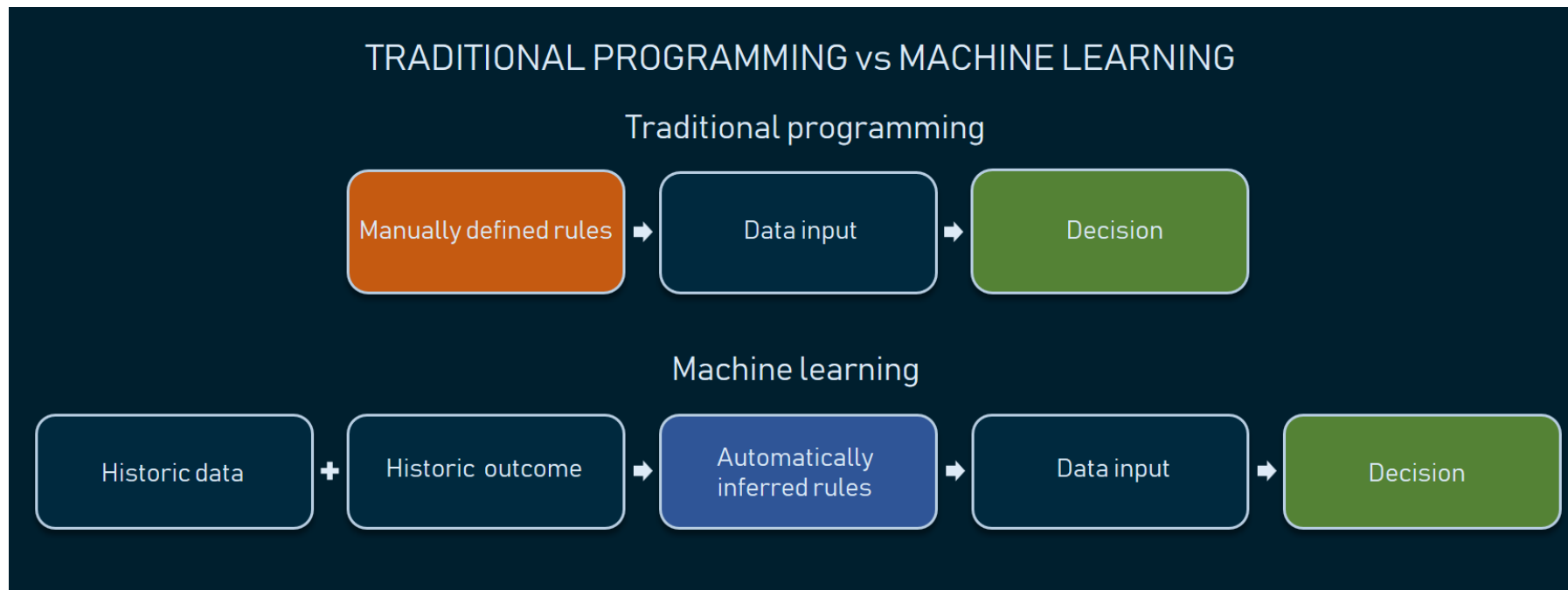# Neural Network in One Hour

Ravi Shankar

1

Traditional Programming
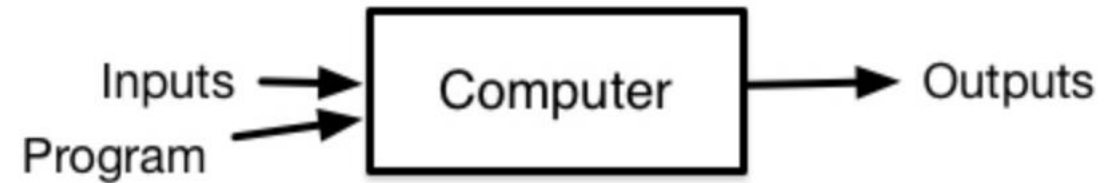

TRADITIONAL PROGRAMMING vs MACHINE LEARNING

Traditional programming

Manually defined rules → Data input → Decision

Machine learning

Historic data + Historic outcome → Automatically inferred rules → Data input → Decision

Traditional Programming vs Machine Learning

▶ If (a > b) return a; else b

▶ Look at the inputs and outputs

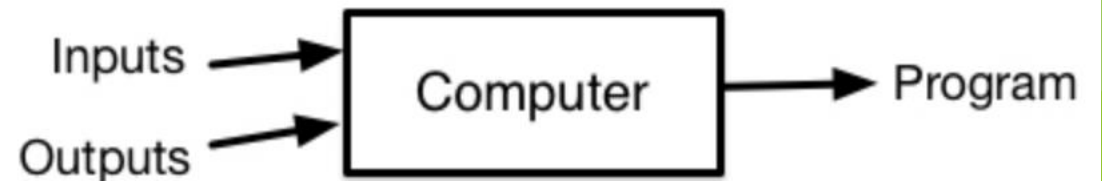| a	b	output
10	5	10
4	8	8
4	4	4
6	2	?

Traditional Programming

Inputs → Computer → Outputs
Program →

Machine Learning

Inputs → Computer → Program
Outputs →

MODEL !!

Inputs

"Hidden" Layer(s)

