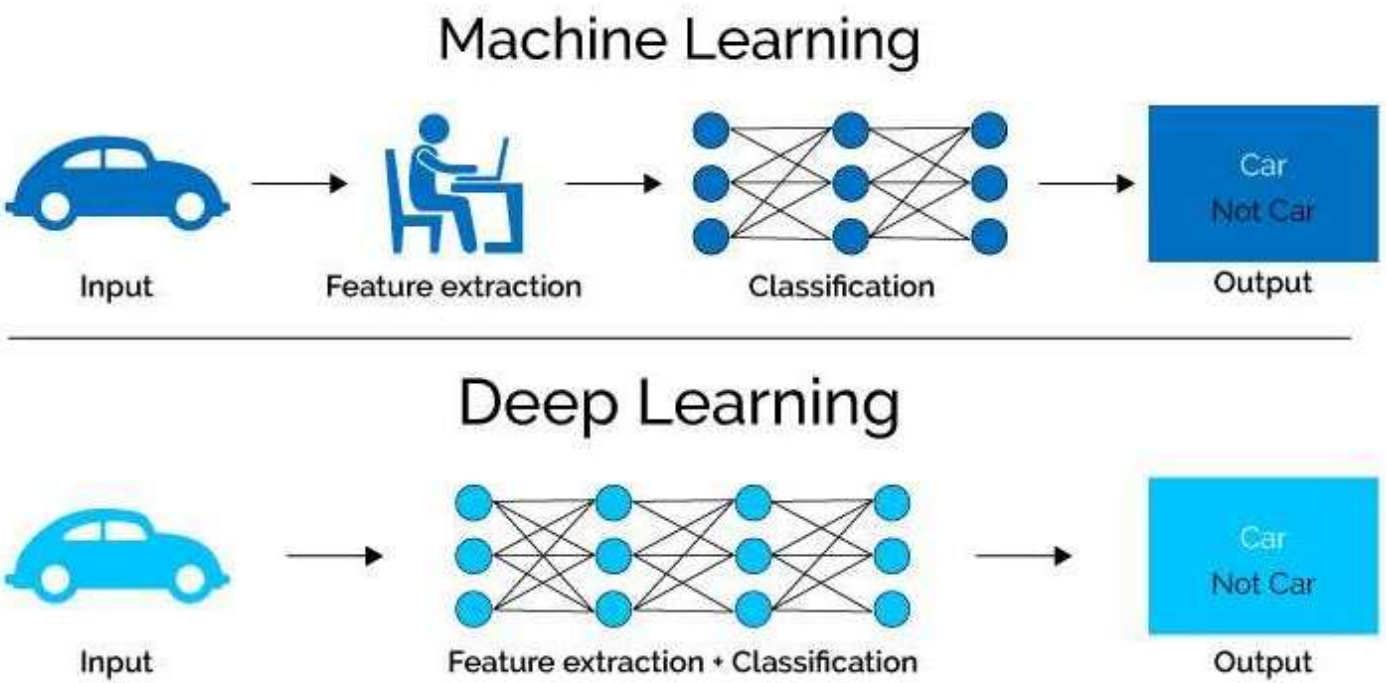
Agenda

- Limitation of Conventional Models
- Neural Networks
- Why Neural Networks
- Hyper parameters
  - Learning rate
  - Epochs
  - Regularization
  - Activation functions
- Play ground

Recap


Machine Learning
Input
Feature extraction
Classification
Car
Not Car
Output
Deep Learning
Input
Feature extraction + Classification
Car
Not Car
Output

Shortcoming of Conventional ML

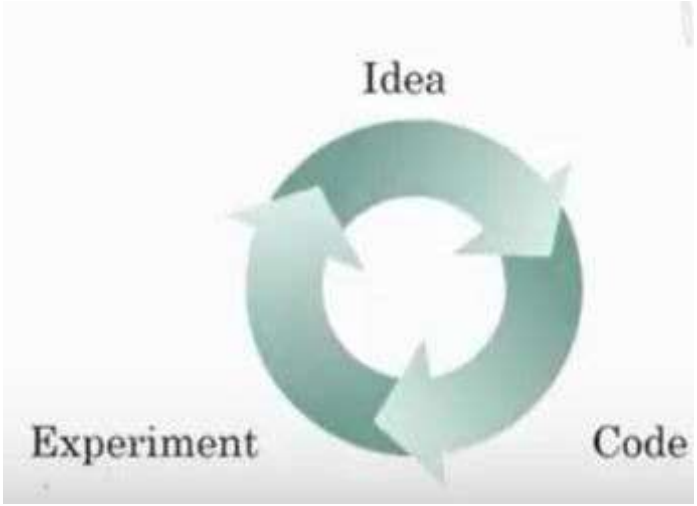
- Can NOT directly work with **unstructured**, **high dimensional data** (Required Feature Engineering)
  Example: Image and Video Data , Genes Data

