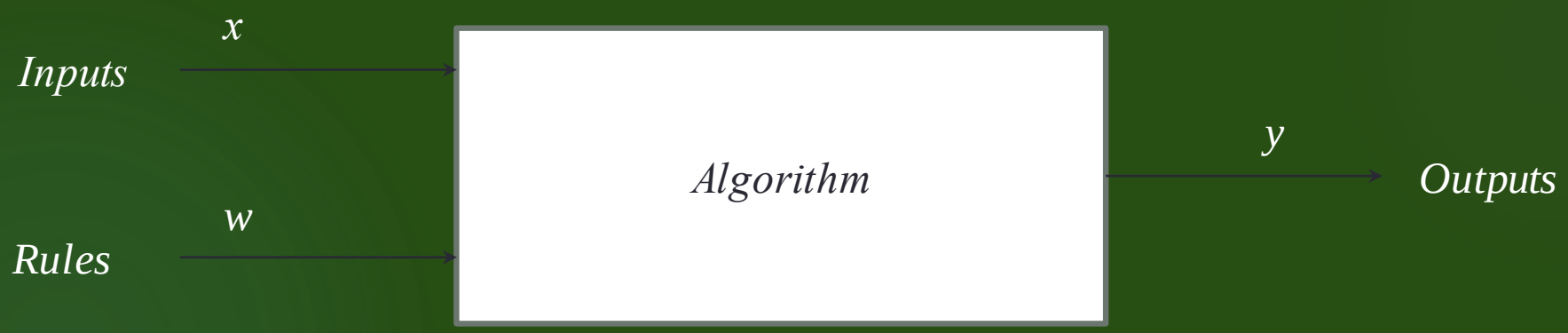
Basic ML/DL

- Introduction of Machine Learning
- Types of ML
- Linear Regression
- Logistic Regression
- Ensemble Learning
- XGBoost

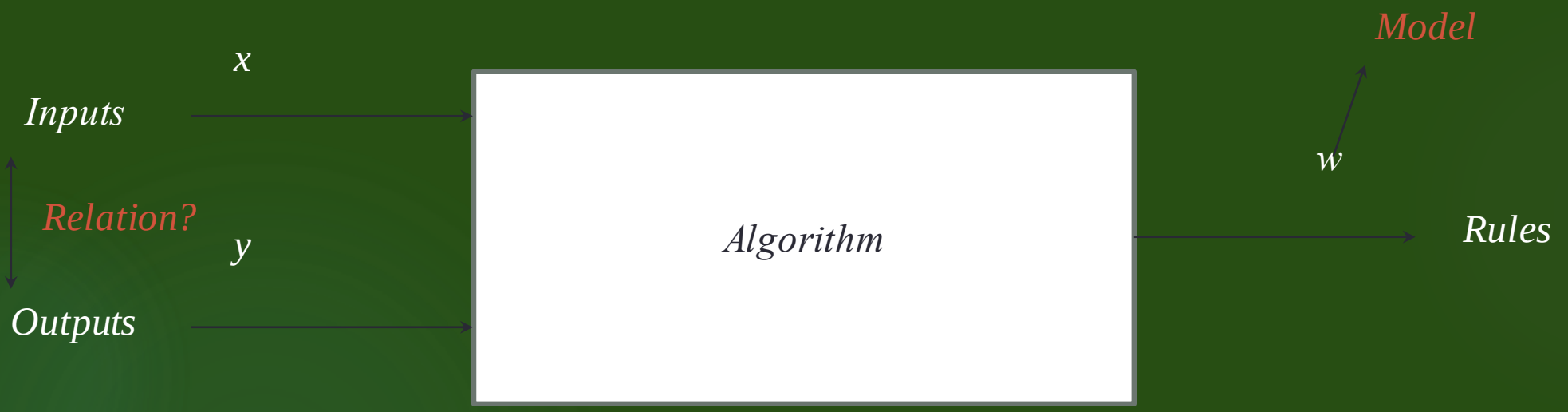
Introduction of Machine Learning

- What is Machine Learning?
- Types of ML
- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning

Classical Programming Approach



ML Programming Approach



ML is about finding out the relation between the Inputs and Outputs.

This relation is nothing but the **rules**- that set of rules is our **Machine Learning Model**

Regression examples

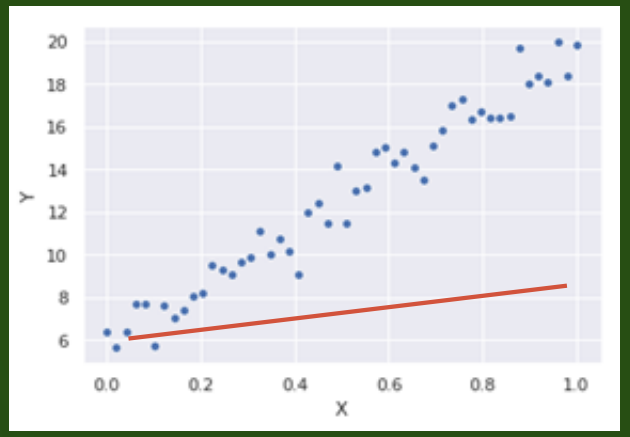
Stock market

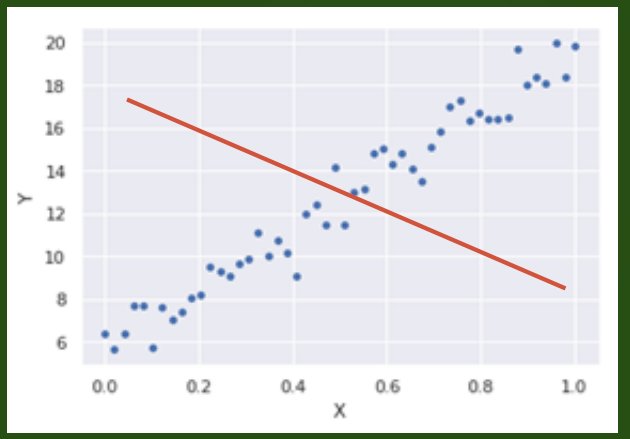
Weather prediction

Predict the temperature at any given location

General Univariate Regression (Simple Linear Regression)

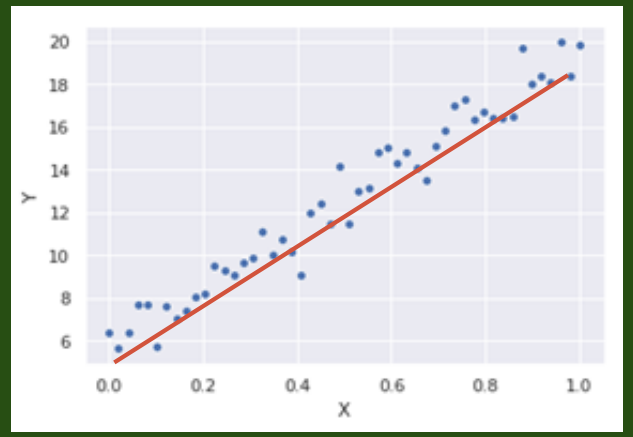
X	Y
$x^{(1)}$	$y^{(1)}$
$x^{(2)}$	$y^{(2)}$
:	:
:	:
$x^{(m)}$	$y^{(m)}$





Linear Regression

- Equation of line: $\hat{y} = mx + c$
- Equation has a form and
- Parameters



- Equation of line: $\hat{y} = m_i x + b$
  - *Equation has a **form** and*
  - ***Parameters***

We have to select those 'm' which minimize J

- For each value of w_0, w_1, we can have infinite number of lines or predicted values ($\hat{y}$)
- So, which one to choose?
- We first introduce loss function and with its help we calculate cost
  - Loss: $L = y_i - \hat{y}_i$
  - Cost: $J = \sum(y_i - \hat{y}_i)^2/n$