Output Layer
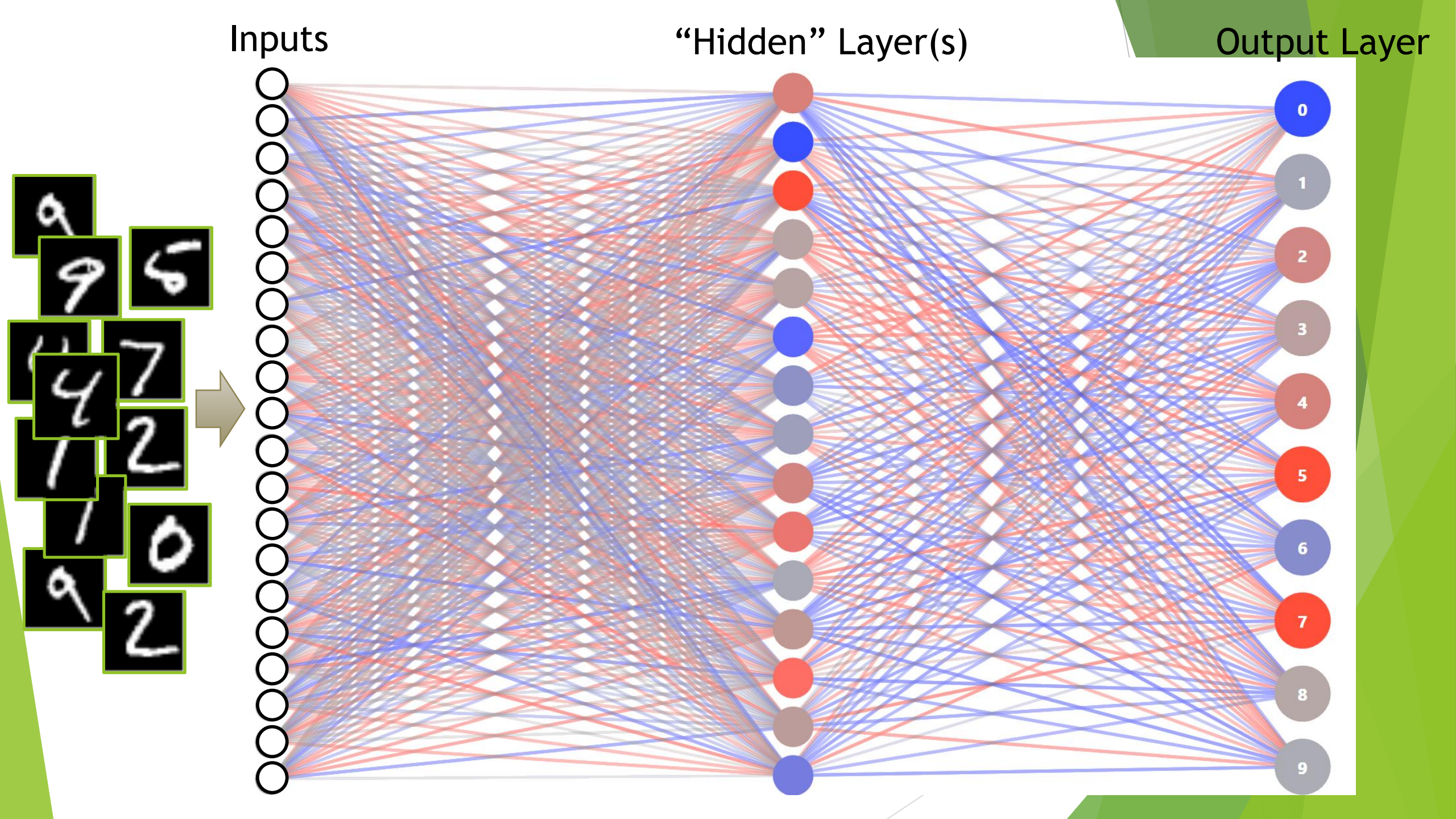
Epoch

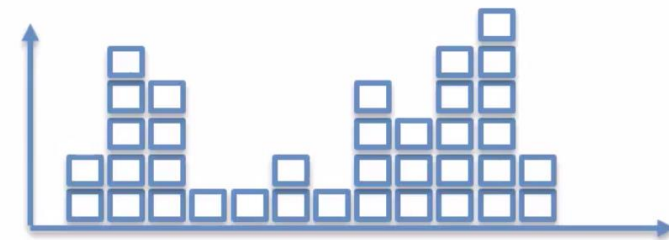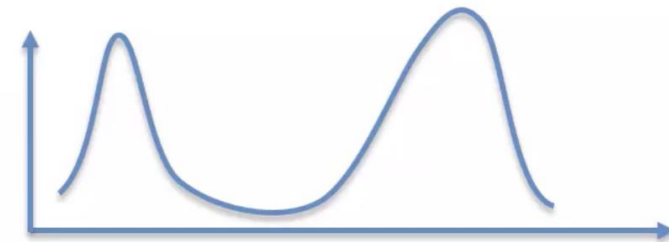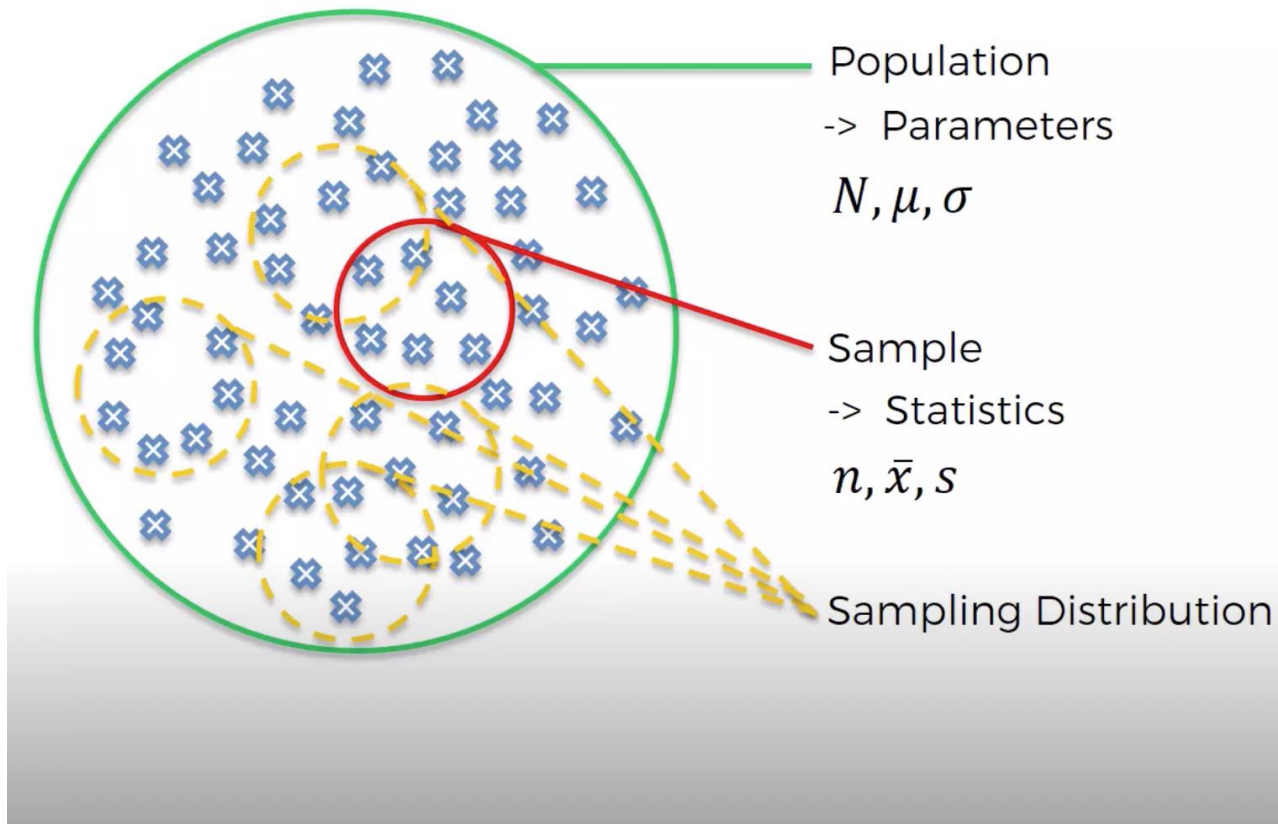