- Can learn only so much from available data (i.e performance does not increase after certain threshold even if more training data is available)

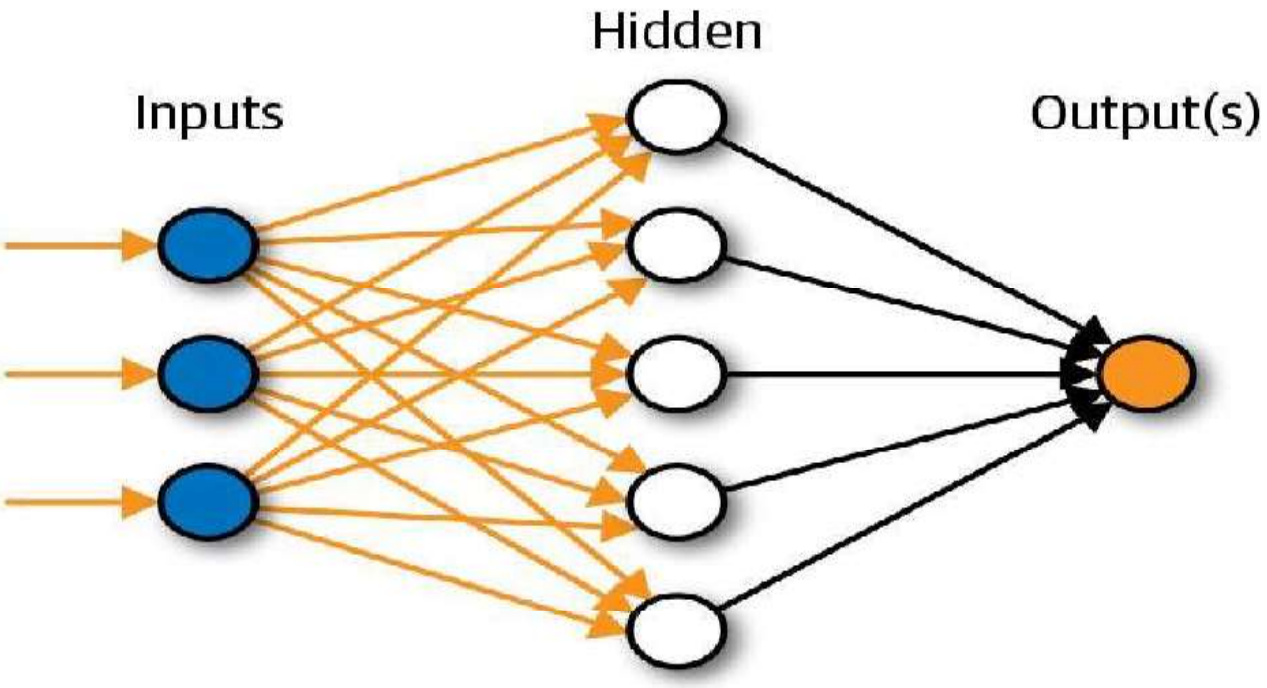
Why is Deep Learning Popular now ?

Most of the techniques used in deep learning are **as old as 70 years**, then why it is suddenly gaining so much traction ?

