Classroom **PROCEDURES**

- Get vaccinated
- Provide vaccination proof
- Do the daily COVID screen
- Don't attend when ill
- Wear a mask
- Leave room promptly
- Wash hands frequently
- Don't consume drinks/food

QUartsci.com/Fall2021

CISC452/CMPE452/COGS400

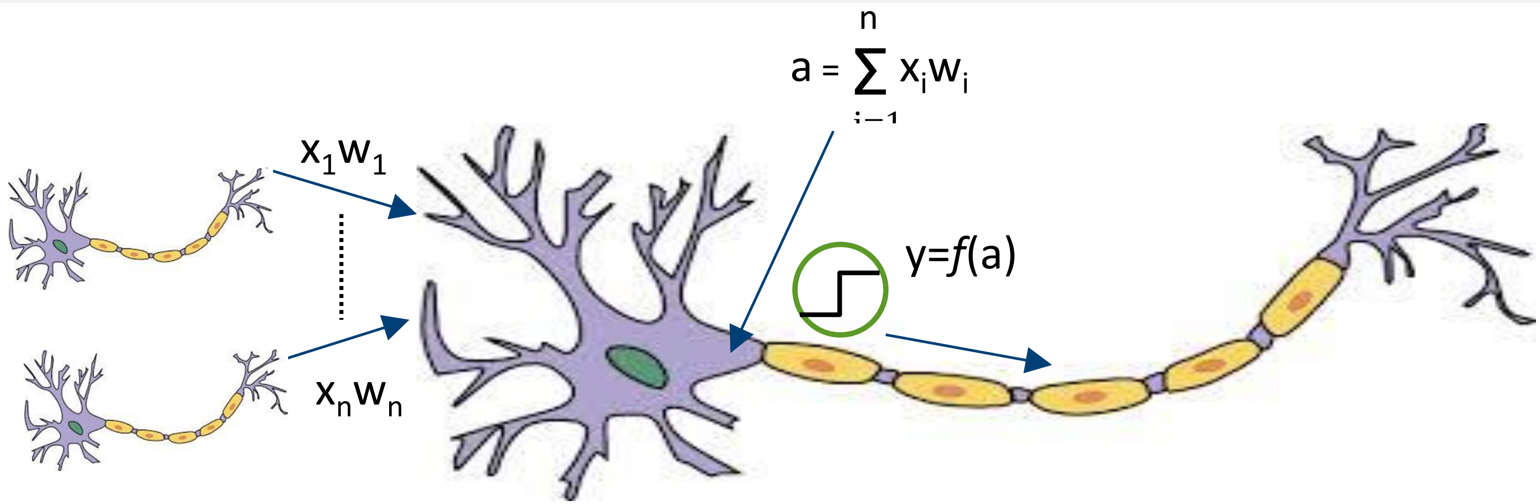
Artificial Neural Networks and Linear Separability

Farhana Zulkernine

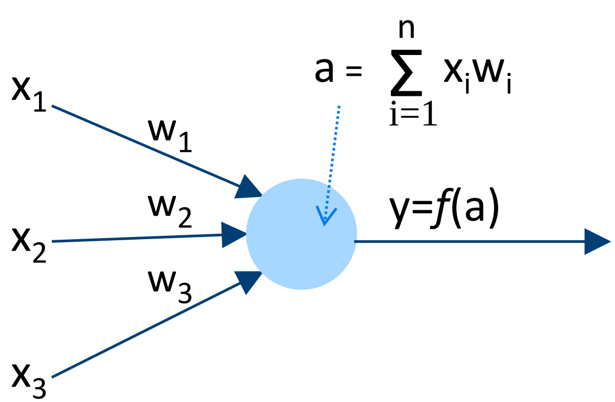
ANN Modeling

- A signal received through the dendrites due to transformation at the synapse is modeled using **connection weights**.
- Total **activation** inside the cell is modeled using the sum of incoming signals.
- The signal generated and transmitted through the axon as a result of the action potential is modeled using an **activation function**.
- Complete information flow is modeled using a **network of connected neurons**.

ANN Model



McCulloch and Pitts Neuron Model



The weights w_i take on *real values* $w_i \in \mathbb{R}$

Activation is the weighted sum of all incoming potentials.

f(a) can be any function that generates a spike (high value) at a given threshold value θ to mimic the scenario of *Action Potential*.

