ZUJO

Session 1: Introduction to Machine Learning Part 1

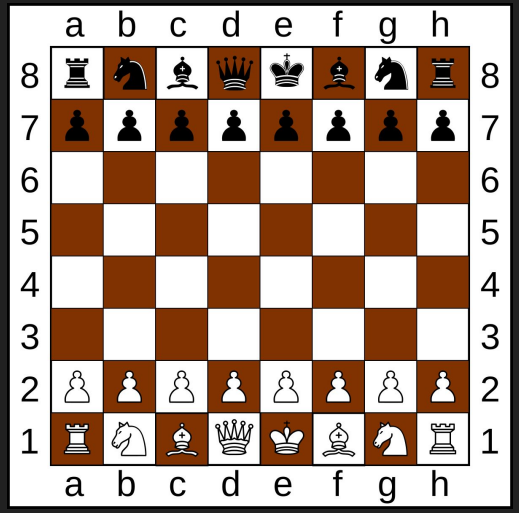
Index

- Introduction to machine learning
- Comparing with human
- Understanding of supervised learning
- Linear Regression

What is Machine Learning?

Machine learning is field of study that gives computers the ability to learn without being explicitly programmed.

- Arthur Samuel (1959)

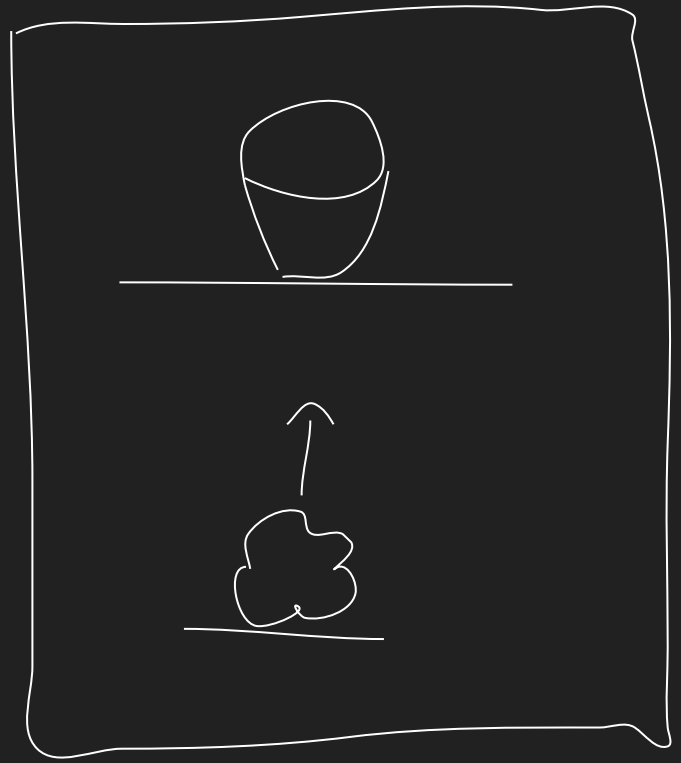


What is Machine Learning?

A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.

- Tom Mitchell (1998)

In simple terms



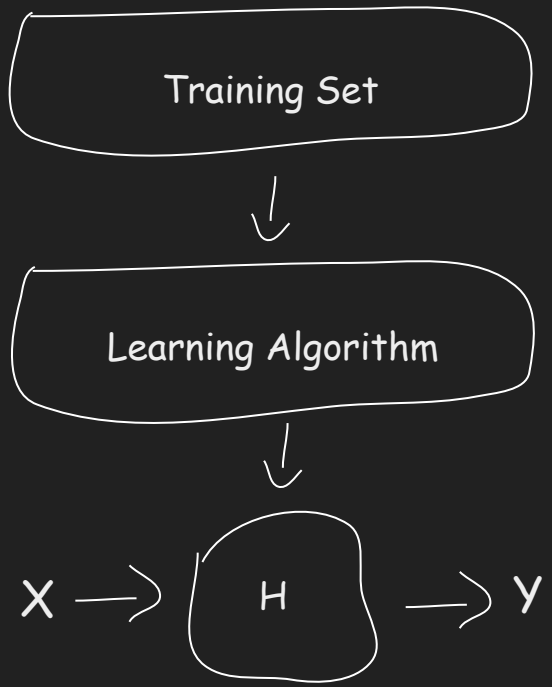
1st Attempt: You have to put too much force.

2nd Attempt: You realise you are closer to target but you need to increase your throw angle.

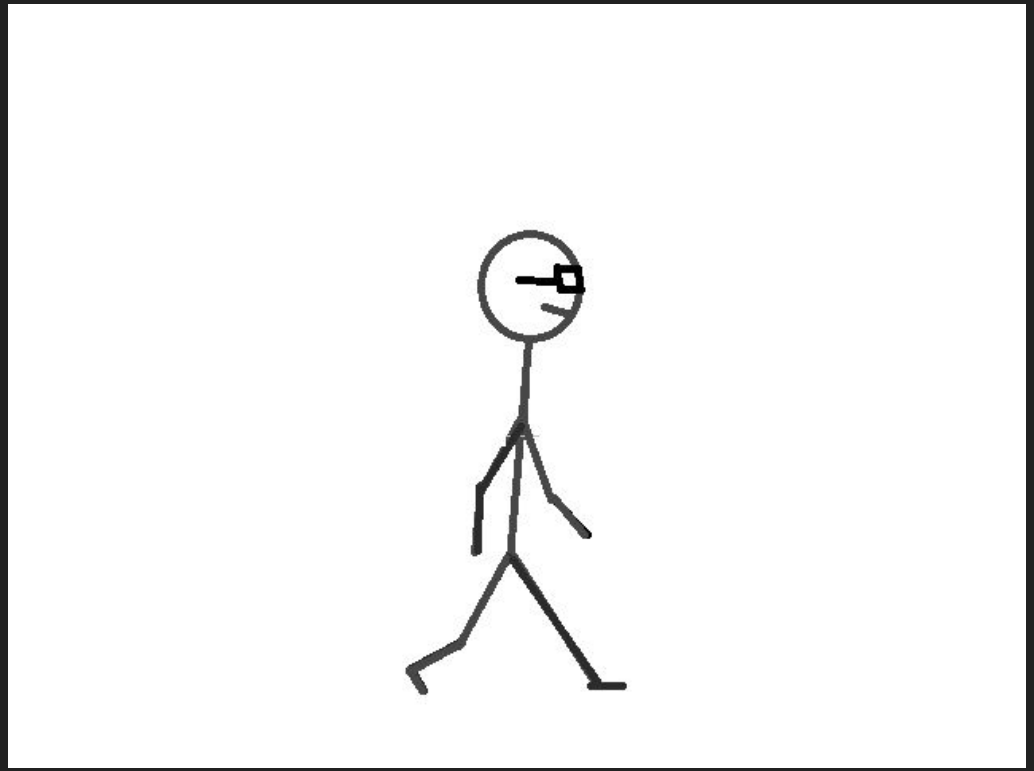
What is happening here is basically after every throw we are learning something and improving the end result.