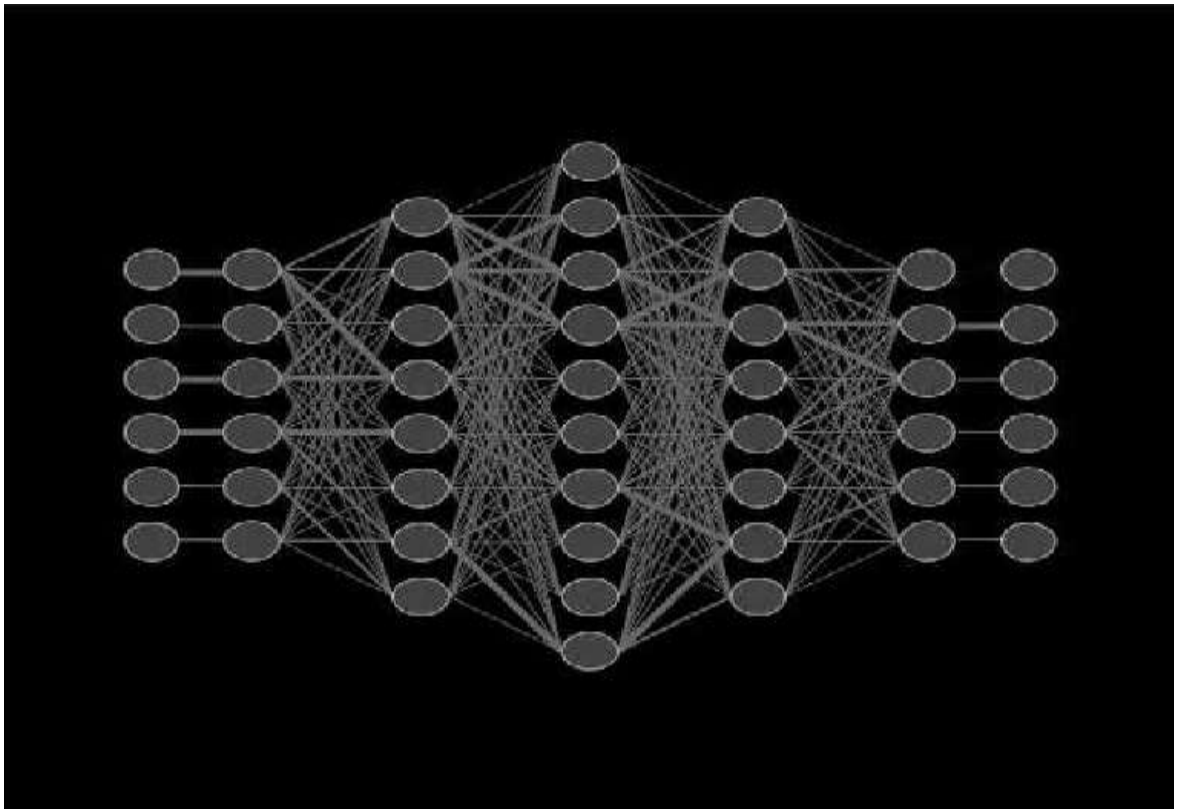
1. Data
2. Compute
3. Algorithm



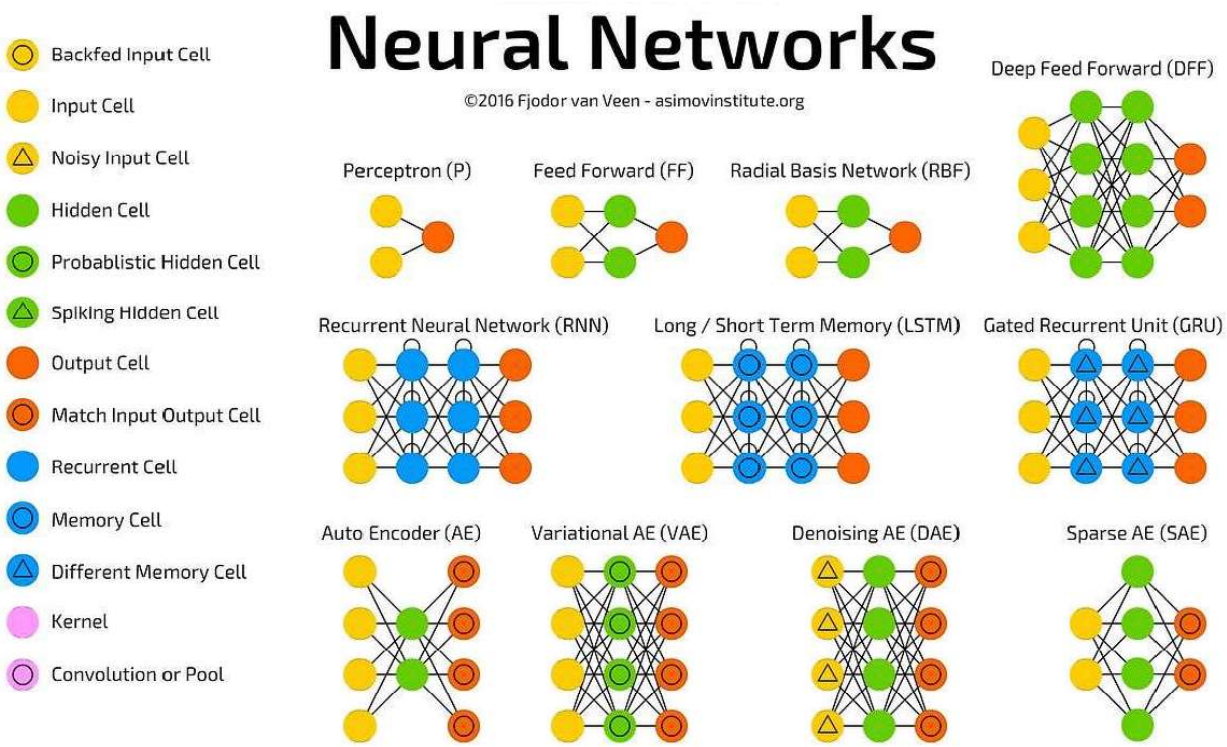
Simple Neural Network



Neural Network

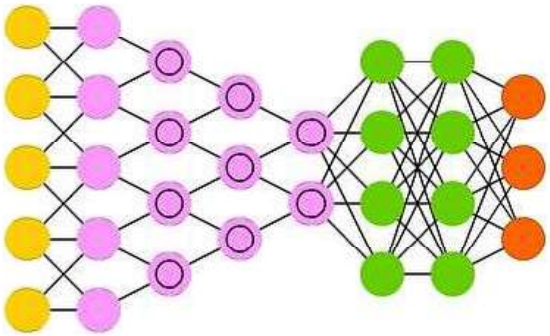


Types of Neural Network