The Activation Function

One possible choice is a threshold function:

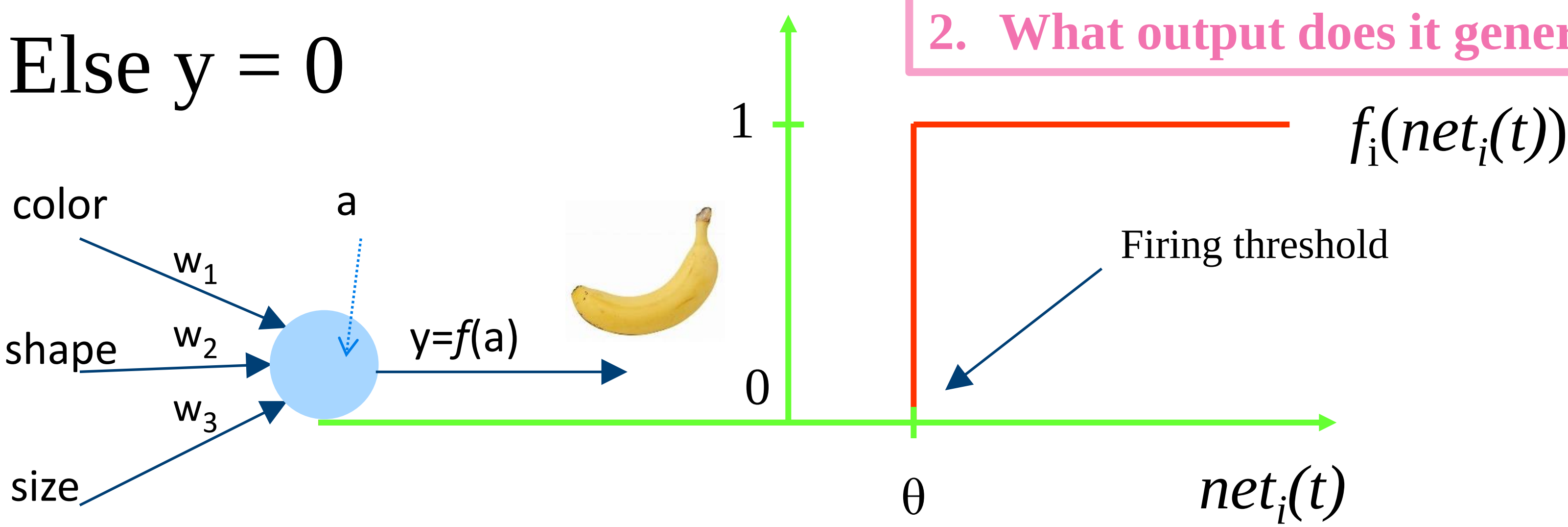
Therefore, we call this a threshold neuron.

$$f_i(\mathrm{net}_i(t)) = 1, \quad \text{if} \quad \mathrm{net}_i(t) \geq \theta$$
$$= 0, \quad \text{otherwise}$$

Recognize: Banana y = 1

Else y = 0

1. **When does it fire?**
2. **What output does it generate?**

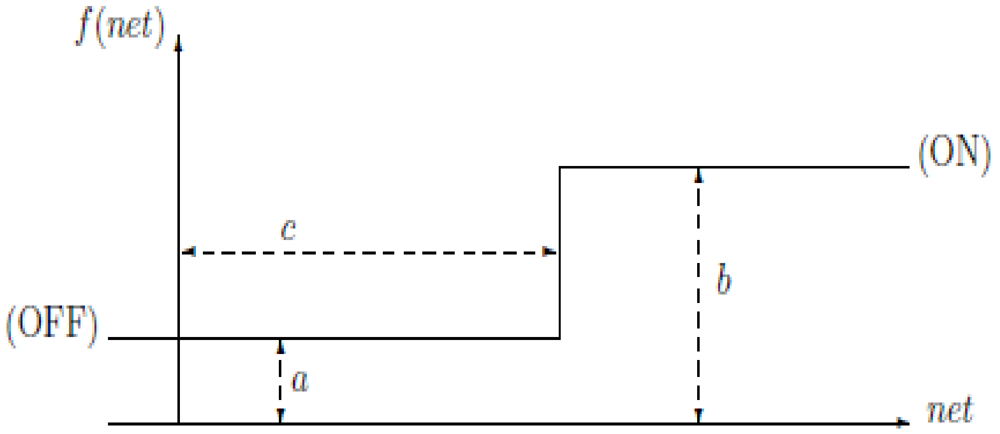


Step Function

- Simplest function that captures the idea of a "firing threshold"
- Can be used as a class identifier
- Problem: Very small change in $net_i(t)$ can cause a spike and hence change the output