How to learn optimal parameter

Gradient Descent

$$\text{Error}_{(m,b)} = \frac{1}{N}\sum_{i=1}^{N}(y_i - (mx_i + b))^2$$

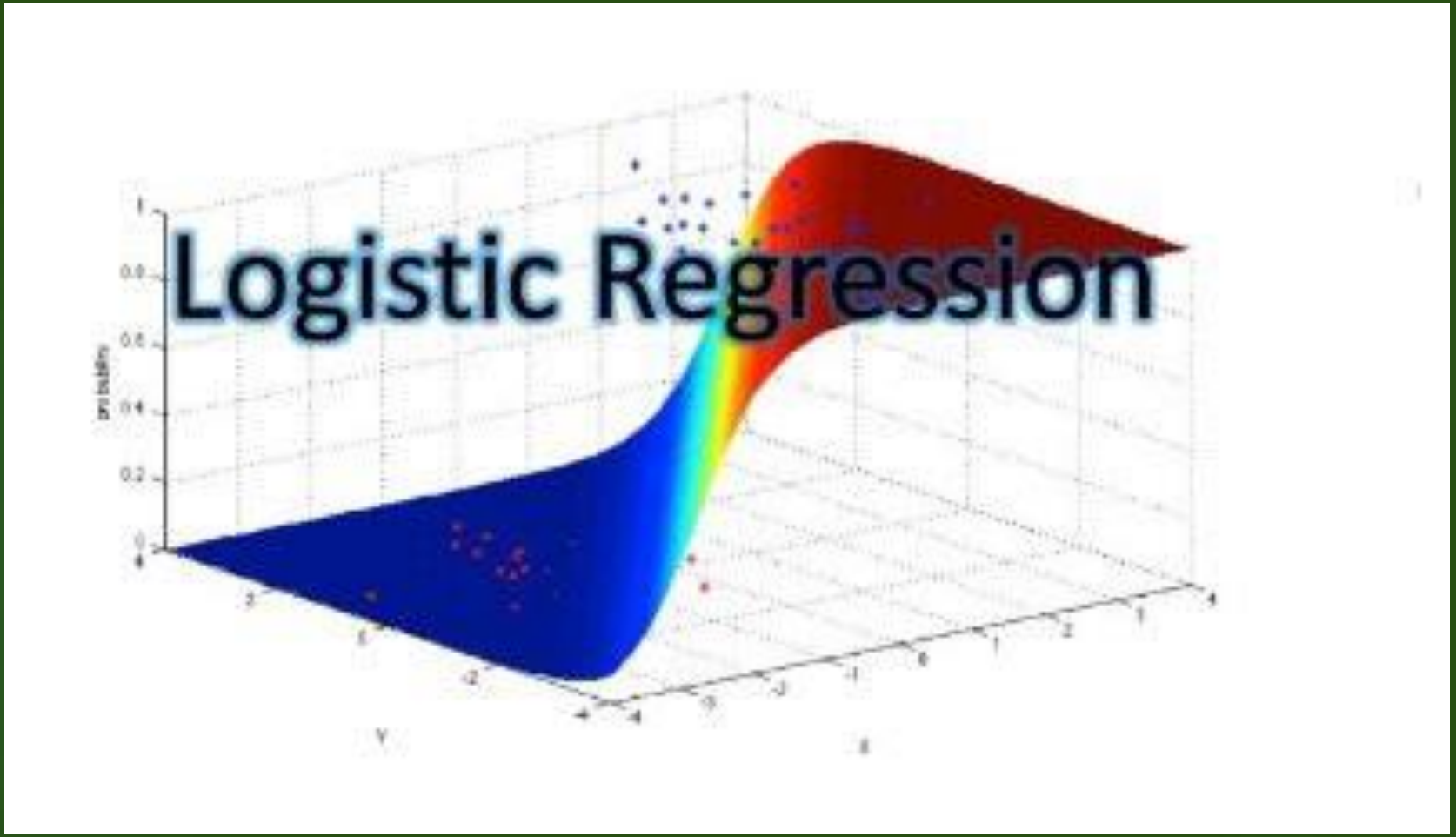
$$\frac{\partial}{\partial \text{m}} = \frac{2}{N}\sum_{i=1}^{N} - x_i(y_i - (mx_i + b))$$

$$\frac{\partial}{\partial \text{b}} = \frac{2}{N}\sum_{i=1}^{N} - (y_i - (mx_i + b))$$