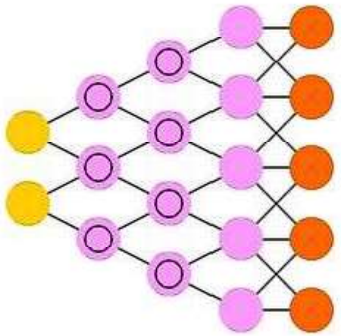


Types of Neural Network

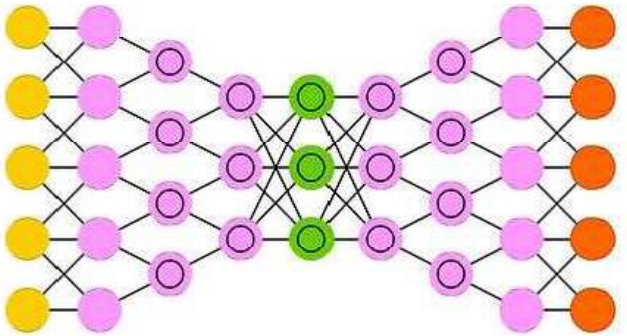
Deep Convolutional Network (DCN)



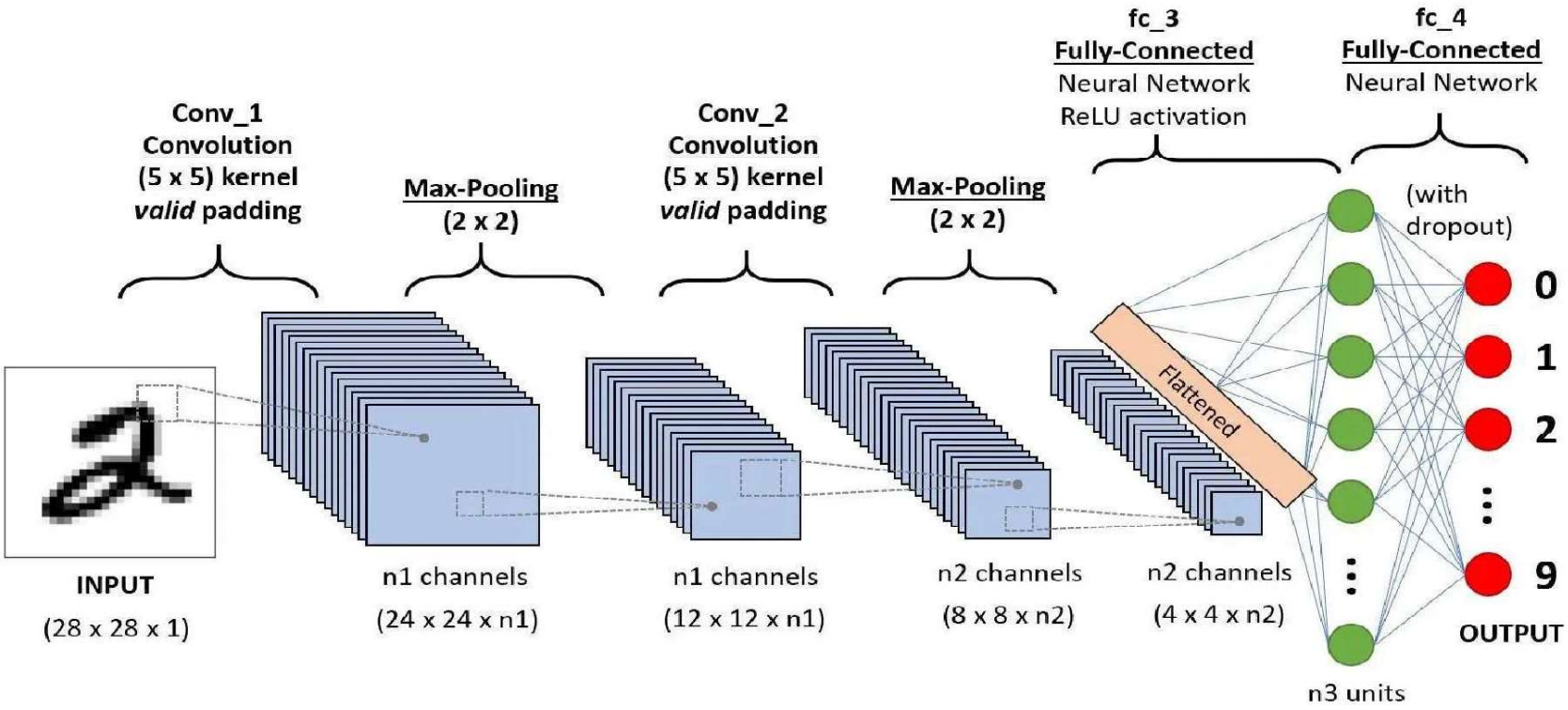
Deconvolutional Network (DN)