f(net) = a if *net* < c

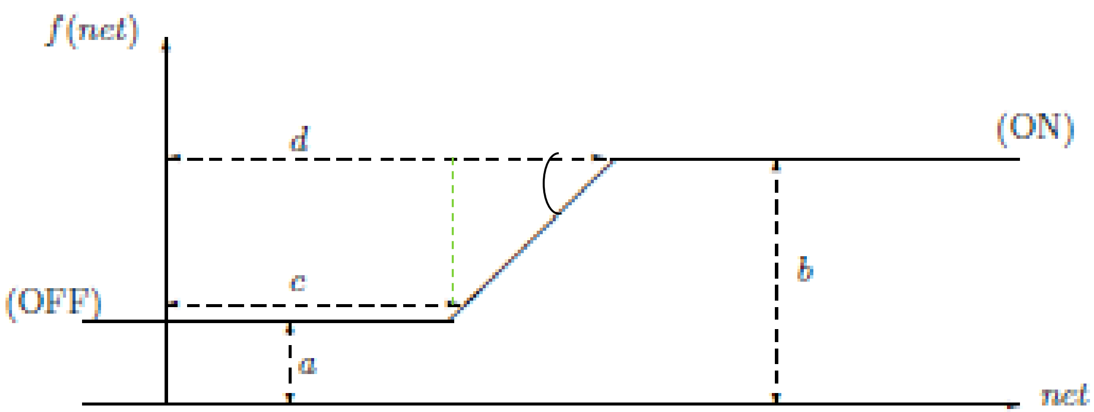
= b if *net* >= c



Ramp Function

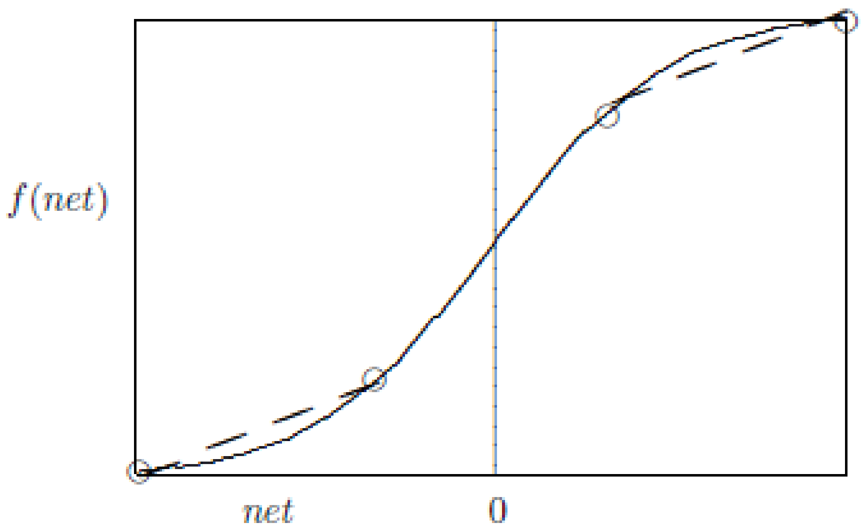
- The ramp function is continuous and almost everywhere differentiable in contrast to the simple ON/OFF description of the output.

$$f(net) = \begin{cases} a & \text{if } net \leq c \\ b & \text{if } net \geq d \\ a + \frac{(net-c)(b-a)}{d-c} & \text{otherwise} \end{cases}$$



Piecewise Linear Functions

- Consist of finite number of linear segments, and are thus differentiable almost everywhere.
- Easier to compute than general nonlinear functions such as sigmoid functions.
- Can be used to avoid sudden change in output like the step function (from 0 to 1).