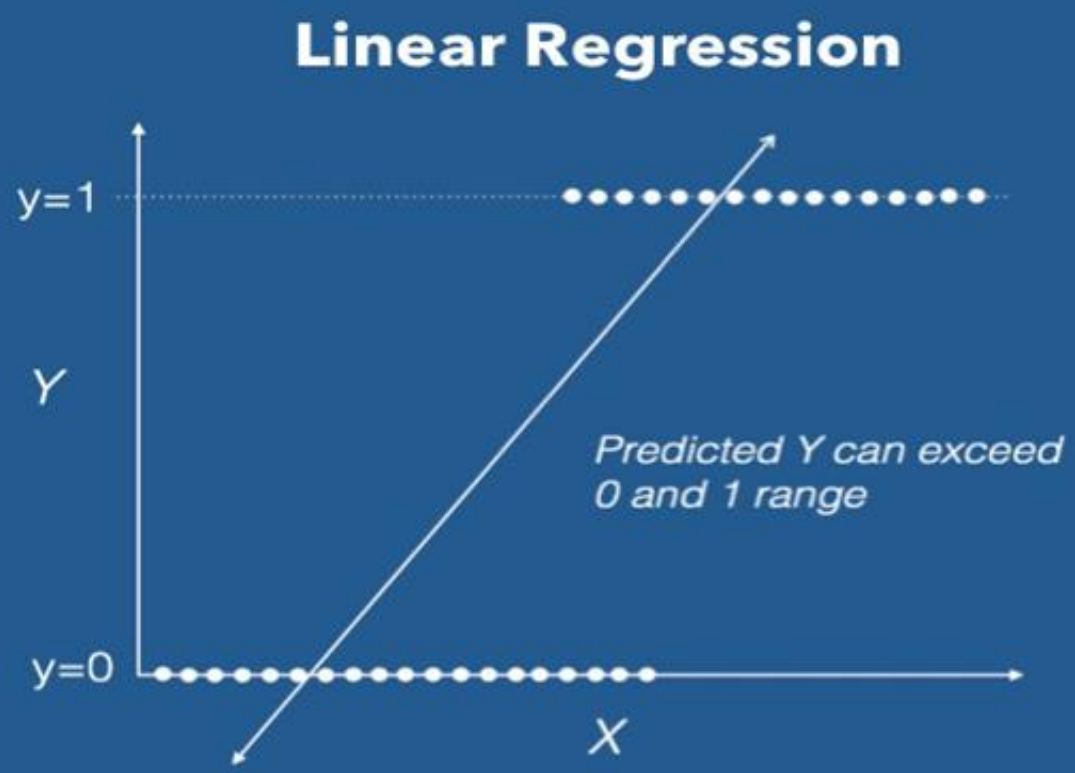
Steps in Gradient Descent

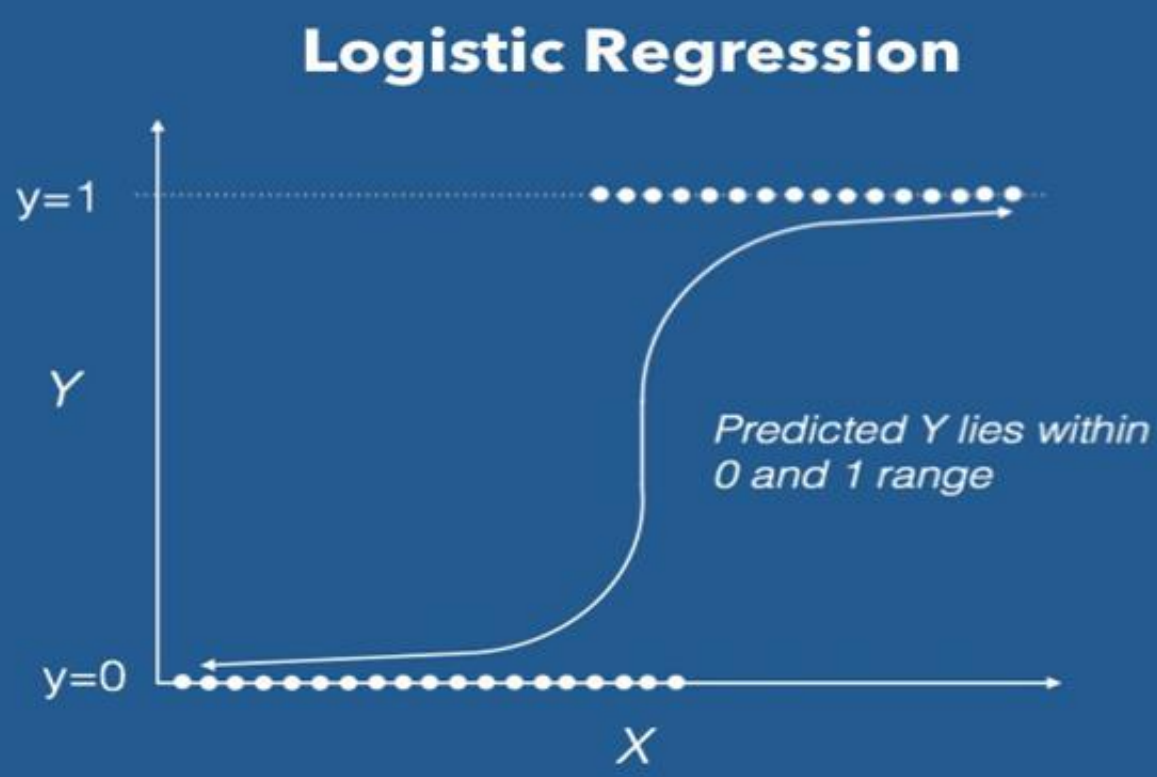
- Choose random initial value for weight and bias.
- Calculate loss
- Do partial derivative of loss wrt to weight and bias.
- Update your weight and bias
  - $m = m - \alpha * \partial J / \partial m$
  - $b = b - \alpha * \partial J / \partial b$
- Where 'α' is learning rate, to control the step size
- Repeat until convergence