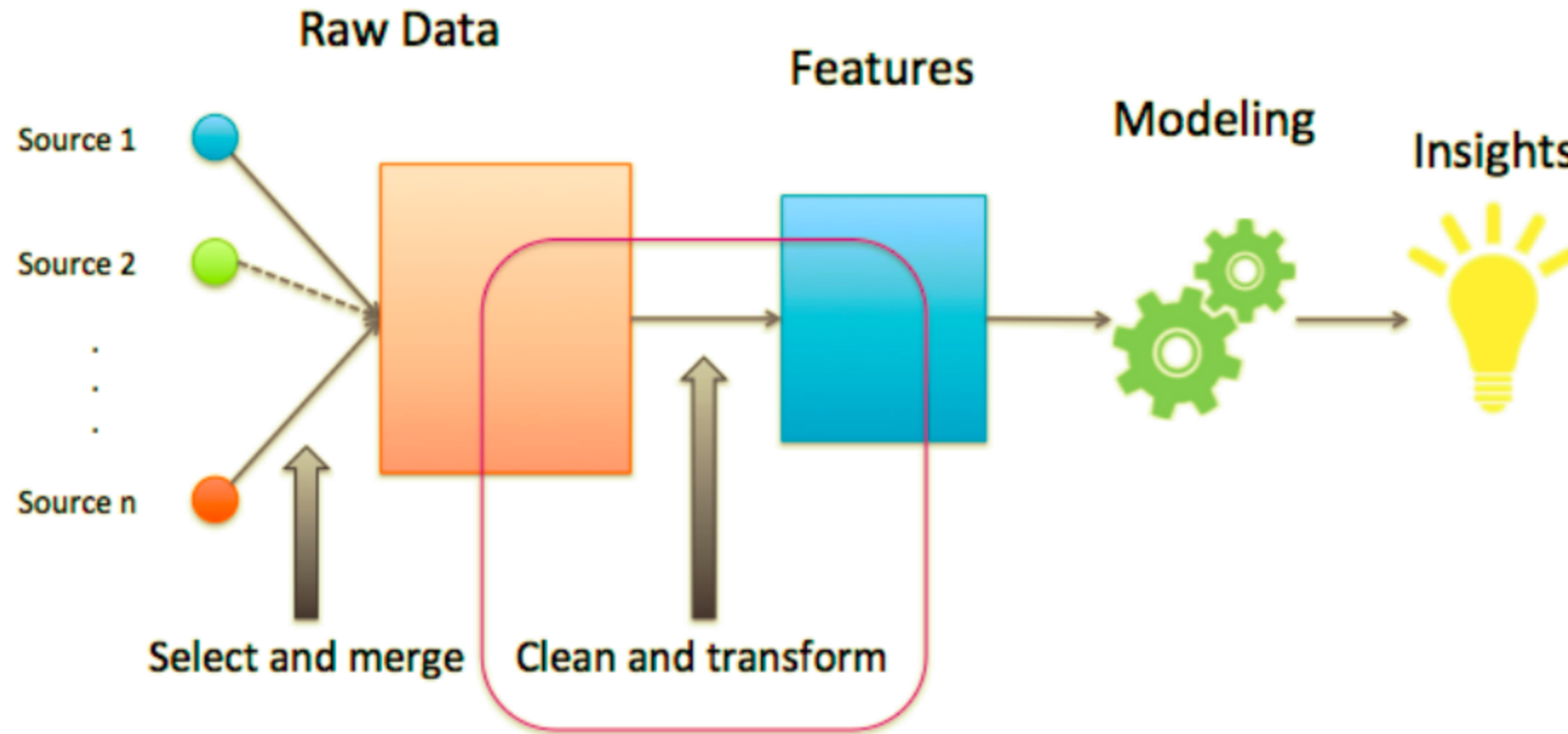The number of epochs is a hyperparameter that defines the number times that the learning algorithm will work through the entire training dataset.