Activation Functions

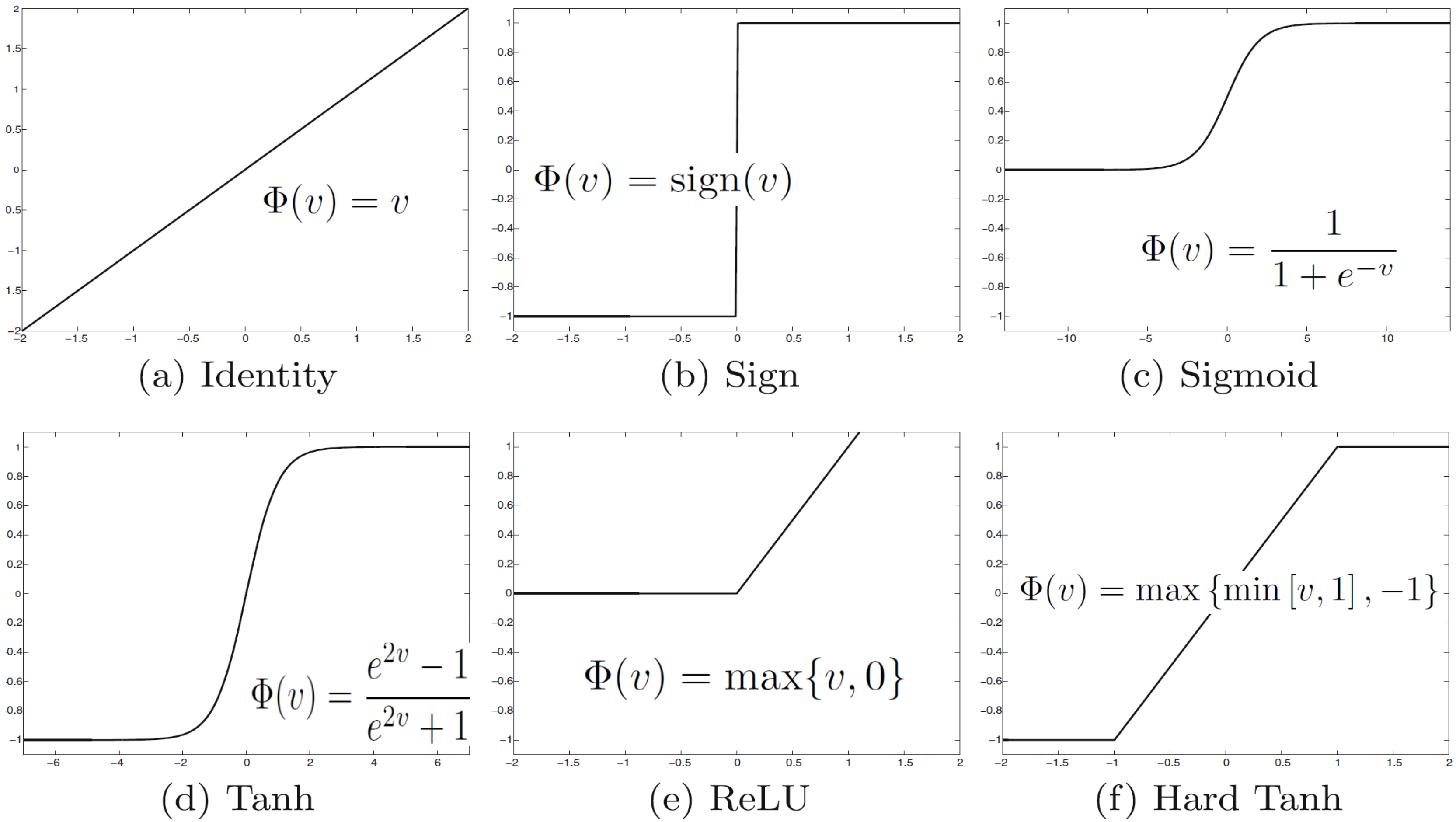
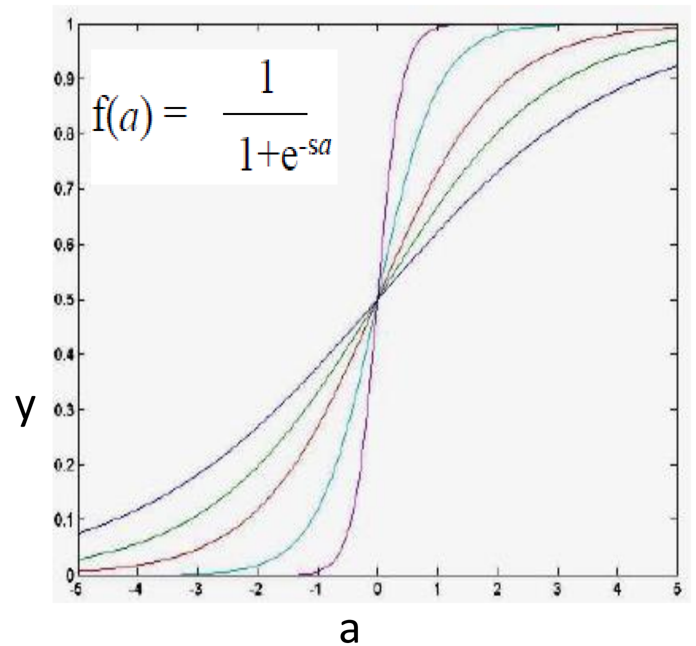