Logistic Regression



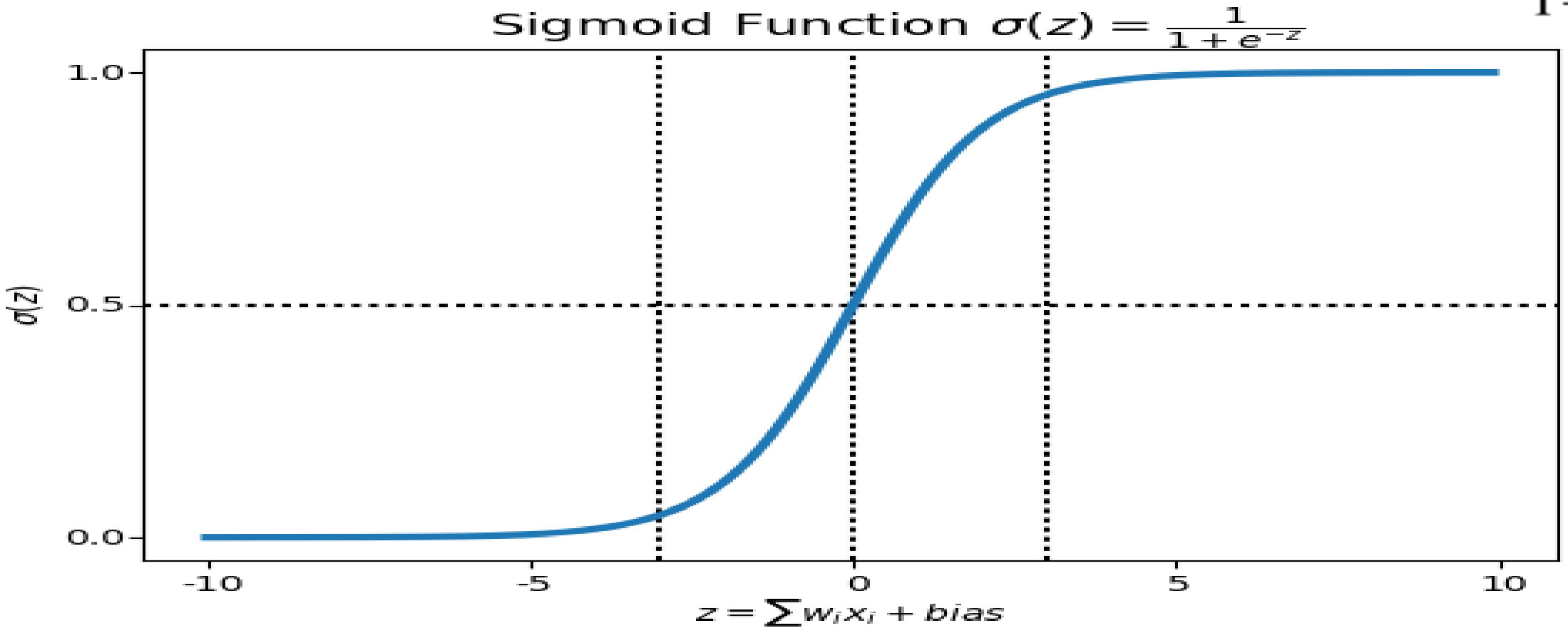
Logistic Regression





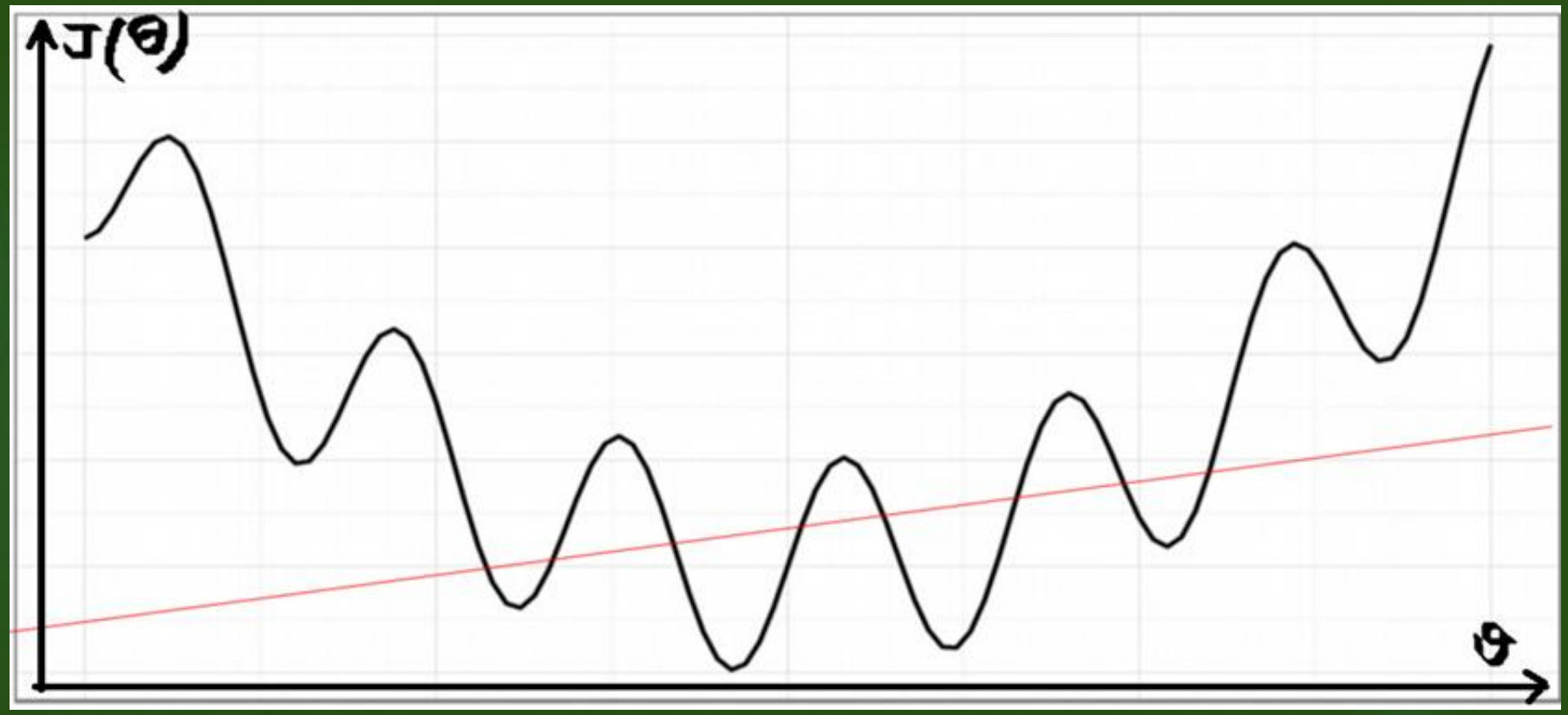
Sigmoid Function

$$f(x) = \frac{1}{1 + e^{-(x)}}$$



Cost Function

$$J(\theta) = \frac{1}{2}\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})^2.$$

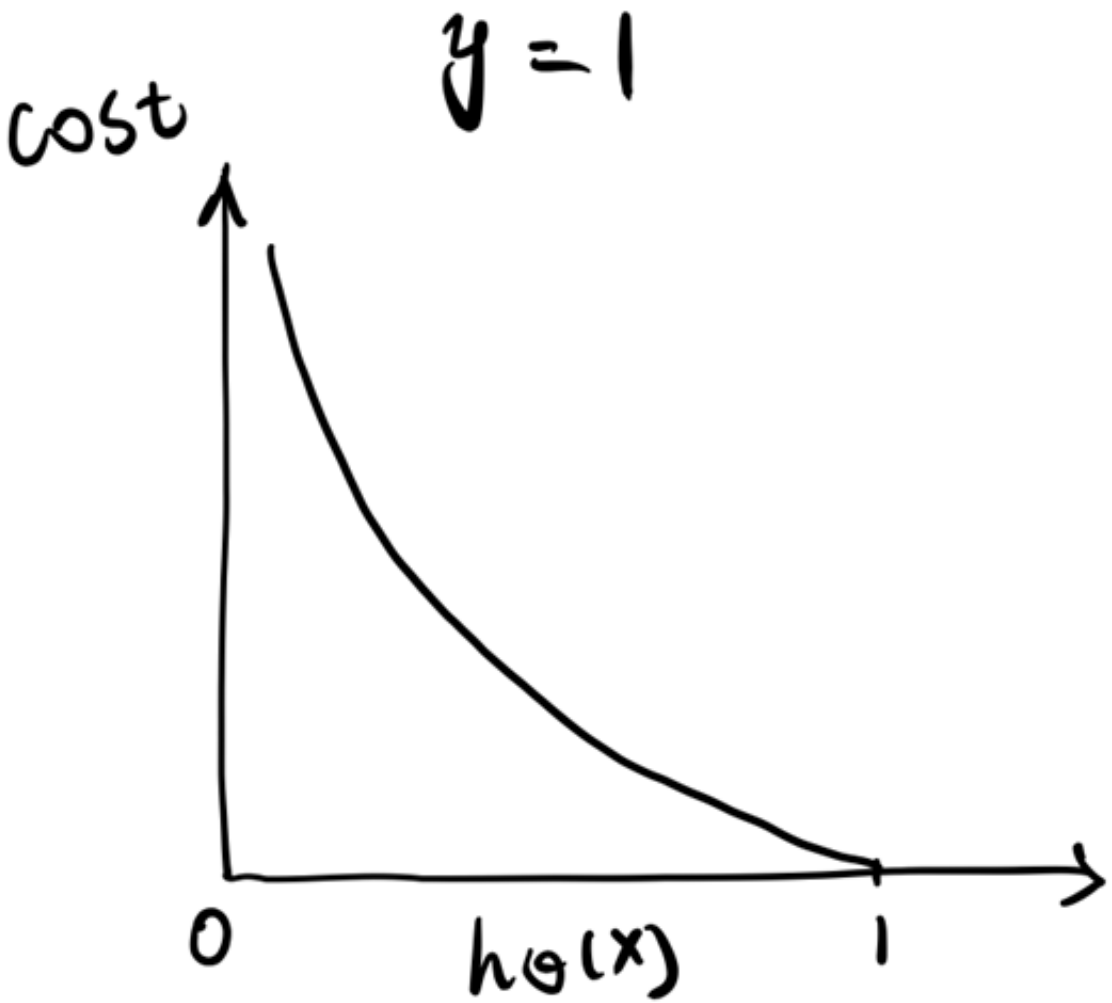


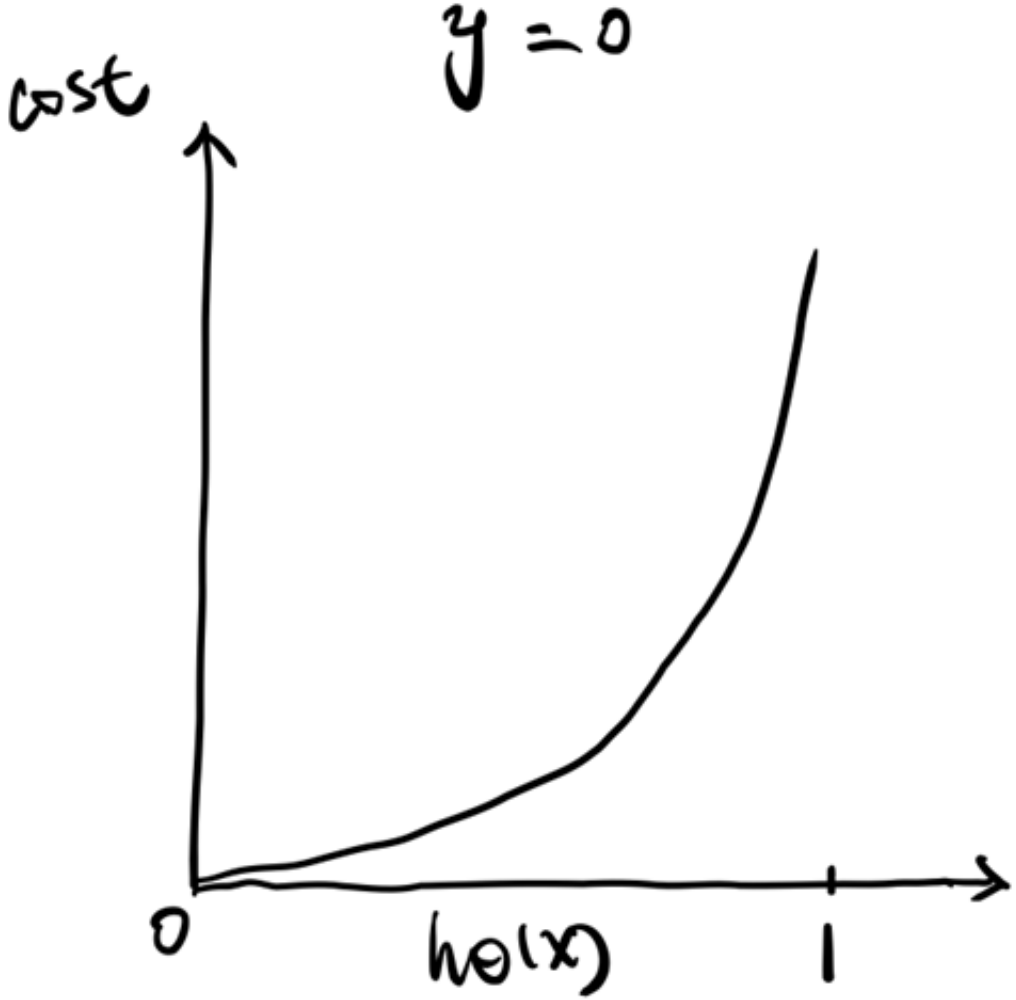
For logistic regression, the Cost function is defined as:

$$h\theta(X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X)}}$$

$$Cost(h_\theta(x), y) = \begin{cases} -log(h_\theta(x)) & \text{if } y = 1 \\ -log(1 - h_\theta(x)) & \text{if } y = 0 \end{cases}$$