We are programmed to learn from our experience.

Learning Process



Compare it with Humans



Human

Speak & Listen

AI

Statistical Learning

Speech Recognition

Human

Write and Read

AI

Statistical Learning

Speech Recognition

NLP

London is the capital and most populous city of England and the United Kingdom. Standing on the River Thames in the south east of the island of Great Britain, London has been a major settlement for two millennia. It was founded by the Romans, who named it Londinium. London's ancient core, the City of London, largely retains its 1.12-square-mile (2.9 km2) medieval boundaries.

Human

See

AI

Statistical Learning

Symbolic Learning

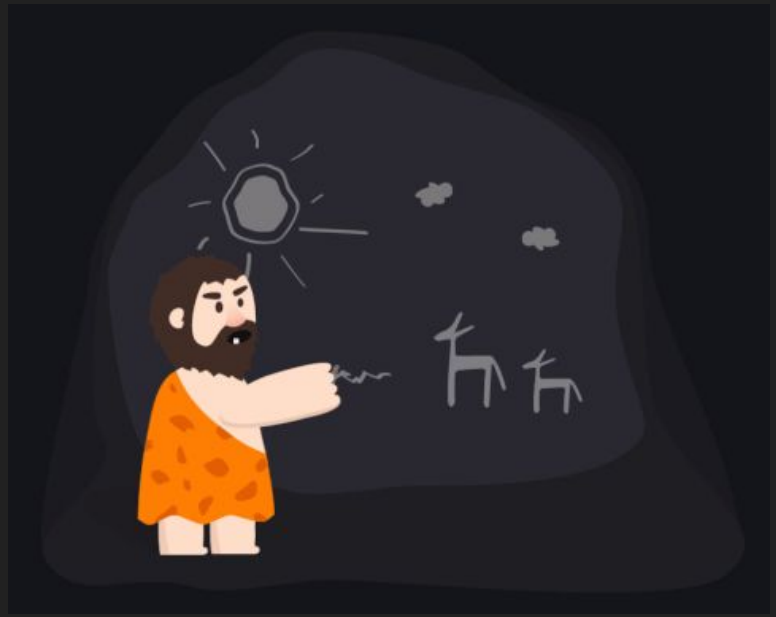
Speech Recognition

NLP

Computer Vision

Human

Process and Visualise



AI

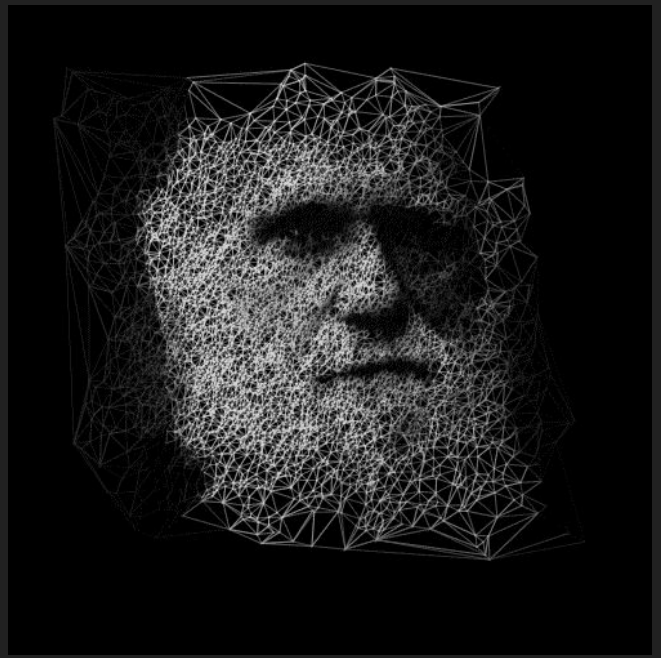
Statistical Learning

Speech Recognition NLP

Image Processing

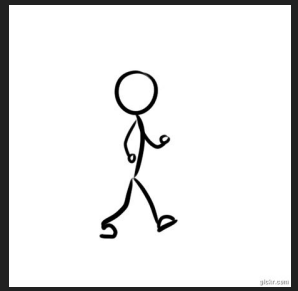
Symbolic Learning

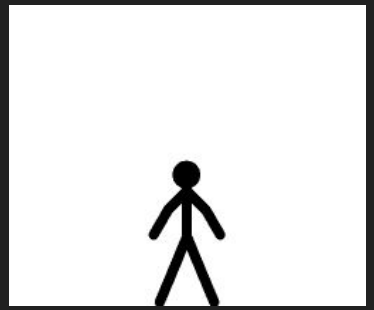
Computer Vision