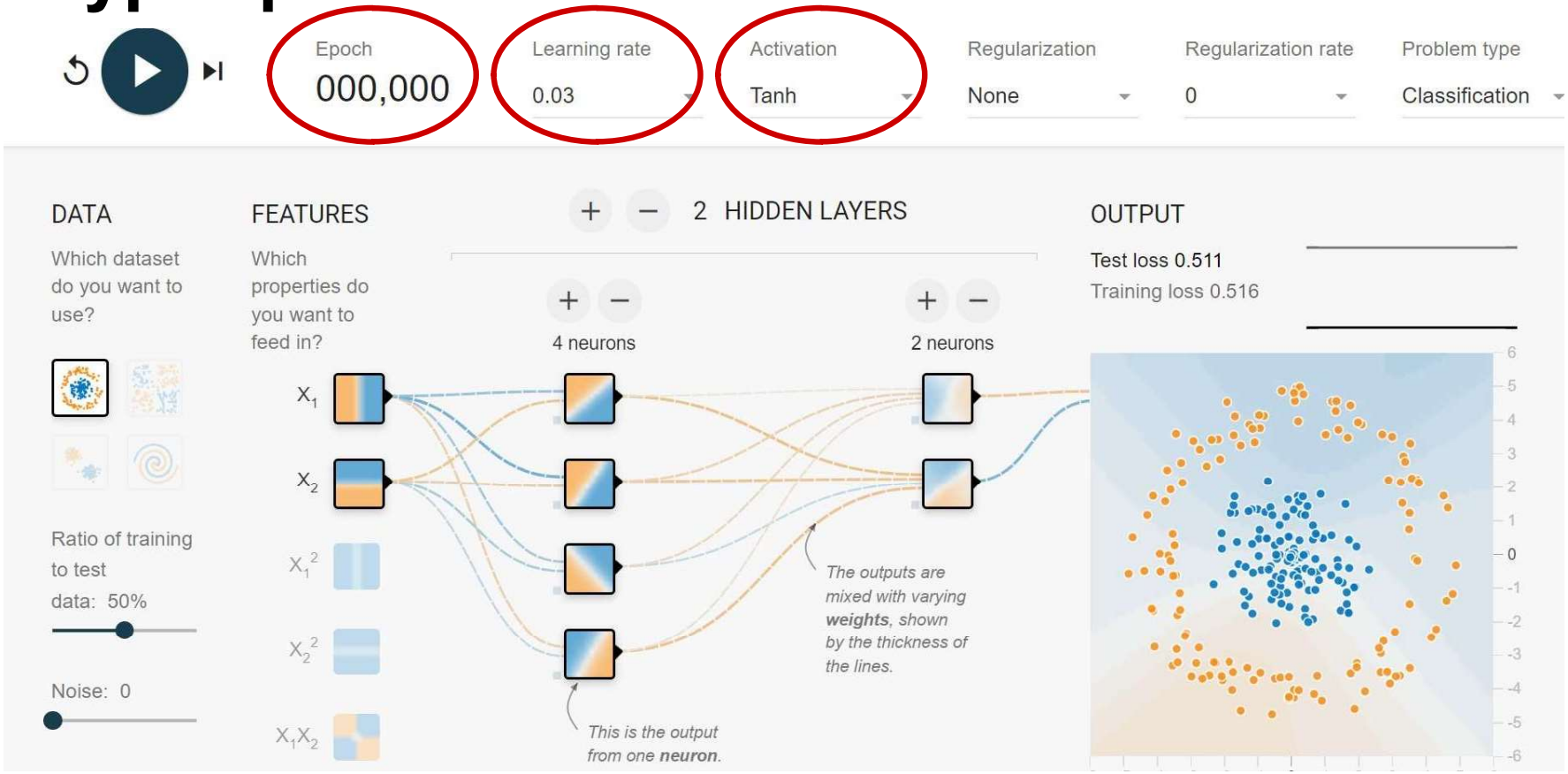
Deep Convolutional Inverse Graphics Network (DCIGN)