Advantage of Log Loss:

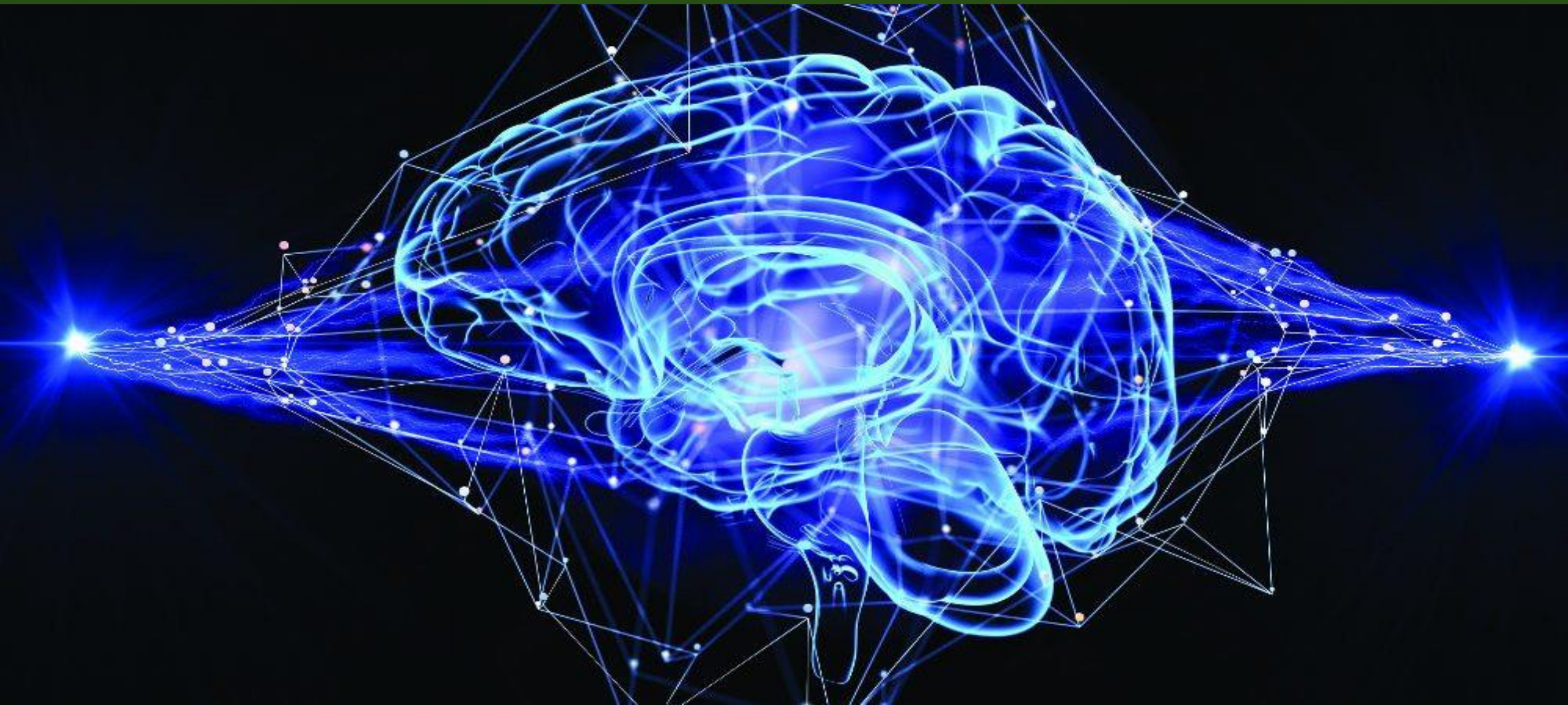




Ensemble Methods

- What is Ensemble Methods
- Types of Ensemble Methods
- Weak Learner
- Strong Learner

Introduction of Deep Learning



Agenda

- What is deep learning
- Difference between ML and DL
- Why we need DL
- When to use Deep Learning
- Application of DL
- What is Neural Networks
- Types of NN
- ANN
- CNN
- RNN

What is Deep Learning

- Subset of machine learning
- Learn like human
- Think and Learn
- Directly from input
- Solve any pattern Recognition