Population and Samples

Population
-> Parameters
N, μ, σ

Sample
-> Statistics
$n, \bar{x}, s$

Sampling Distribution

Central Limit Theorem   http://onlinestatbook.com/stat_sim/sampling_dist/
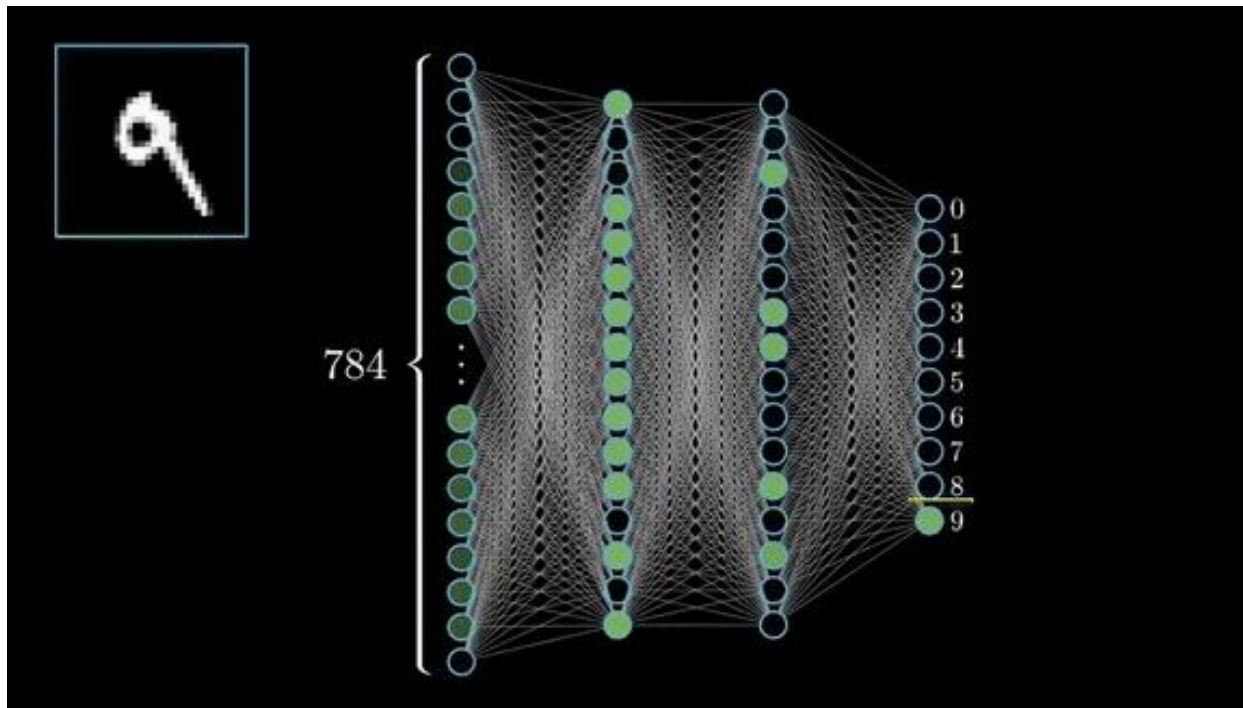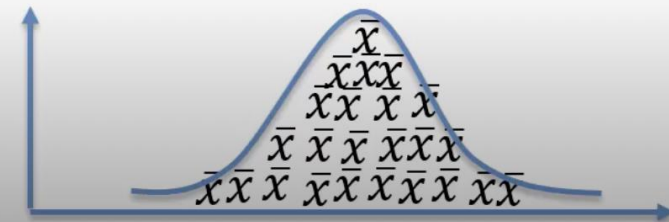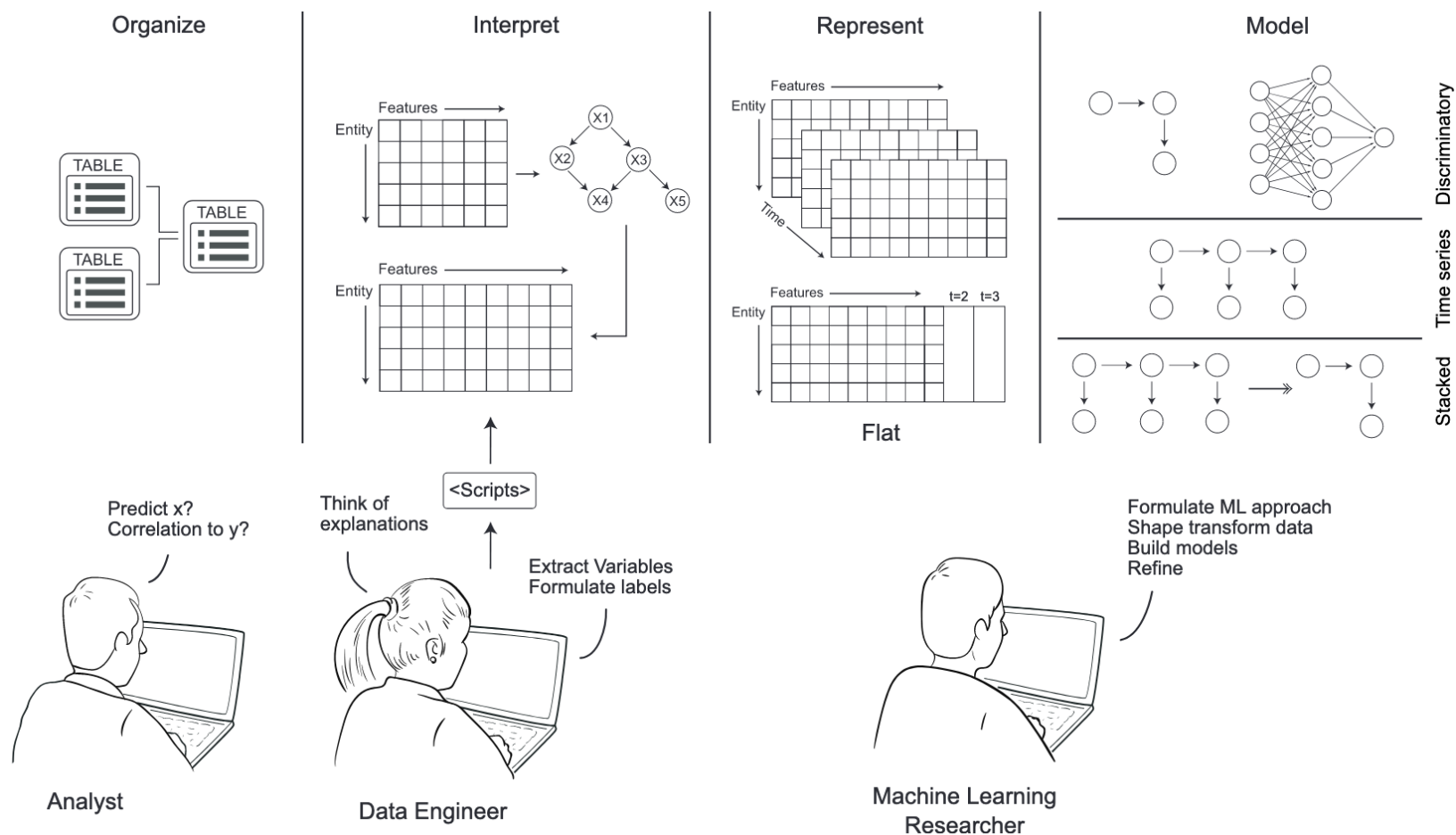
Ideal Input Identification

Deep Feature Synthesis (DFS)

You can automate data science endeavors – AutoML plays less role, but auto featuring is amazing

Deep Feature Synthesis (DFS)

