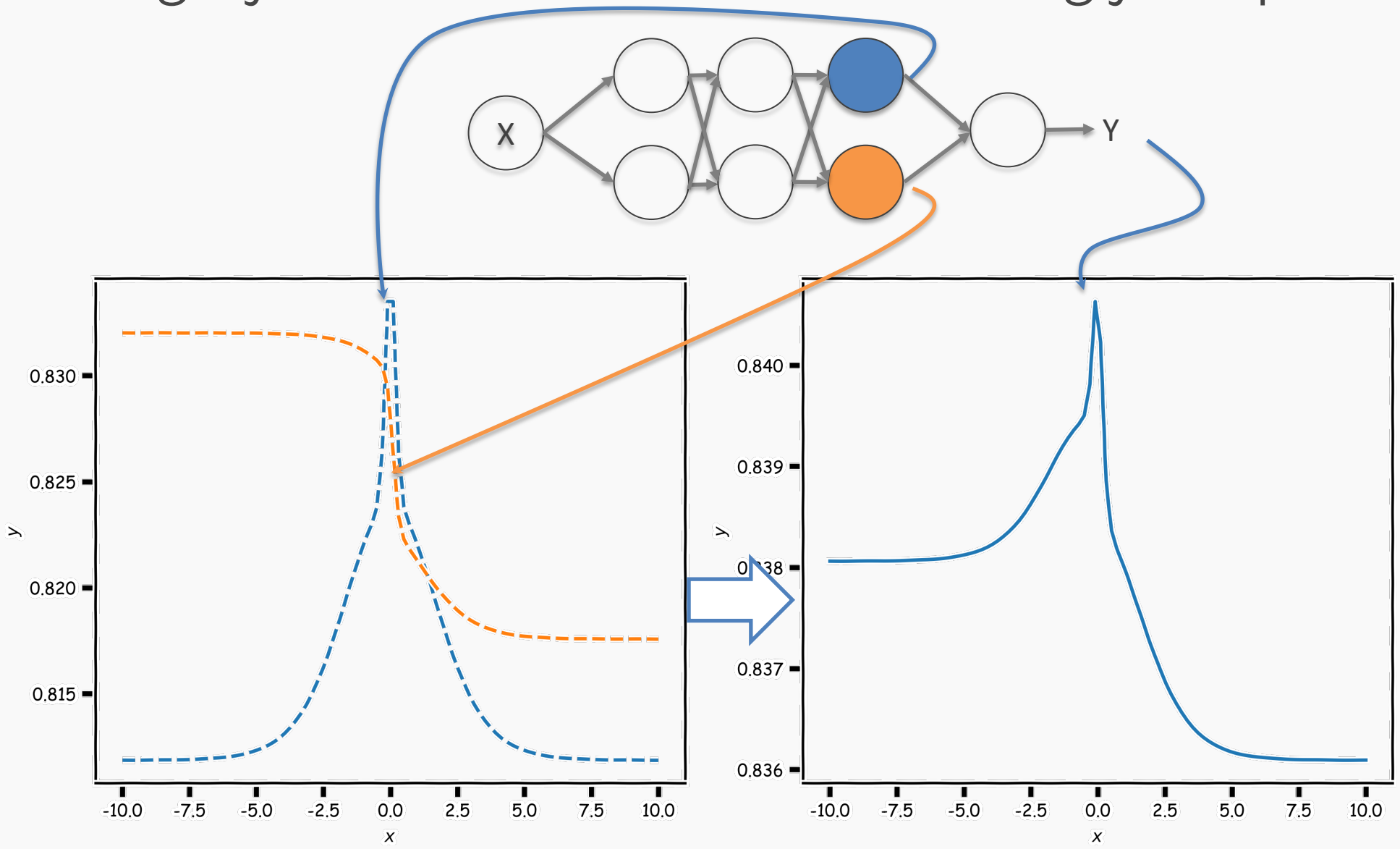
Different weights change the shape and position