Difference between ML and DL

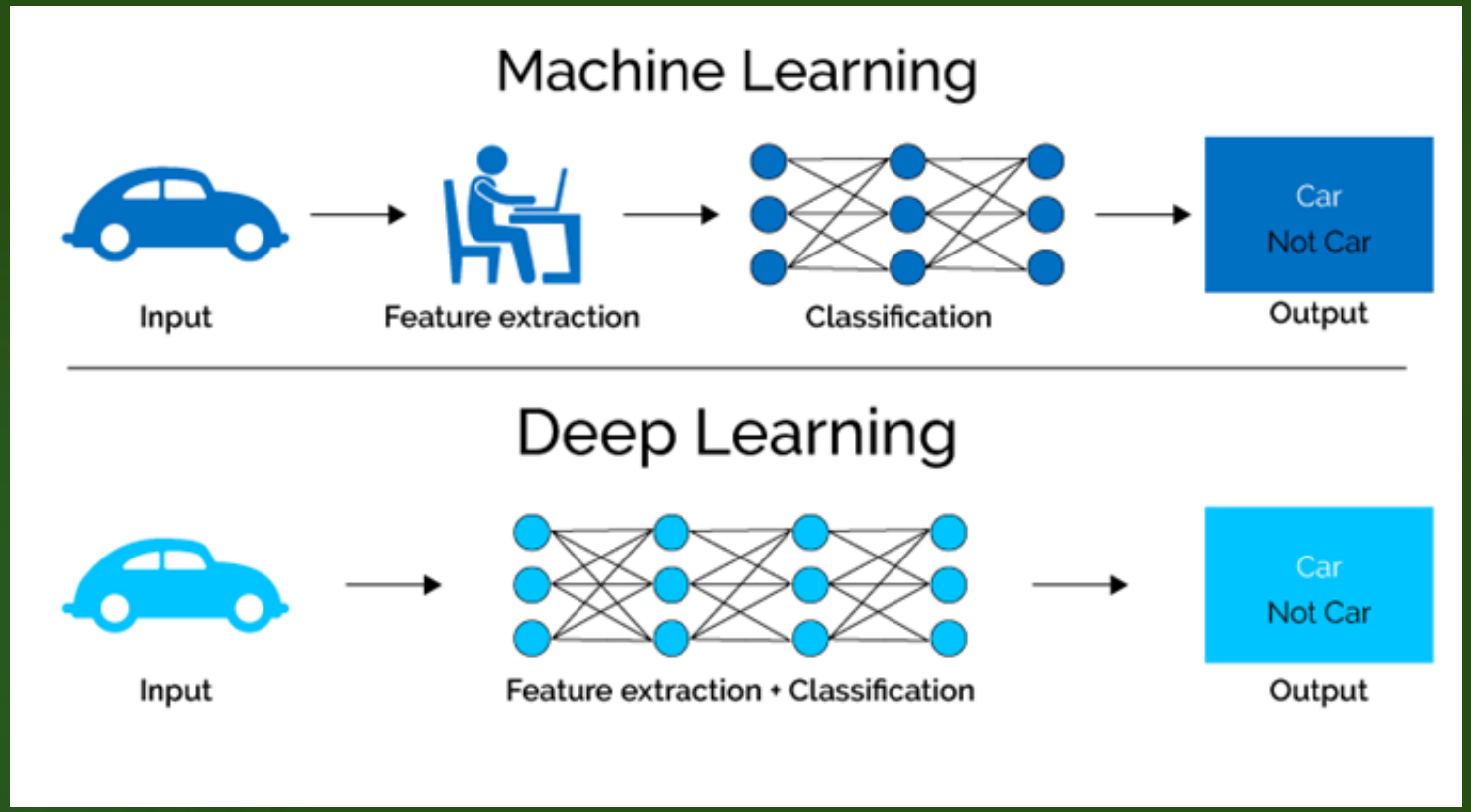
Machine Learning

- Small – medium amount of Data
- Not Dependent on high-end machine
- Less training time
- Works on specific algorithms

Deep Learning

- Large amount of Data
- Heavily dependent on high-end machine
- Longer time to train
- Its create own algorithms

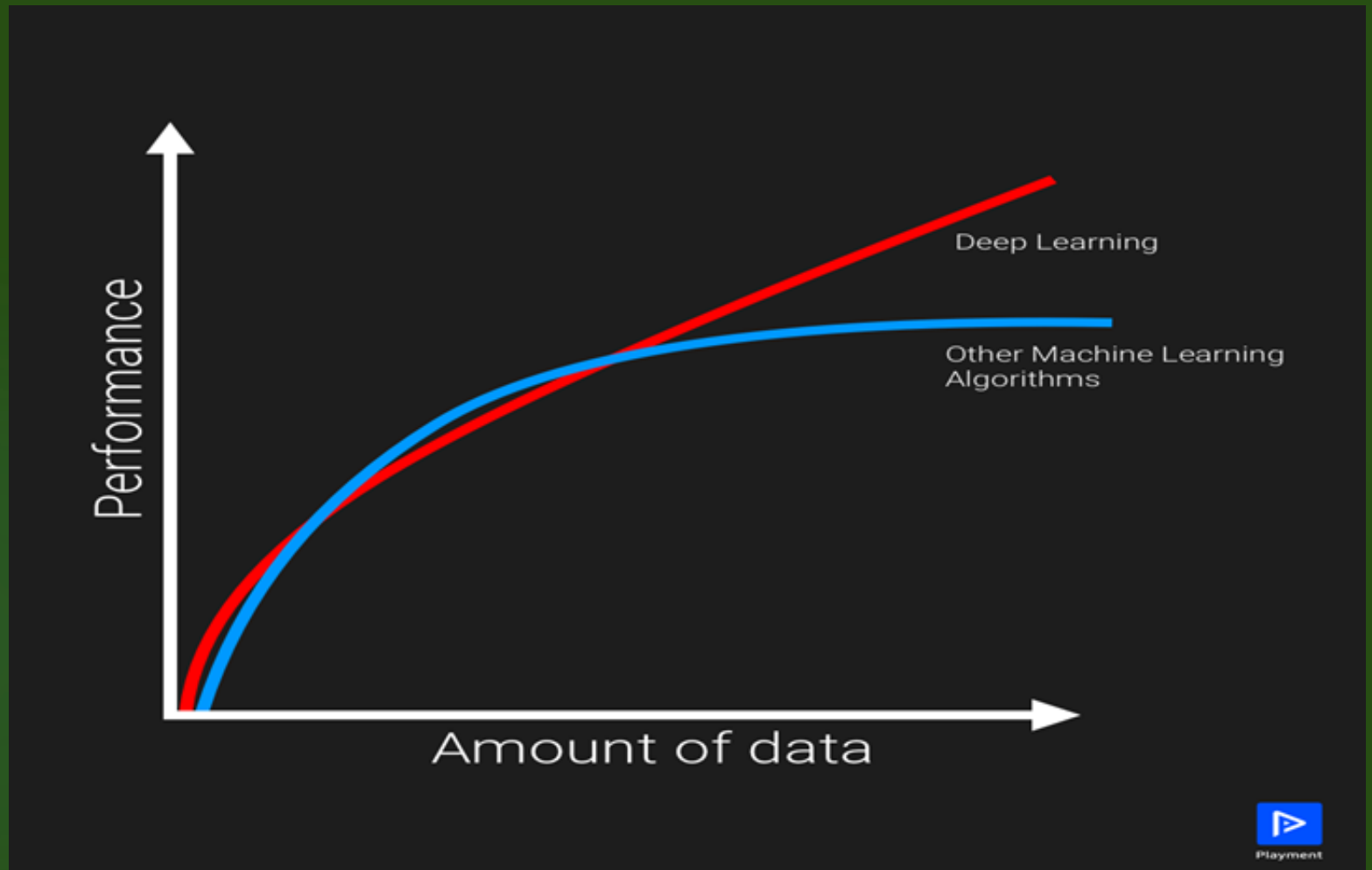
Difference between ML and DL



Comparing deep learning with machine learning

Source: https://www.xenonstack.com/blog/data-science/log-analytics-with-deep-learning-and-machine-learning

Why we need Deep Learning



Source: https://becominghuman.ai/what-is-deep-learning-and-why-you-need-it-9e2fc0f0e61b