Deep Feature Synthesis algorithm, which is capable of generating features that express a rich feature-space. We can develop an end-to-end Data Science Machine which can:
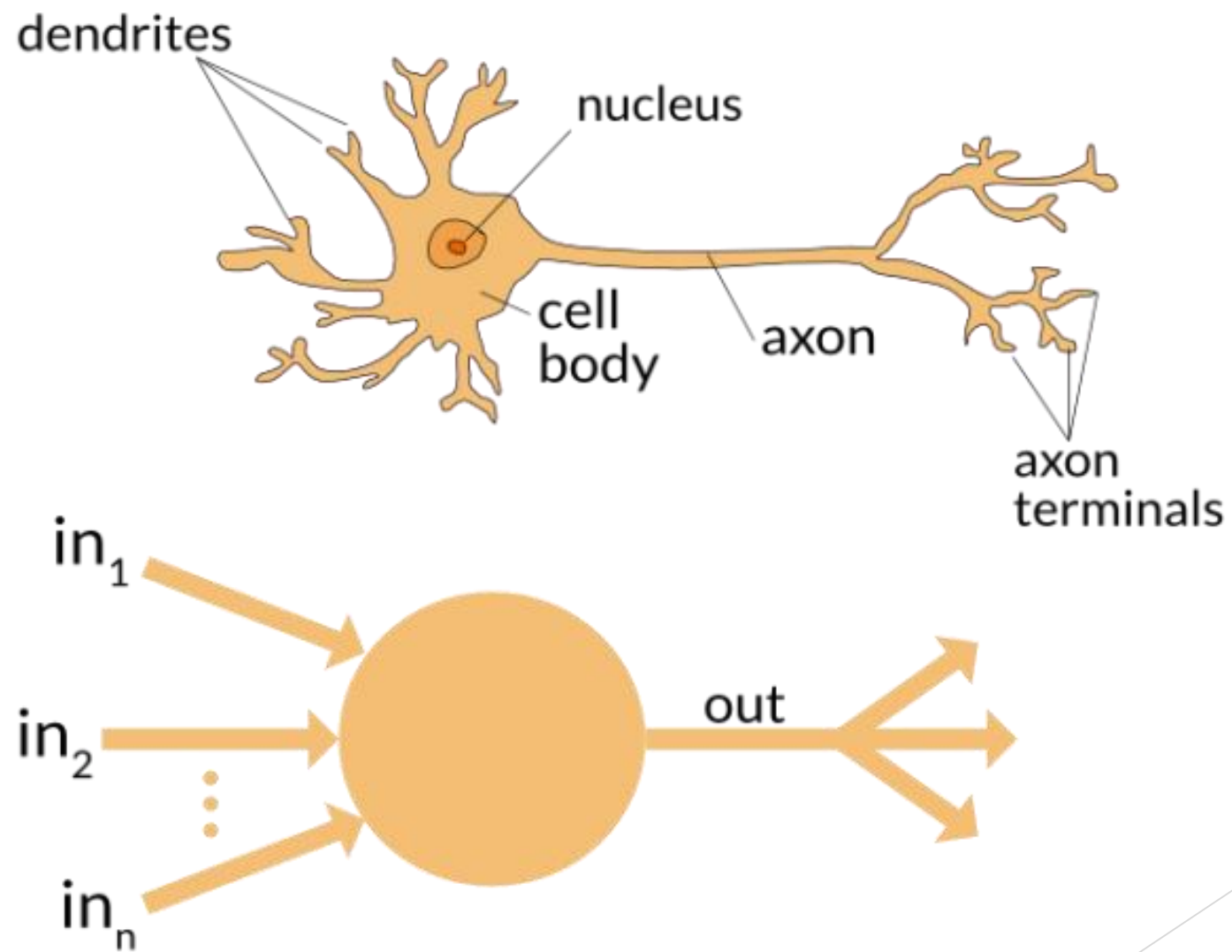
(a) automatically generate features via Deep Feature Synthesis

(b) autotune a machine learning pathway to extract the most value out of the synthesized features

(c) produce submissions for online data science competitions. Matched human level performance when competing in data science competitions using the Data Science Machine

http://www.jmaxkanter.com/static/papers/DSAA_DSM_2015.pdf
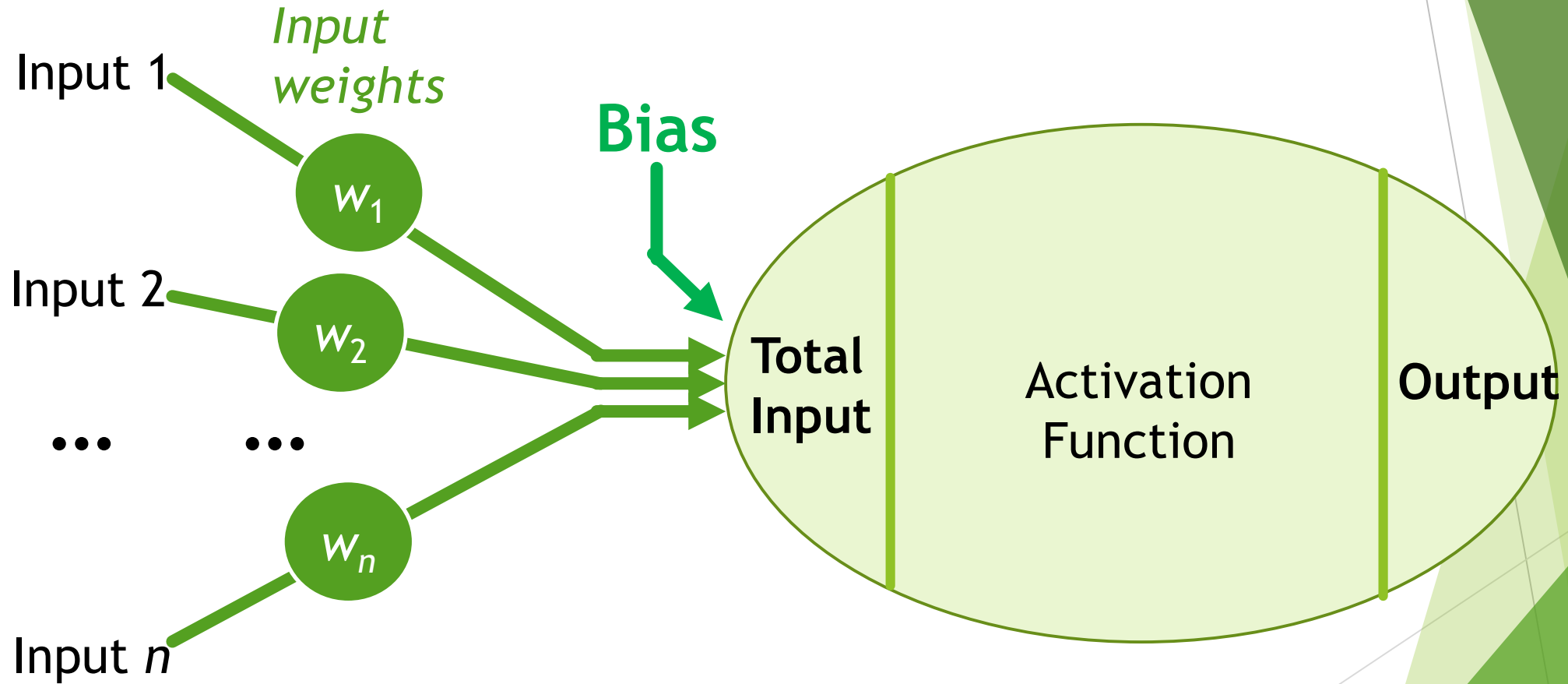
The Best Places to End Up