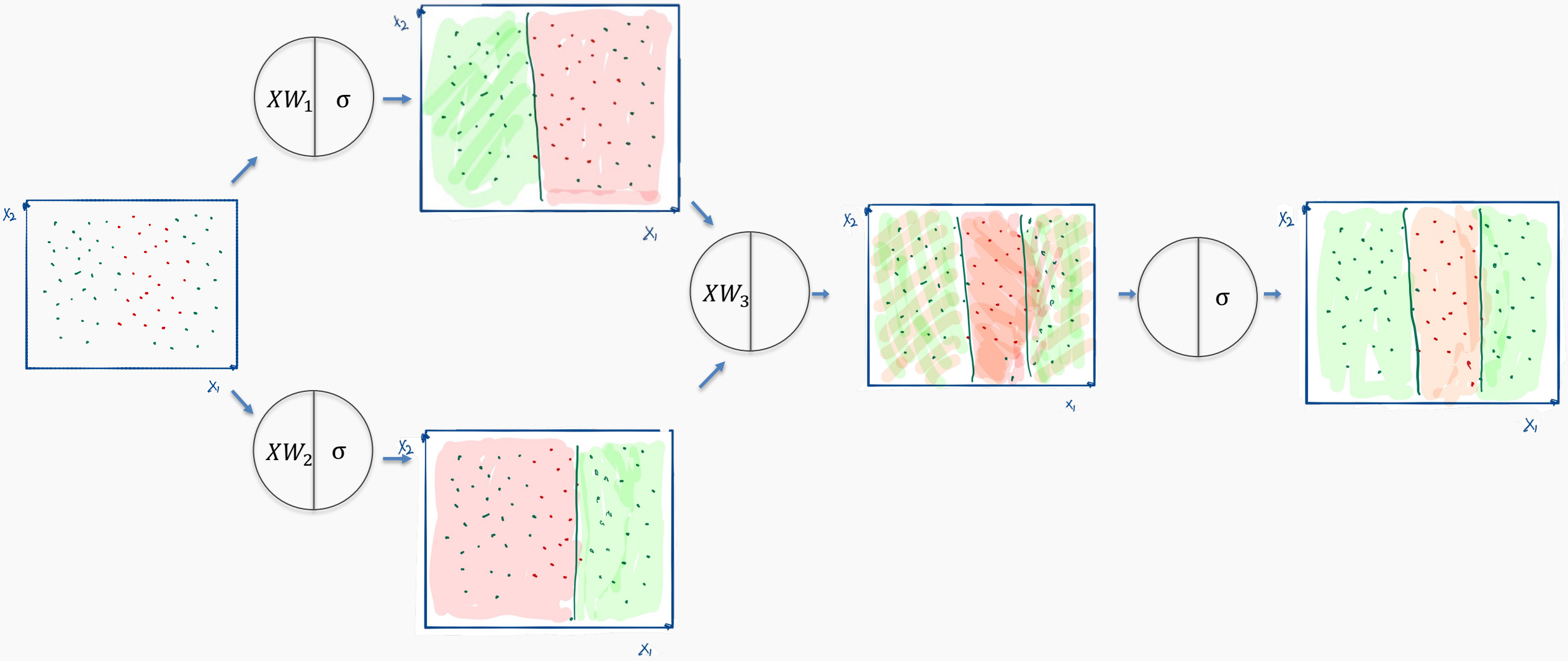
Neural networks can model *any* reasonable function



Adding layers allows us to model increasingly complex functions



For 2-D input the same idea applies.



Summary

So far:

- A single neuron can be a logistic regression or linear unit. We will soon see other choices of activation functions.
- A neural network is a combination of logistic regression (or other types) units.
- A neural network can approximate non-linear functions either for regression or classification.

Summary

So far:

- A single neuron can be a logistic regression or linear unit. We will soon see other choices of activation functions.
- A neural network is a combination of logistic regression (or other types) units.
- A neural network can approximate non-linear functions either for regression or classification.

Next:

- What kind of activations, how many neurons, how many layers, how to construct the output unit and what loss functions are appropriate?

Summary

So far:

- A single neuron can be a logistic regression or linear unit. We will soon see other choices.
- A neural network is a combination of logistic regression (or other types) units.
- A neural network can approximate non-linear functions either for regression or classification.

Next:

- What kind of activations, how many neurons, how many layers, how to construct the output unit and what loss functions are appropriate?

Following two lectures on NN:

- How do we estimate the weights and biases?
- How to regularize Neural Networks?