Human

Understand Environment





AI

Image Processing

Statistical Learning

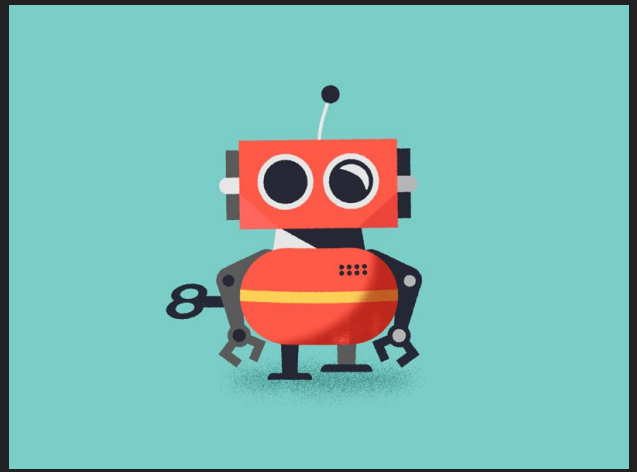
Symbolic Learning

Speech Recognition

NLP

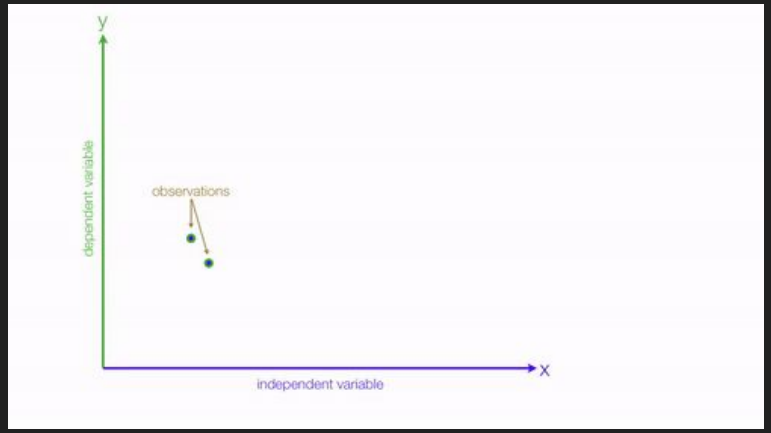
Computer Vision

Robotics



Human

Recognise Patterns



AI

Machine Learning

Pattern Recognition

Image Processing

Symbolic Learning

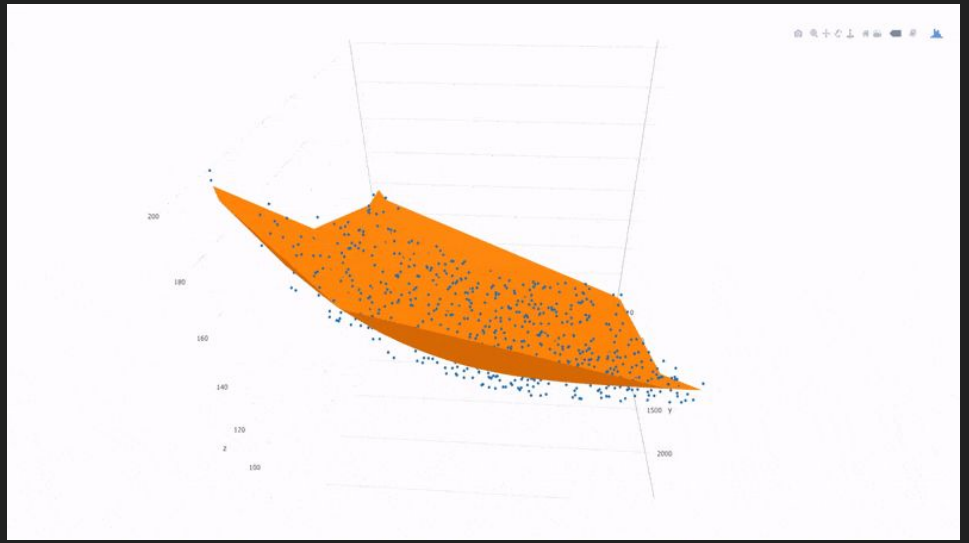
Robotics

Computer Vision

Statistical Learning

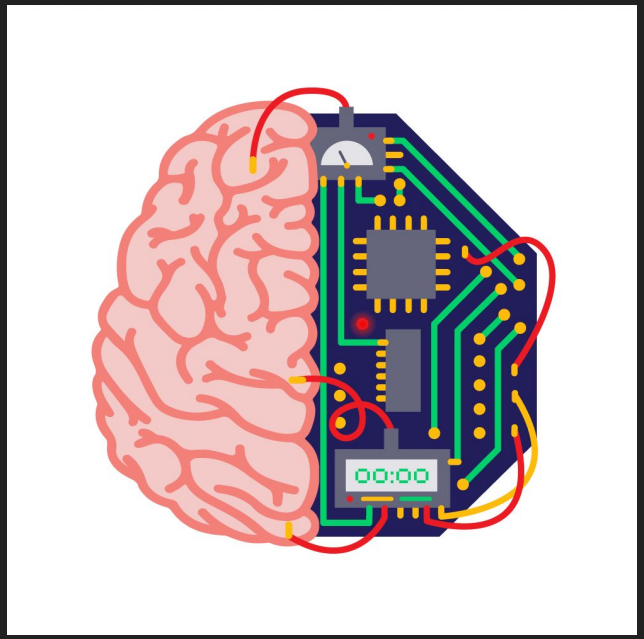
Speech Recognition

NLP



Human

Brain



AI

Pattern Recognition

Machine Learning

Image Processing

Symbolic Learning

Statistical Learning

Robotics

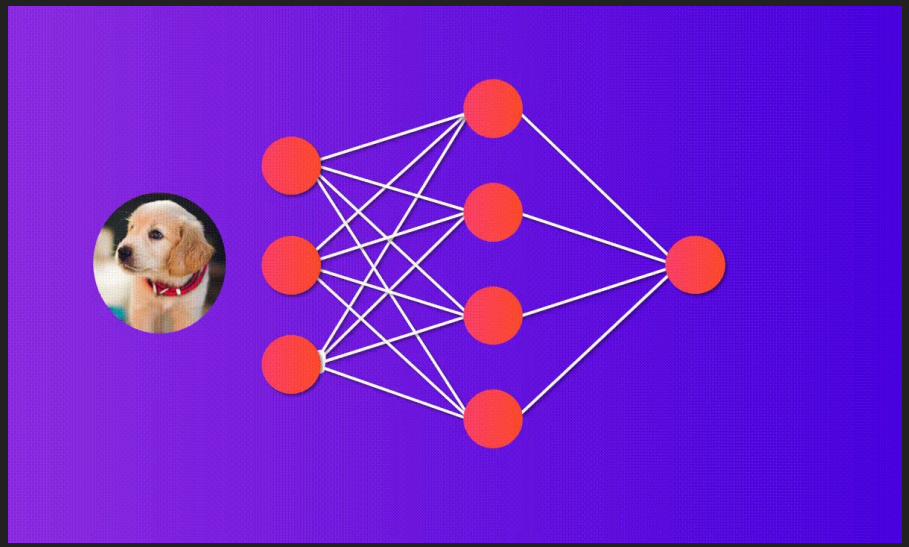
Computer Vision

Speech Recognition

NLP

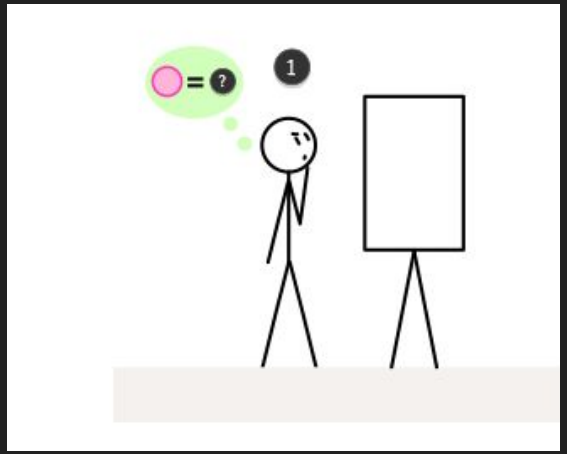
Deep Learning

Neural Networks



Human

Remember Scene



Pattern Recognition

AI

Machine Learning

Image Processing

NN

Symbolic Learning

Statistical Learning

Deep Learning

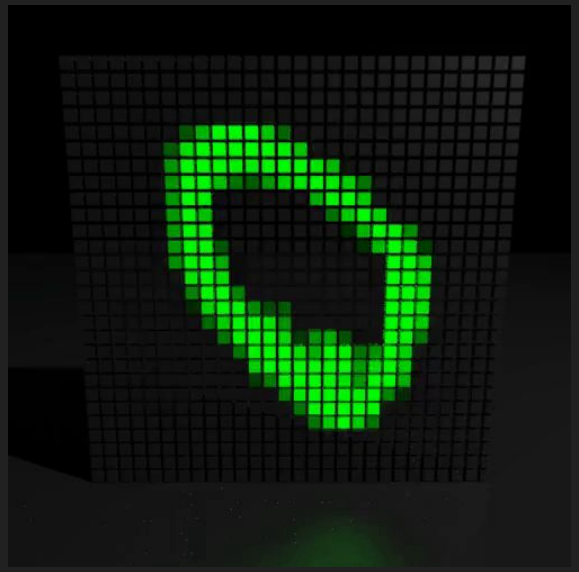
Robotics

Computer Vision