- Many options available

  - TensorFlow

  - Microsoft Cognitive Toolkit

  - Keras

  - Pytorch

  - Caffe

  - Anaconda

Inspired by Human Brain

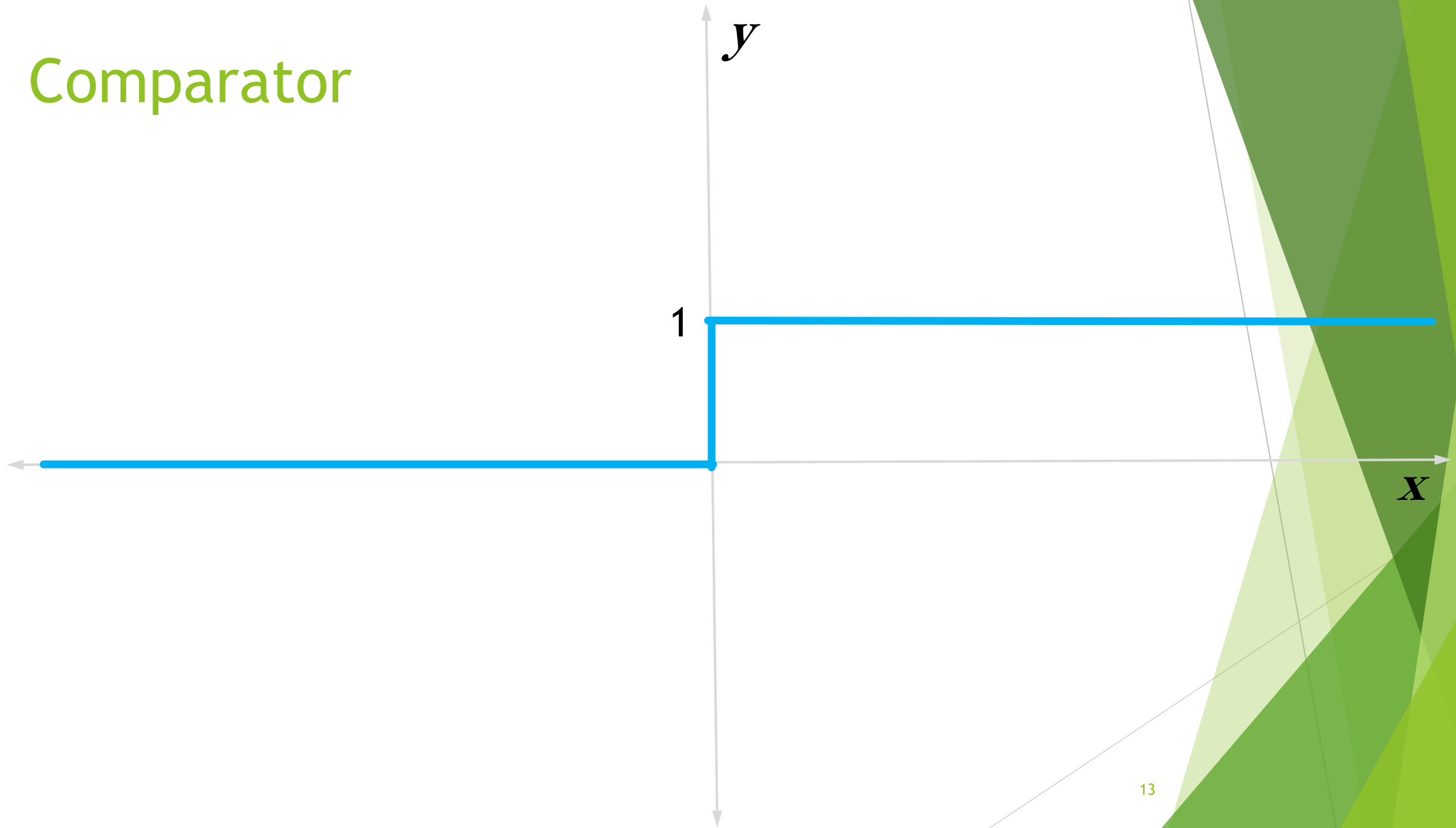# The AI Neuron

Input 1

Input weights

w_1

Input 2

w_2

... ...