Types of Neural Network



Hyper-parameters

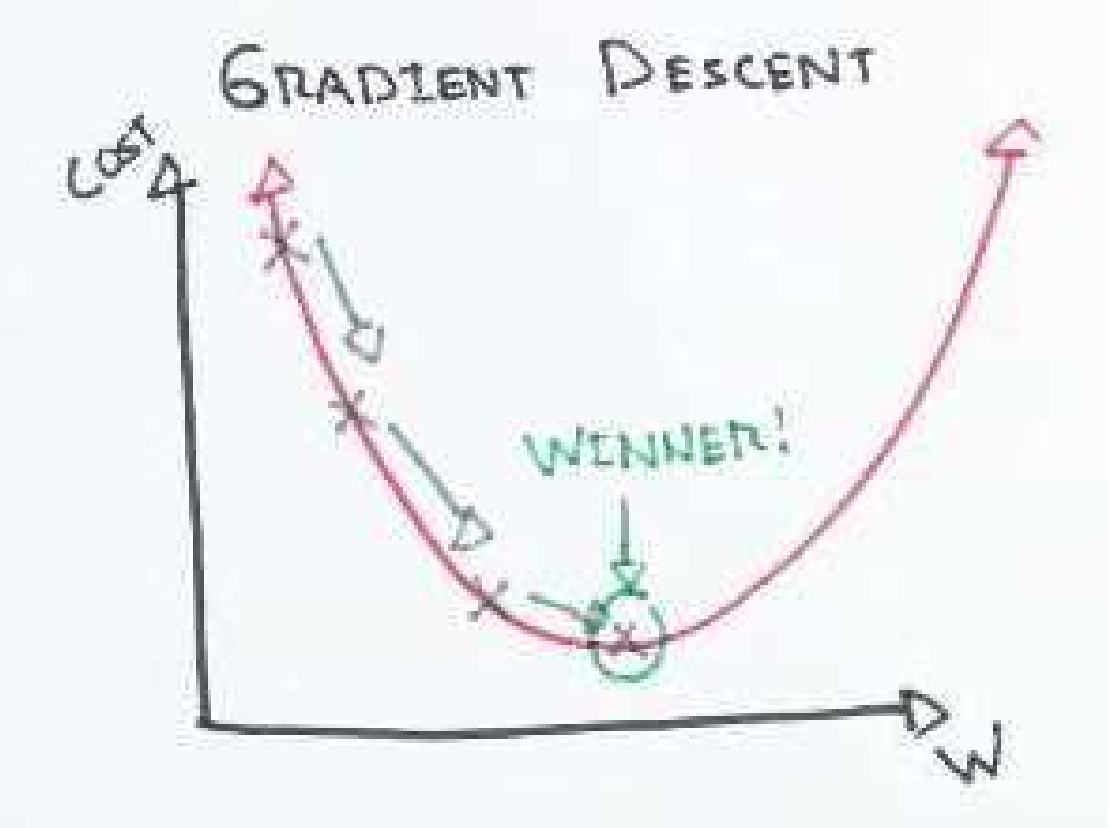


Learning Rate

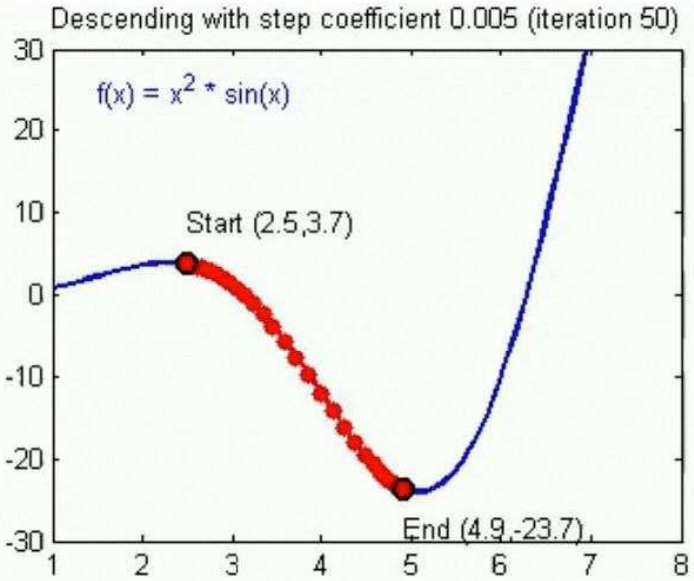
A hyperparameter that controls the **step size** during **gradient descent** optimization. It determines how quickly the model adjusts its parameters in the direction that **reduces** the **loss.**

A **higher learning** rate might lead to **faster convergence** but **risks overshooting**, while a **lower rate** might slow down convergence.