Speech Recognition

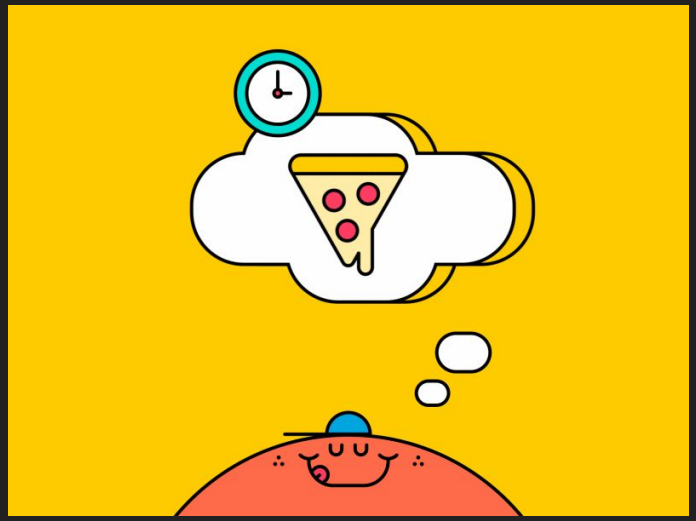
NLP

CNN



Human

Remember Past



Pattern Recognition

AI

Machine Learning

Image Processing

NN

Statistical Learning

Deep Learning

Symbolic Learning

Robotics

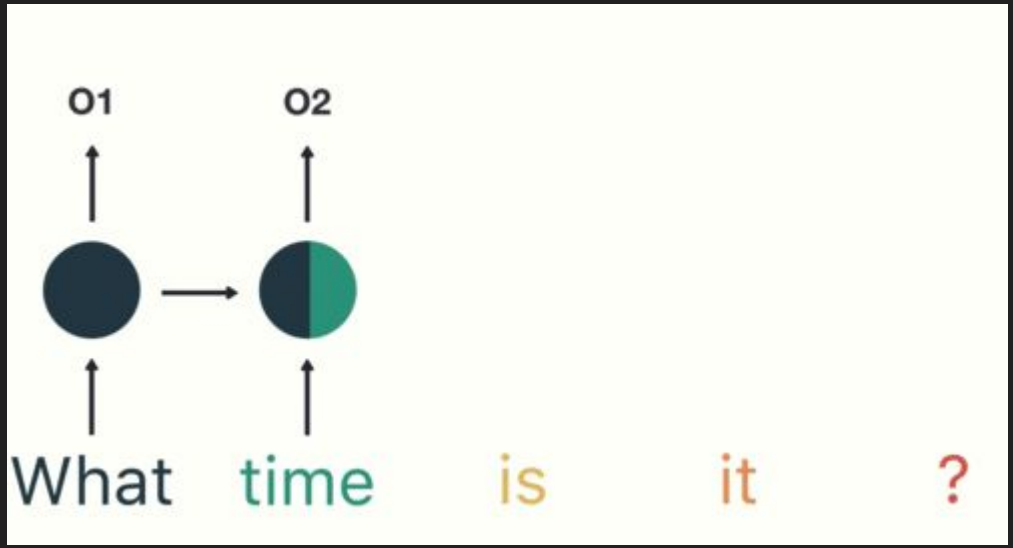
Computer Vision

Speech Recognition

NLP

CNN

RNN



Human

Dream

Pattern Recognition

AI

Machine Learning

Image Processing

NN

Symbolic Learning

Statistical Learning

Deep Learning

Robotics

Computer Vision

Speech Recognition

NLP

CNN

RNN

GAN

Pattern Recognition

Machine Learning

Statistical Learning

Deep Learning

Speech Recognition

NLP

CNN

RNN

GAN

Classification

Prediction

AI

Image Processing

Symbolic Learning

Robotics

Computer Vision

Supervised Learning

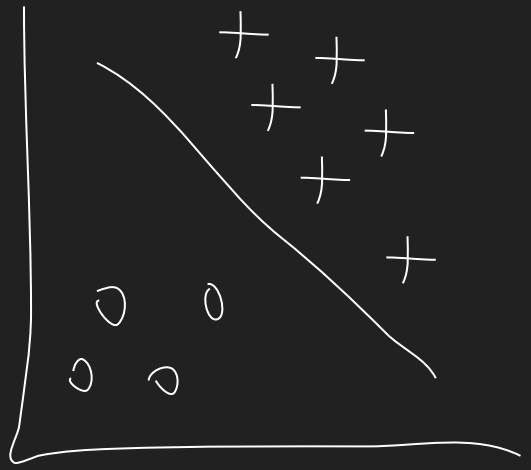
Unsupervised Learning

Reinforcement Learning

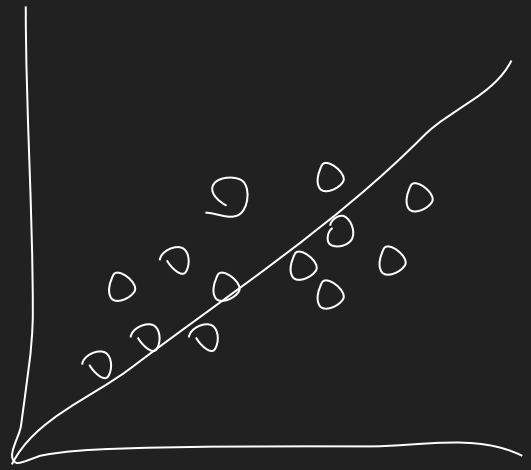
Understanding each type

- Supervised Learning
- Unsupervised Learning
- Semi-supervised Learning
- Reinforcement Learning