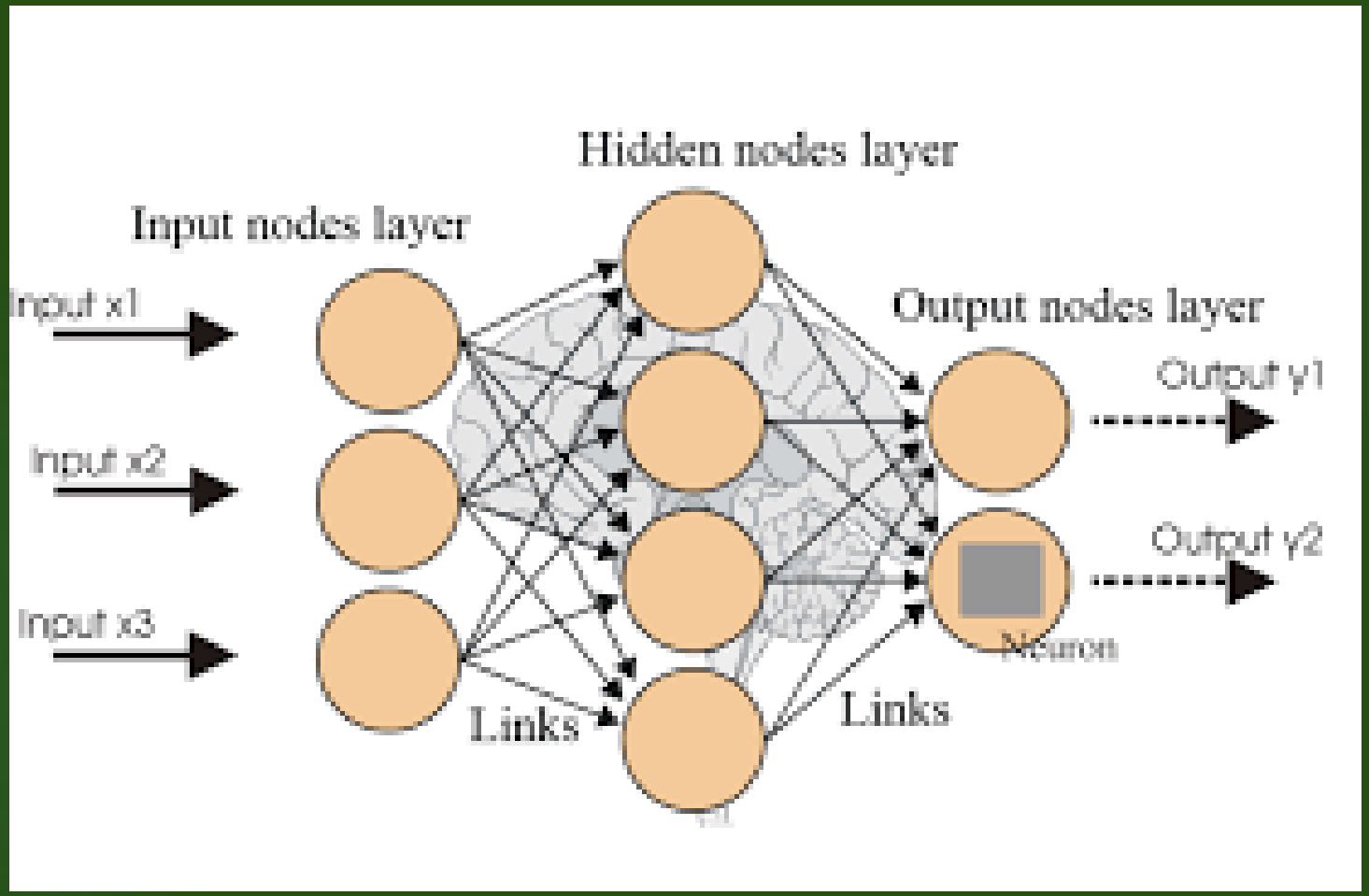
When to use Deep Learning

- Large dataset
- High end infrastructure
- Lack of domain knowledge
- Complex problem

Application of Deep Learning

- Virtual Assistants
- Translation
- Image colorization
- Face recognition
- Medicine and Pharmaceuticals