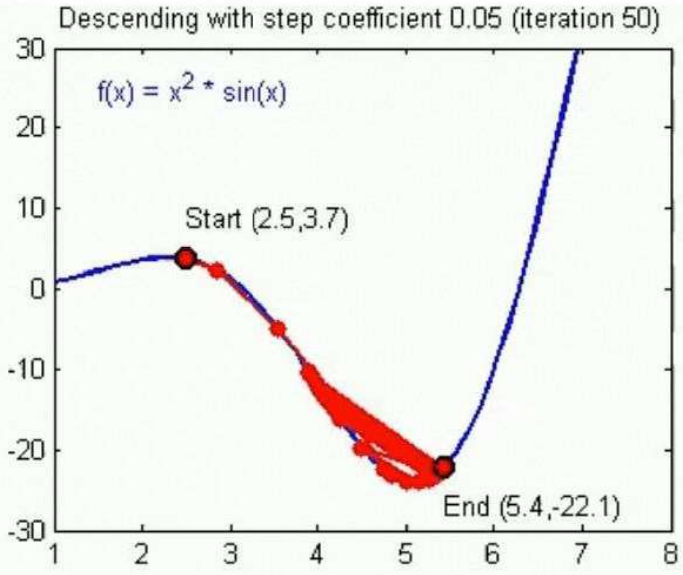
Gradient Descent



Convergence

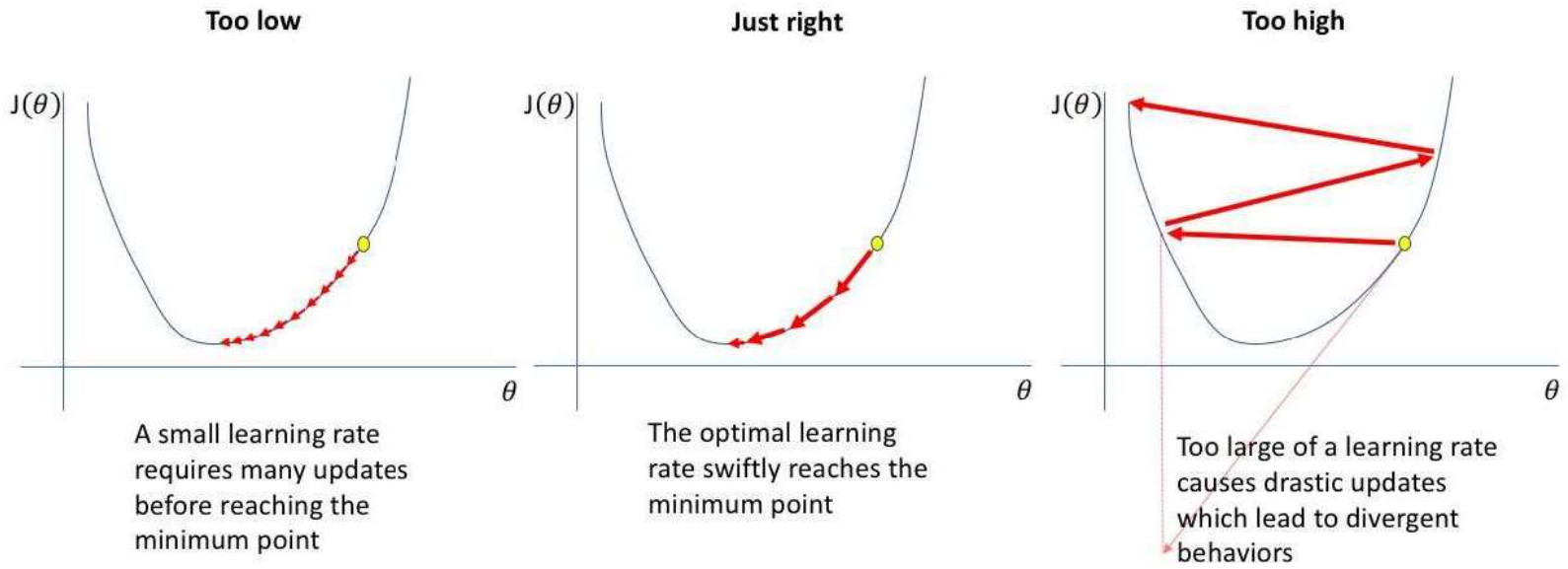


Divergence



Learning Rate

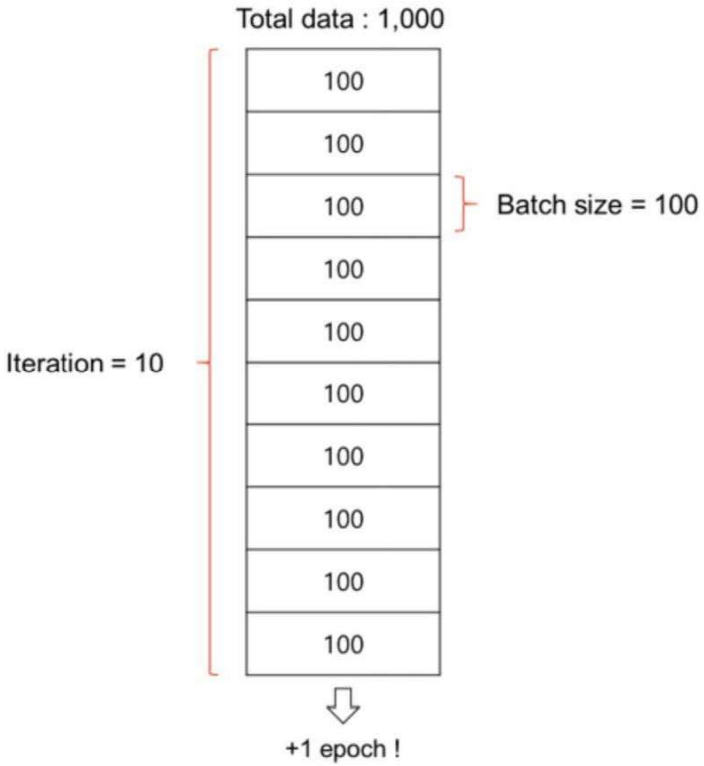
A **higher learning** rate might lead to **faster convergence** but **risks overshooting**, while a **lower rate** might slow down convergence.



Batch Size

Batch is a **subset** of the **training dataset** used in **each iteration** of the training process.

Instead of processing the **entire dataset** at once, we **divide** it into **smaller batches**.



Batch Size

Batch size is the number of **training examples** in each **batch.**