Figure 1.8: Various activation functions

Reference: Page 13 Book of Charu Aggarwal

Sigmoid Function

- These functions are continuous and differentiable everywhere, and asymptotically approach saturation values (0 and 1 as shown in the picture)
- The parameter *s* controls the slope of the sigmoid function. The greater the value of *s*, the steeper the curve will be.



$$\lim_{net \to +\infty} f(net) = 1$$

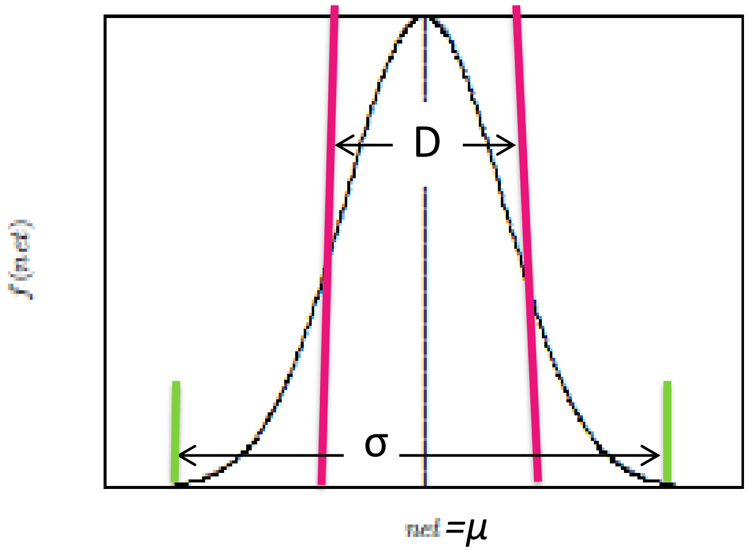
$$\lim_{net \to -\infty} f(net) = 0$$

Gaussian Functions

- Continuous bell-shaped functions.
- Also called 'radial-basis' function.
- *f*(net) asymptotically approaches 0 (or some constant) for large magnitudes of *net*, with a single maximum for

  $net = \mu$, say $\mu = 0$.

  Greater σ → wider curve.



$$f(net) = \frac{1}{\sqrt{2\pi}\sigma} exp[-\frac{1}{2}(\frac{net-\mu}{\sigma})^2]$$

Application

- Classification – Grouping and recognizing the label attached to the group
  - Label is known
  - Identify letters as ‘A’ and ‘B’ which denote class labels
- Clustering – Grouping based on distinctive and common features
  - Label is not known
  - Identify both as letters but different based on shape, name of the letter is not known, define central properties
- So we need logic implementation to divide the data into categories and groups