Supervised Learning

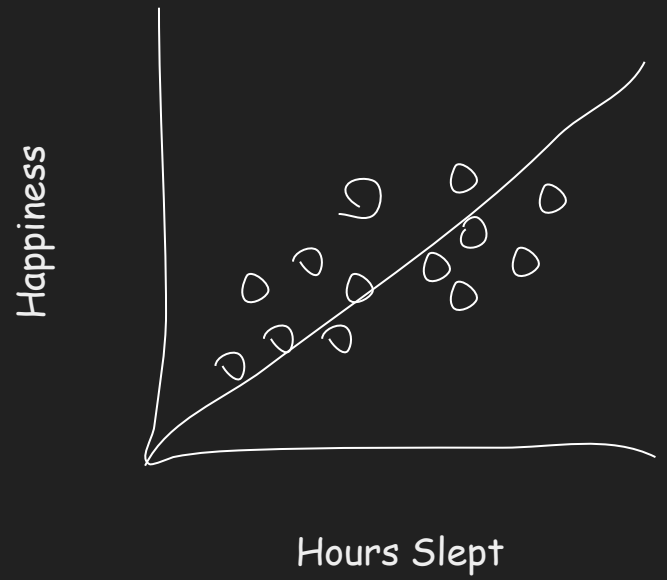


Classification

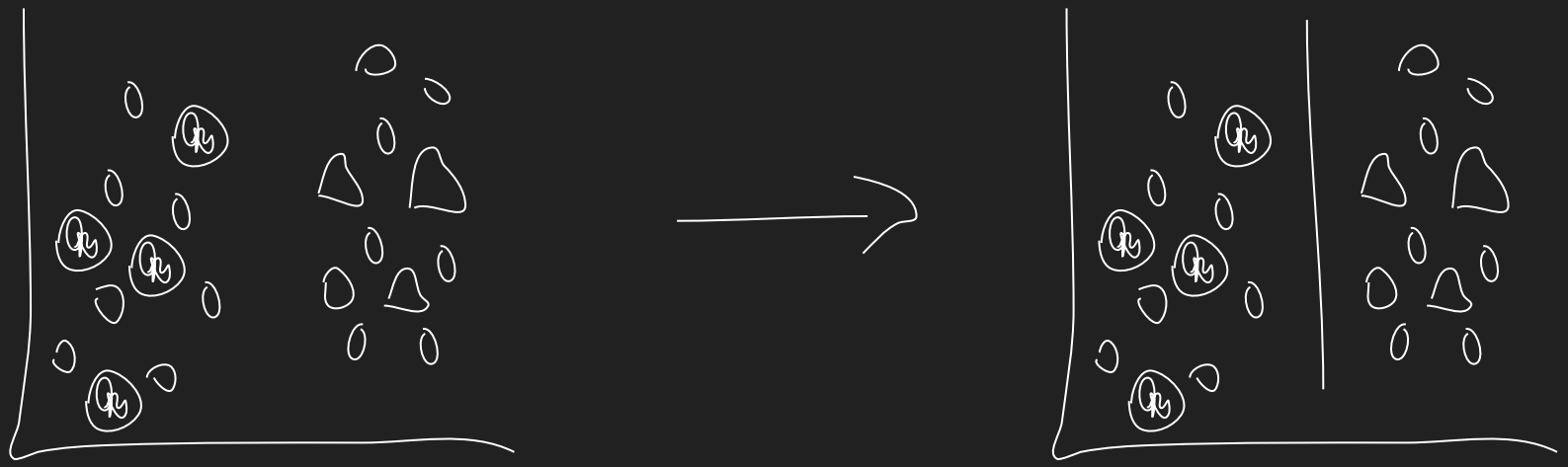


Regression

Supervised Learning Example



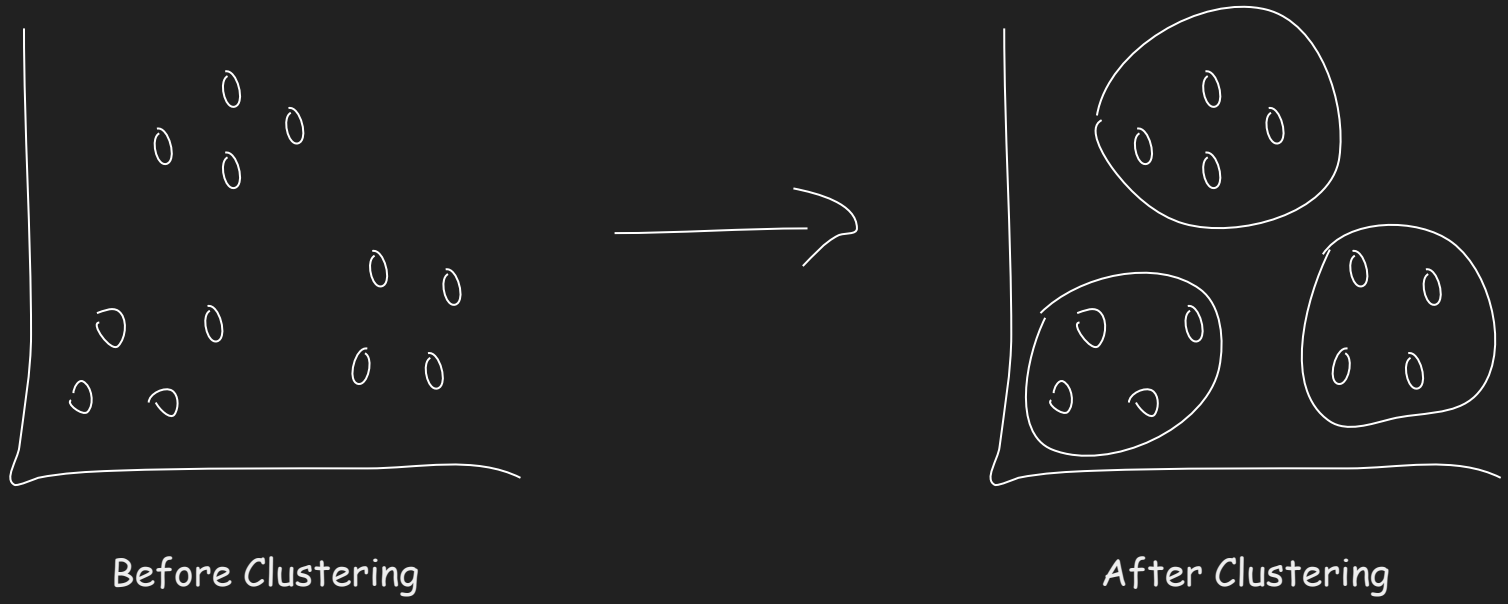
Semi-supervised Learning



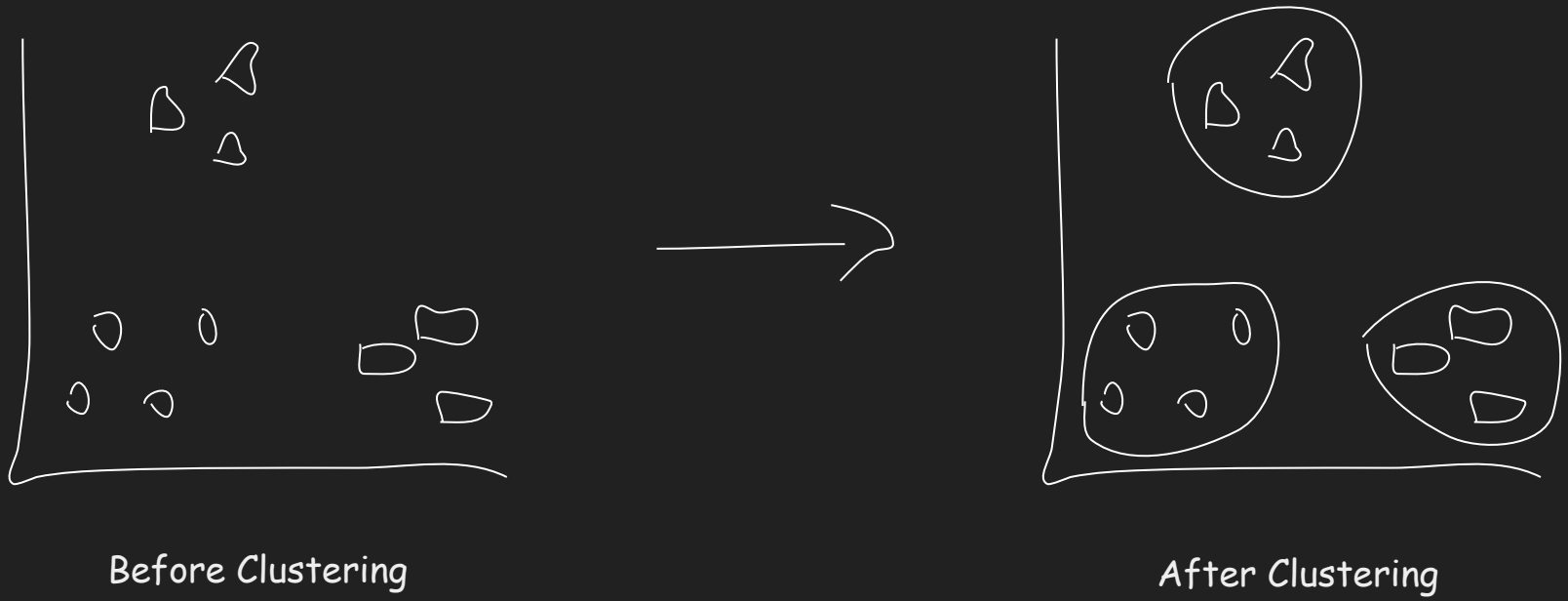
Labeled and Unlabeled data

Semi-supervised Learning

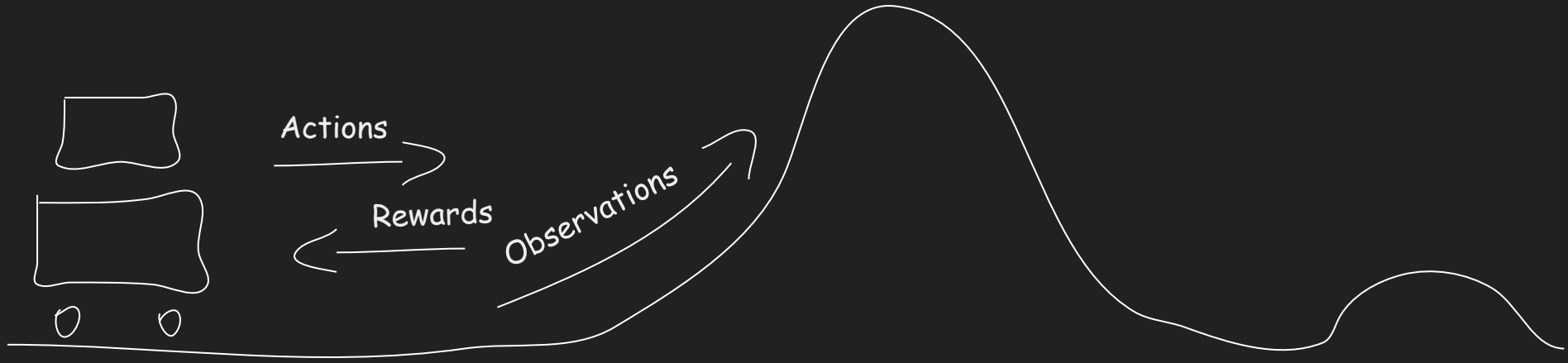
Unsupervised Learning



Unsupervised Learning Example



Reinforcement Learning



Supervised Learning