Input n

w_n

Bias

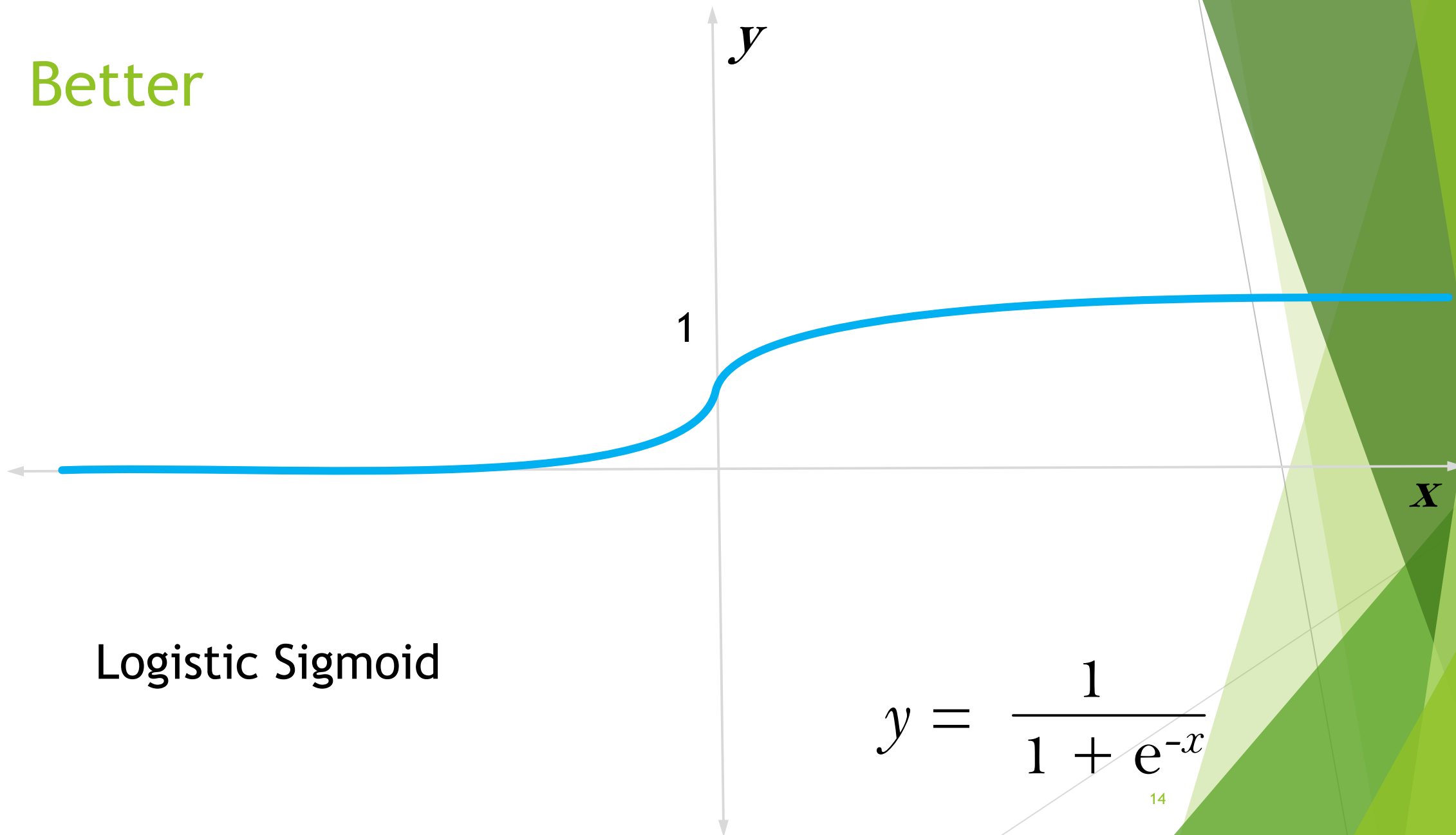**Total Input**

Activation Function

Output
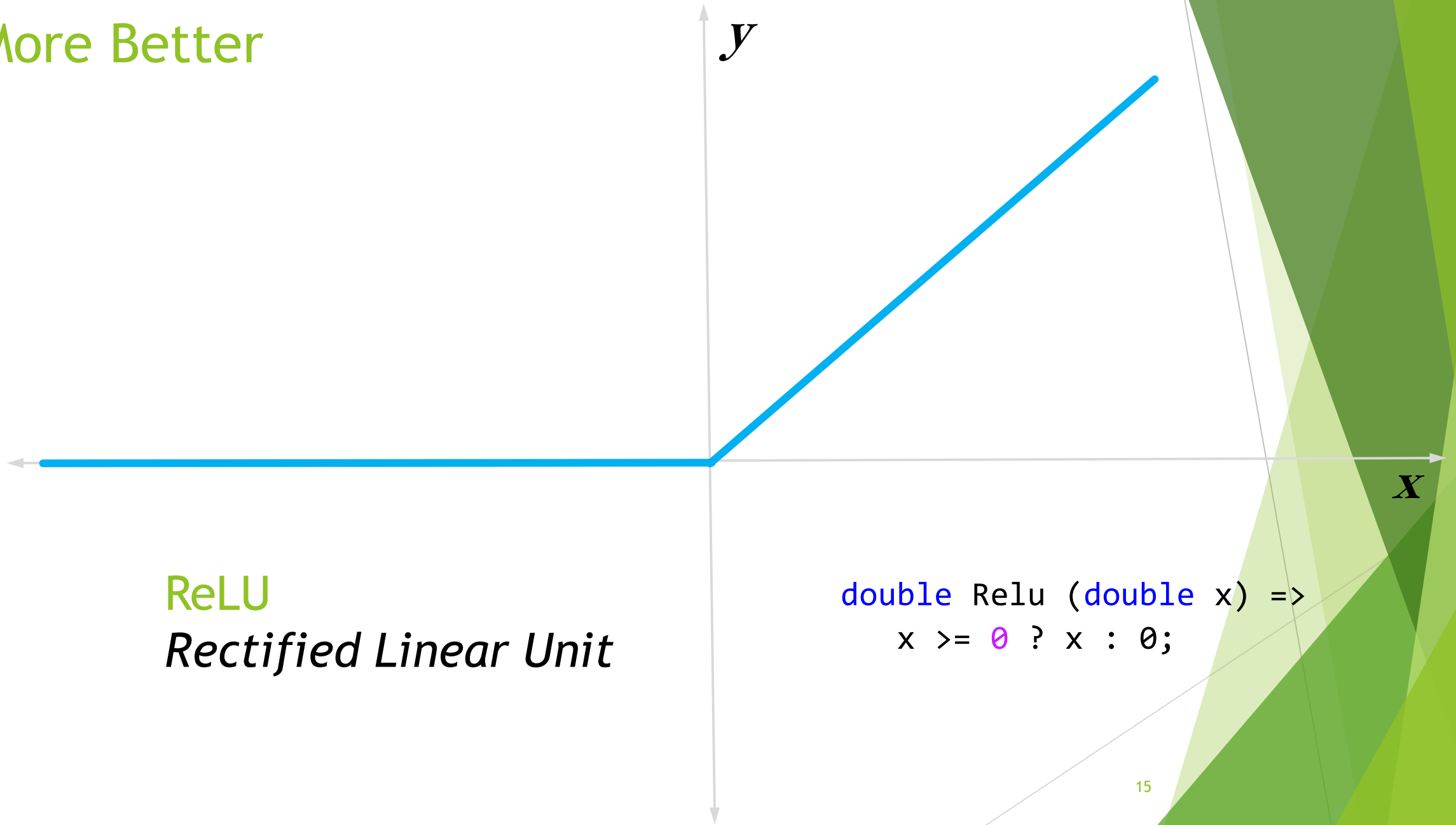
All values are numbers

Comparator

Better

Logistic Sigmoid

$$y = \frac{1}{1 + e^{-x}}$$
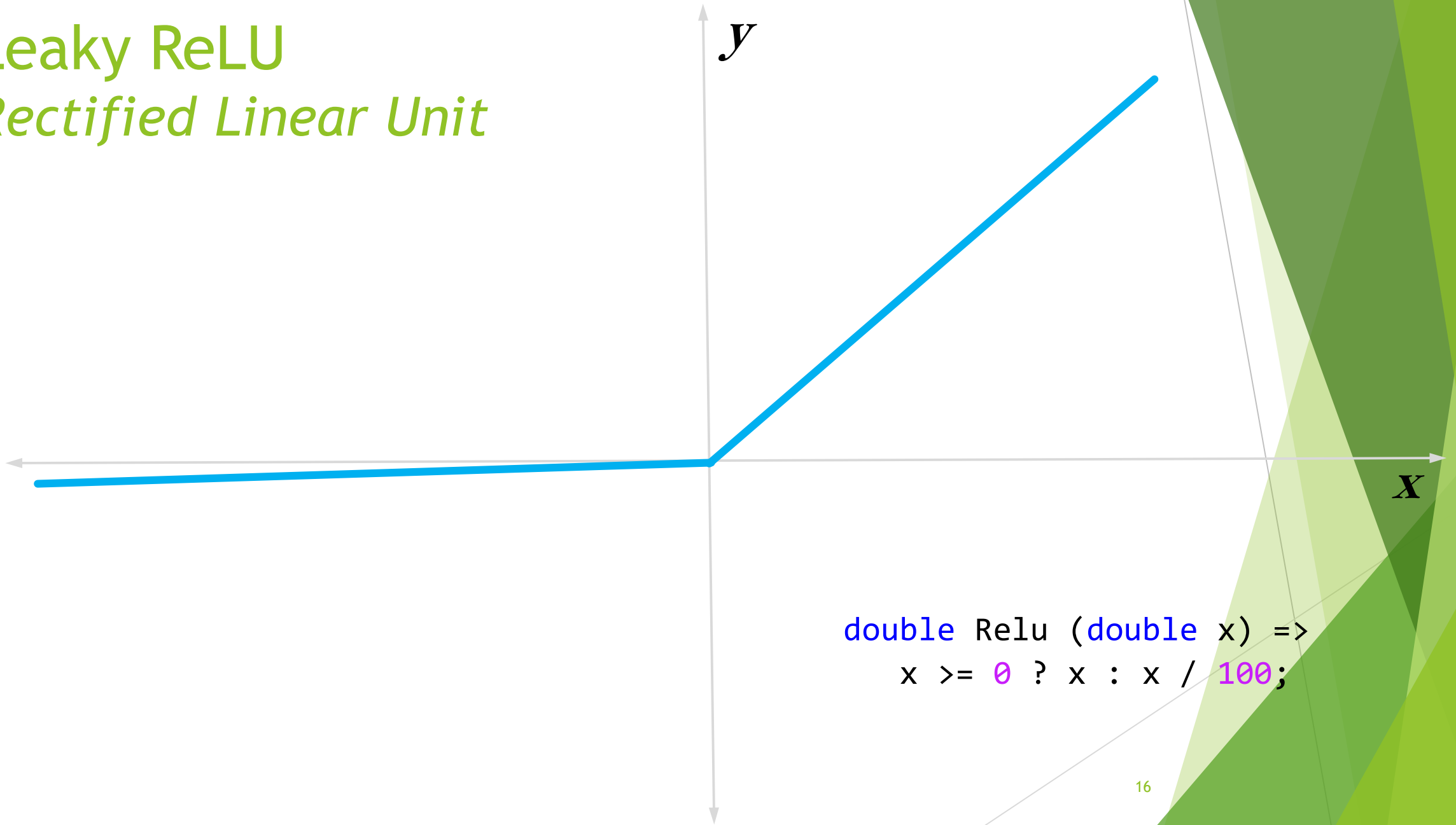
More Better