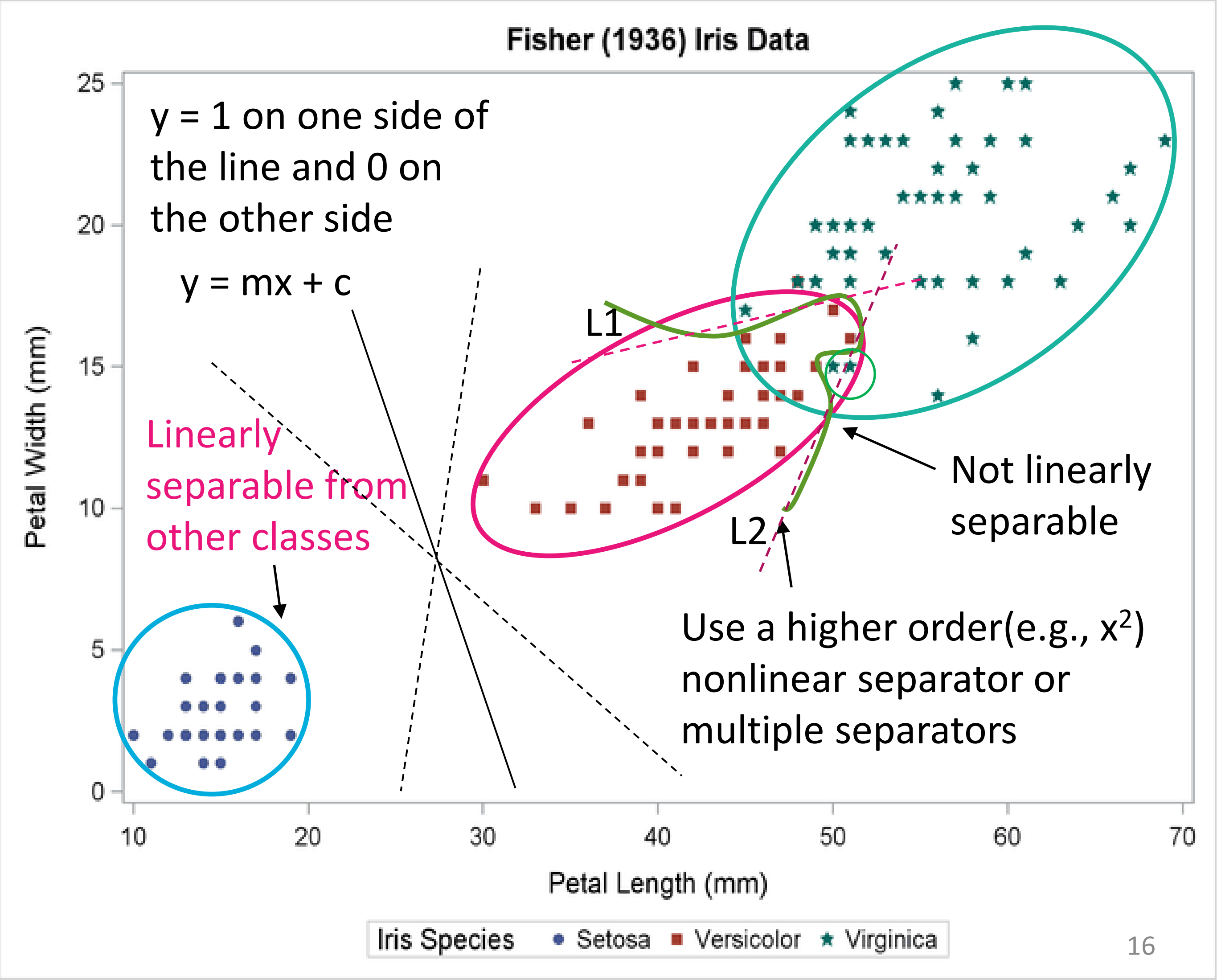
  if …. Then … category [class label] or group

Iris Dataset

- Iris is a flowering plant with 260-300 different species. One of the most popular datasets in machine learning.
- Contains measurements of 3 different species of Iris: Setosa, Versicolour and Virginia.
- The machine learning problem requires creating a model that can learn (be trained) to recognize the **class label** from the features of the flower.

Separability

- 4 features are given
  - Sepal length
  - Sepal width
  - Petal length
  - Petal width
- Only two features are used here.
- *x* represents the data point, *y* is the output of the model
- *m* is learned – can have different values



Summary

- ANNs model biological neurons, the structure and the functionality to create machine intelligence.
- Variation in input is modeled using weights associated with incoming links.
- Activation inside the neuron is modeled using weighted sum of inputs.
- Activation function is the output function that transforms the activation to an output value and can vary based on model requirements.