- **Linear regression**
- Logistic regression
- Naive bayes
- **Support vector machine**
- Linear discriminant analysis
- Decision trees
- K-nearest neighbor
- **Neural networks (Multi-layer perceptron)**
- Similarity learning

Unsupervised Learning Algorithms

- K-means clustering
- KNN (k-nearest neighbors)
- Hierarchical clustering
- Anomaly detection
- **Neural Networks**
- Principal Component Analysis
- Independent Component Analysis
- Apriori algorithm
- Singular value decomposition

Semi-supervised Learning Algorithms

- Generative models
- Low-density separation
- Graph-based methods
- Heuristic approaches

Reinforcement Learning Algorithms

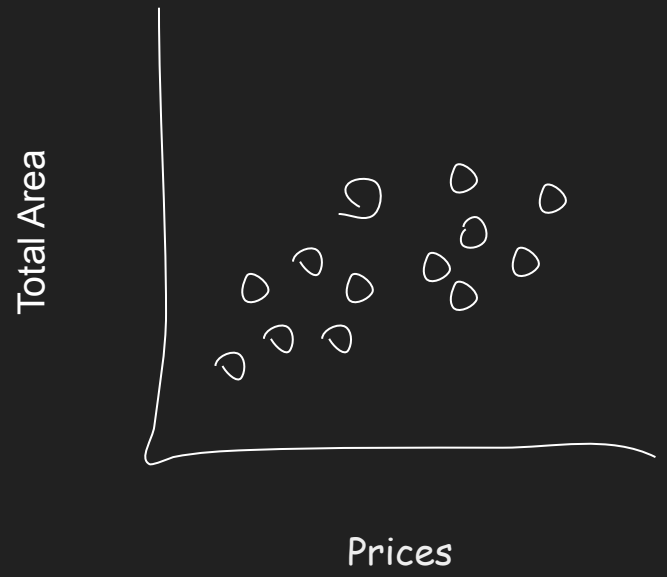
- Monte Carlo
- Q-learning
- SARSA (State–action–reward–state–action)

Linear Regression

Dataset

Total Area	No of bedrooms	Prices (Lakh)
1056	2 BHK	39.07
2600	4 Bedroom	120
1440	3 BHK	62
1521	3 BHK	95
1200	2 BHK	51
1170	2 BHK	38
2732	4 BHK	204
3300	4 BHK	600