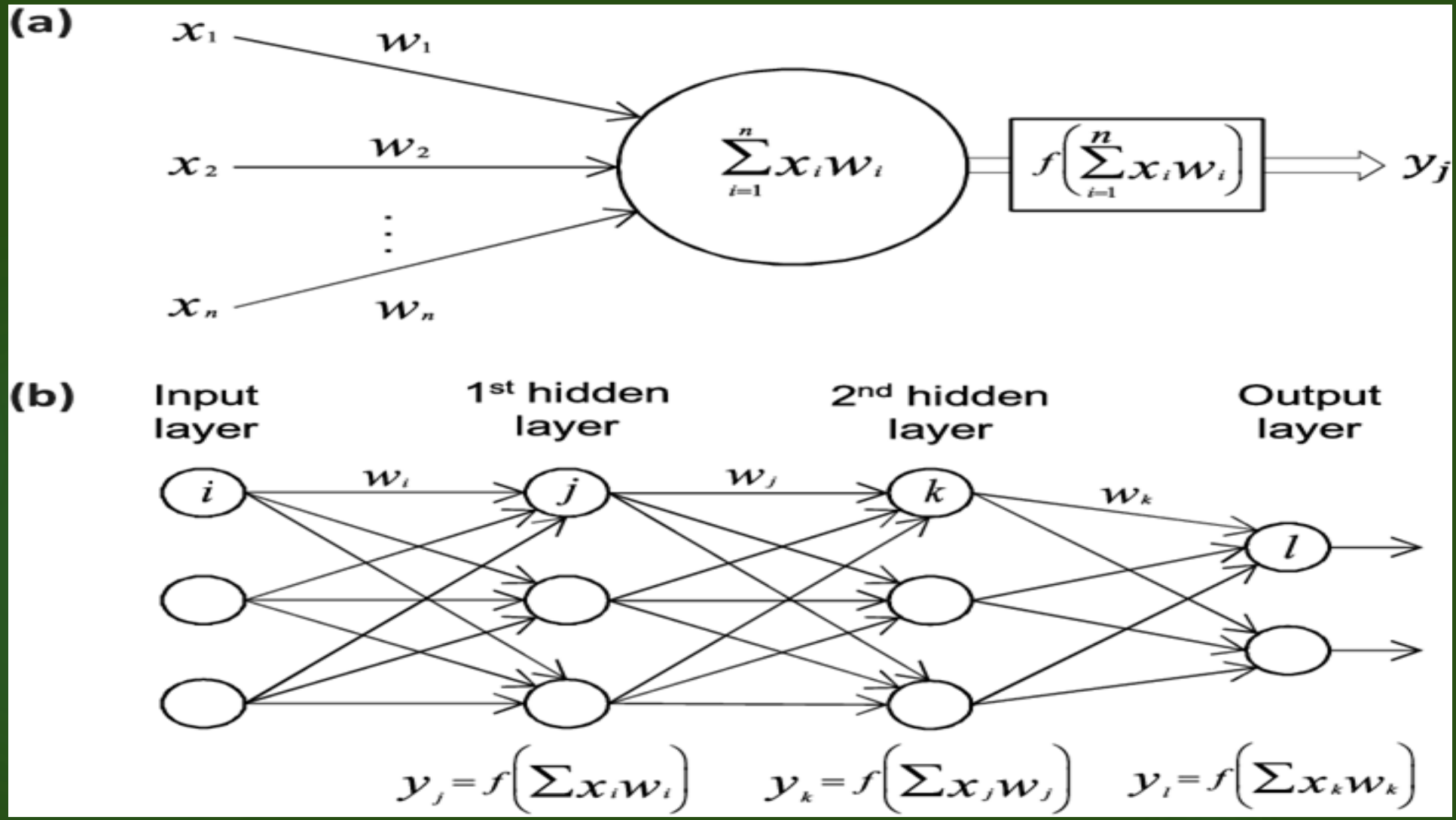
What is Neural Network



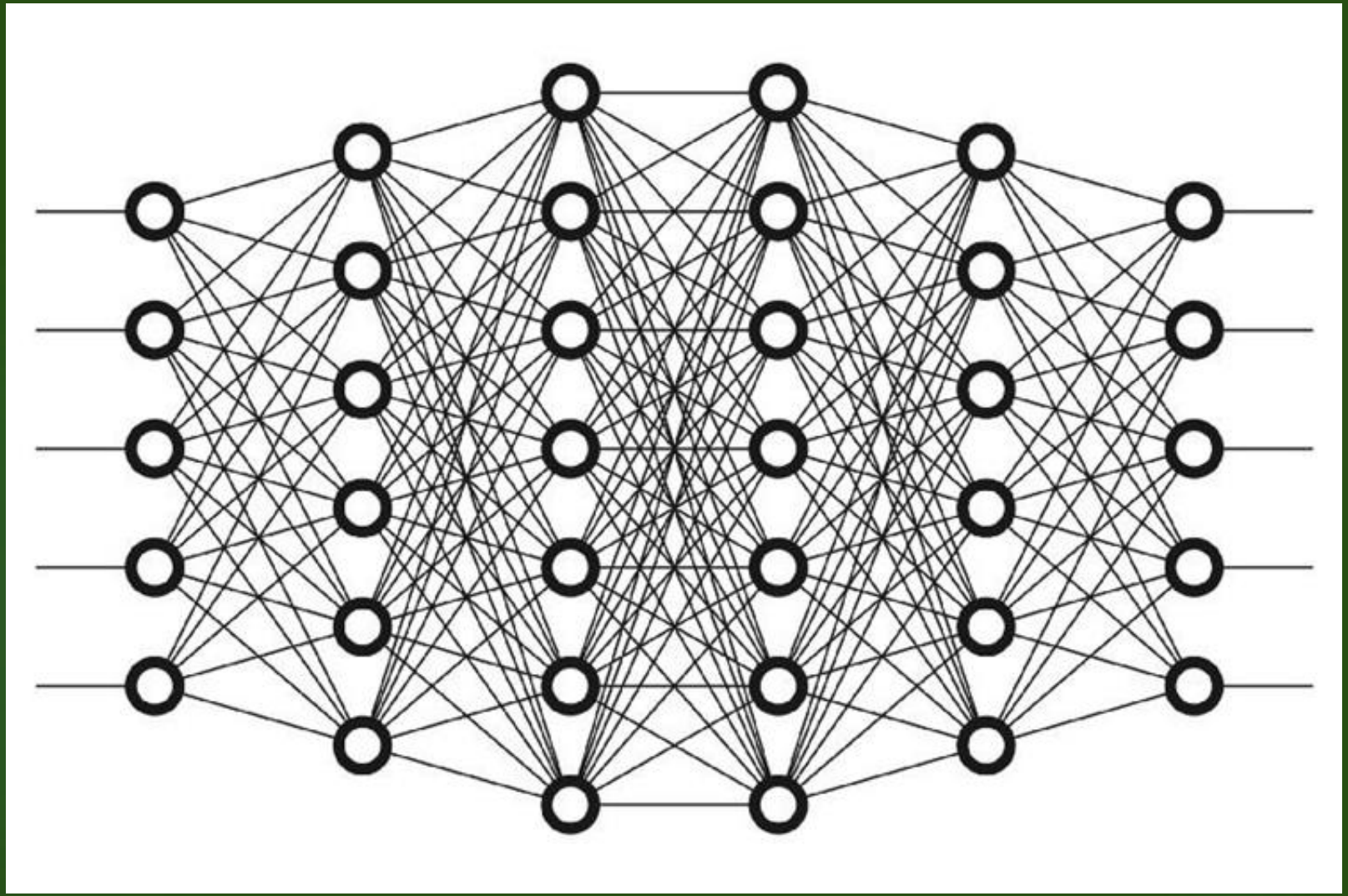
Basic Types of Neural Networks

- Artificial Neural Network (ANN)
- Convolutional Neural Network (CNN)
- Recurrent Neural Networks (RNN)

Artificial Neural Networks



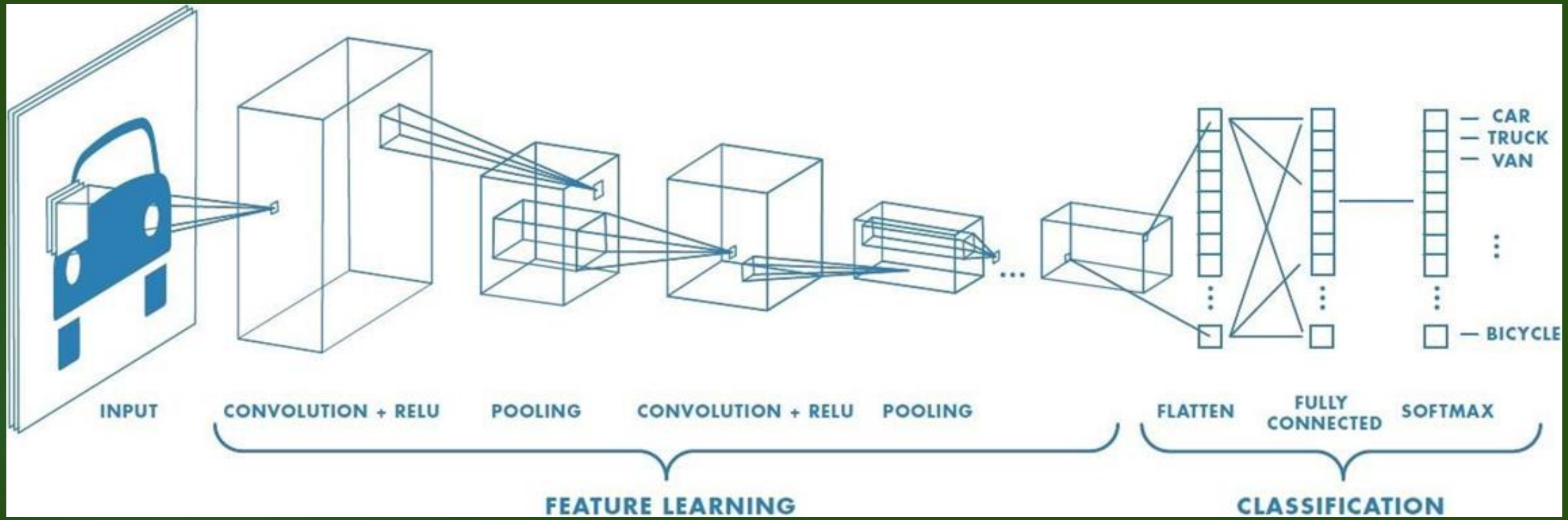
Deep Neural Network



How ANN works

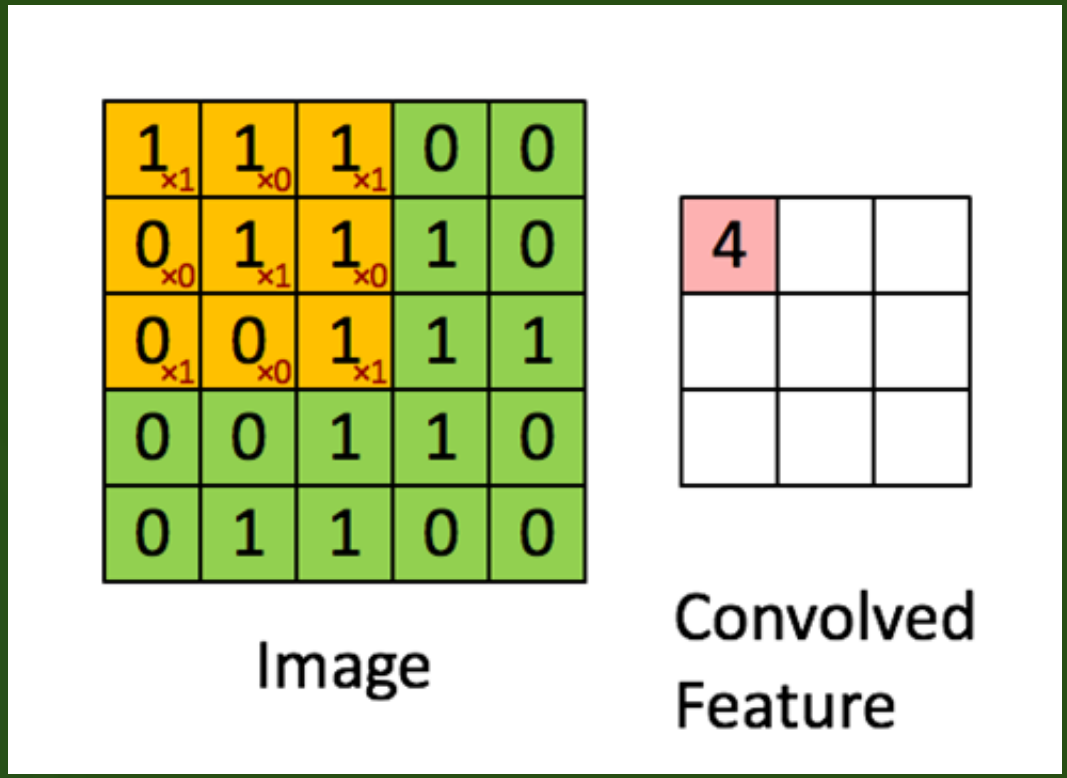
- Initialize random weight in node.
- Feed input in input layer.
- Feed forward network
- Predict output
- Calculate error
- Update weights for minimize loss
- Repeat this till the convergence criterion is met

Convolution Neural Network (CNN)



Why we need CNN over ANN?

Convolution Layer – The Kernel



Recurrent Neural Network