It's a **hyperparameter** that can be adjusted based on **hardware limitations** and **dataset characteristics**.

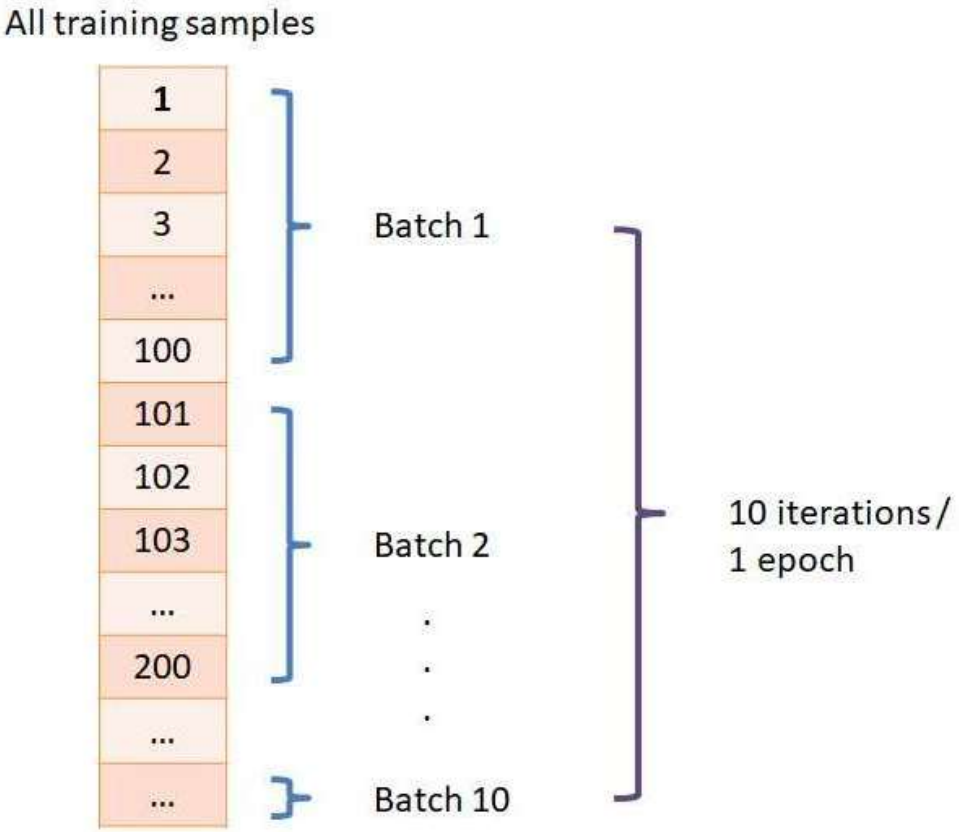
Larger batch sizes may **speed** up **training** but require **more memory**.

Number of Epochs

The number of **times** the entire training **dataset** is **seen** by the **model** during **training**.

Too **few epochs** might result in **underfitting**, while too **many epochs** can lead to **overfitting**.

Number of Epochs



Activation Functions (Neuron)

It determine the functions applied to the outputs of each neuron to **introduce non-linearity**.

Common activation functions include

- **ReLU (Rectified Linear Unit),**
- **Softmax**