ReLU
Rectified Linear Unit

```
double Relu (double x) =>
    x >= 0 ? x : 0;
```

Leaky ReLU
Rectified Linear Unit


```
double Relu (double x) =>
    x >= 0 ? x : x / 100;
```

Some Adjustments

Rather than:

$$\text{Loss} = \text{sum (error)}^2$$

We *can* consider:

$$\textbf{\textit{Loss}} = \textbf{\textcolor{red}{½}} \textbf{\textit{ sum (error)}}^2$$

▶ We could instead use average (mean)

▶ *Cost* is a synonym for loss

We don't usually calculate the loss!

Stochastic Gradient Descent

foreach training sample:

Determine slope of the loss function at current position
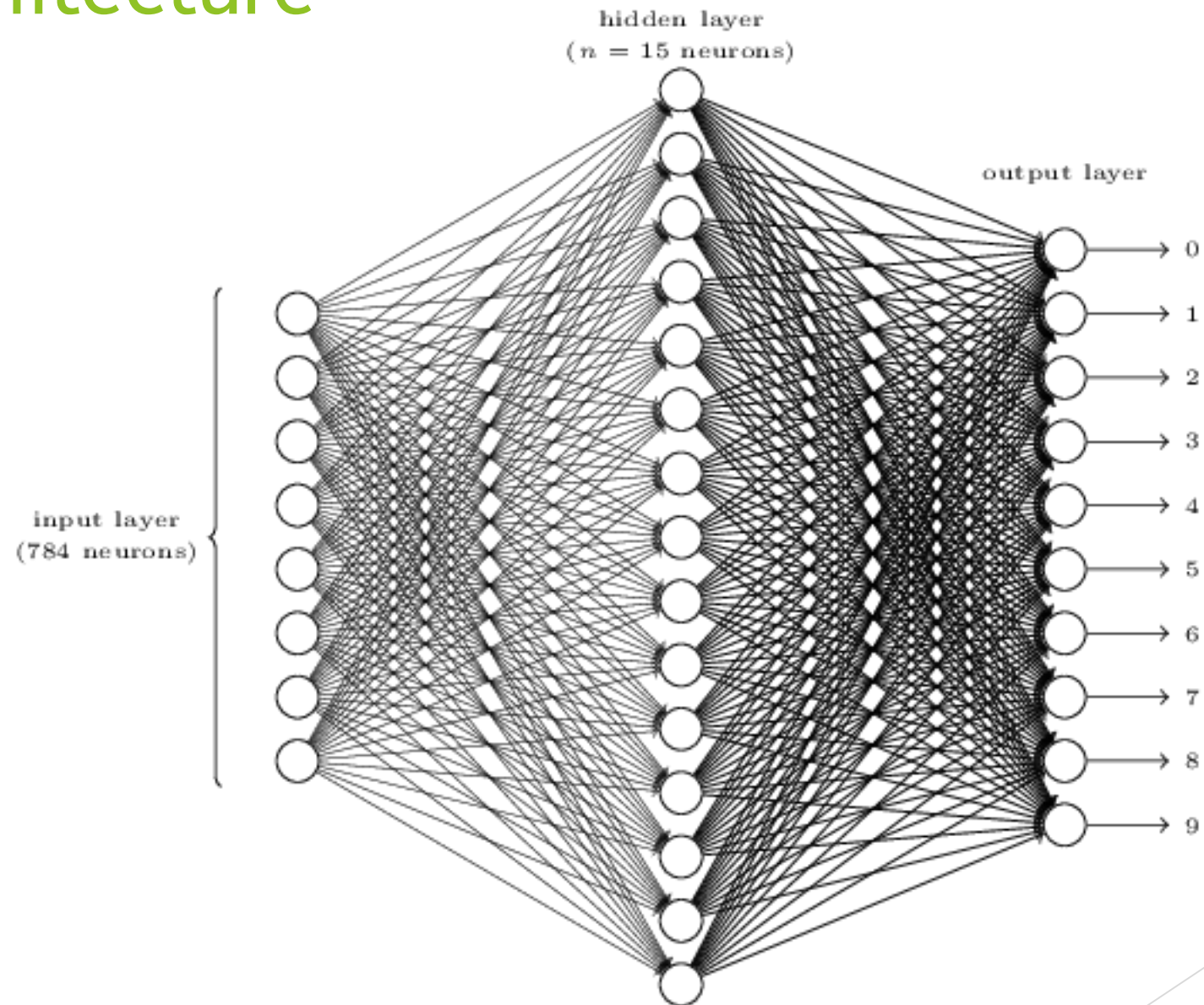
    (the error tells us the slope!)

    Step the following amount:

        – slope * learning rate

Architecture

hidden layer
($n = 15$ neurons)

output layer

input layer
(784 neurons)

0
1
2
3
4
5
6
7
8
9

Ideal

Future: GANs

Apple hires Google AI expert Ian Goodfellow to direct machine learning

Goodfellow is best known for inventing generative adversarial networks (GANs), which pair two AI algorithms together with the goal of continuously improving one another — one AI could be tasked with creating realistic images, while the other AI acts as a judge of real versus fake images, such that both AIs spur each other to become better over time. His research also involved combating adversarial attacks on neural networks that could undermine a GAN's ability to perform its functions, including methods that were undetectable to human observers.

Thank You & CodeOntheBeach-2019