Dataset



Hypothesis

$$h(x) = \theta_0 + \theta x1 + \theta x_2$$

$$x_1 = Size$$

$$x_2 = Bedrooms$$

Hypothesis

$$h(x) = \sum_{j=1}^{2} \theta_j x_j$$

Where, $x_0 = 1$

$$\theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{bmatrix} \quad x = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix}$$

$\theta = parameters$

$m = \#\ training\ examples$

$x = inputs$

$y = output$

$(x, y) = training\ example$

$(x^{(i)}, y^{(i)}) = n^{th}\ training\ example$

$n = \#\ features$

How to choose theta?

$$h_\theta(x) = \theta_0 + \theta_1 x$$

Hypothesis

Parameters → $(\theta_0 + \theta_1)$

Simplified Math

Predicted Output

Actual Output

Row Index

Cost Function →

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

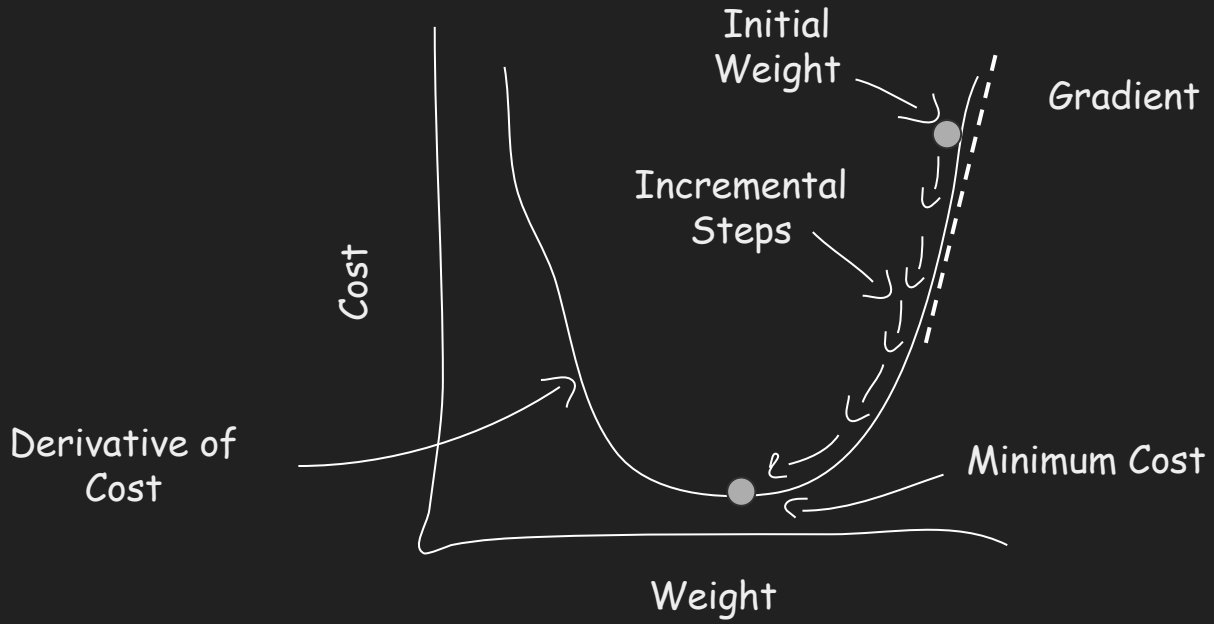
Summation of all the samples

Total Squared Error

Goal → $Minimize\ J(\theta)$

Gradient Descent

Gradient descent is an iterative optimization algorithm for finding the local minimum of a function.



Gradient Descent Algorithm

The goal of the gradient descent algorithm is to minimize the given function (say cost function). To achieve this goal, it performs two steps iteratively:

1. Compute the gradient (slope), the first order derivative of the function at that point
2. Make a step (move) in the direction opposite to the gradient, opposite direction of slope increase from the current point by alpha times the gradient at that point

Gradient Descent Algorithm

$$repeat\ until\ convergence\{$$

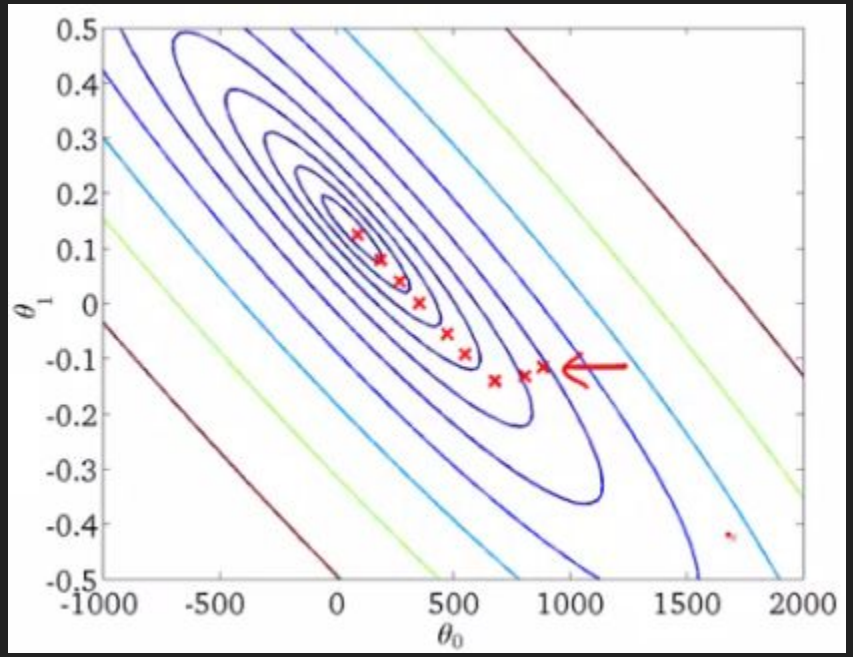
$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0 . \theta_1)$$

$$(for\ j = 1\ and\ j = 0)$$

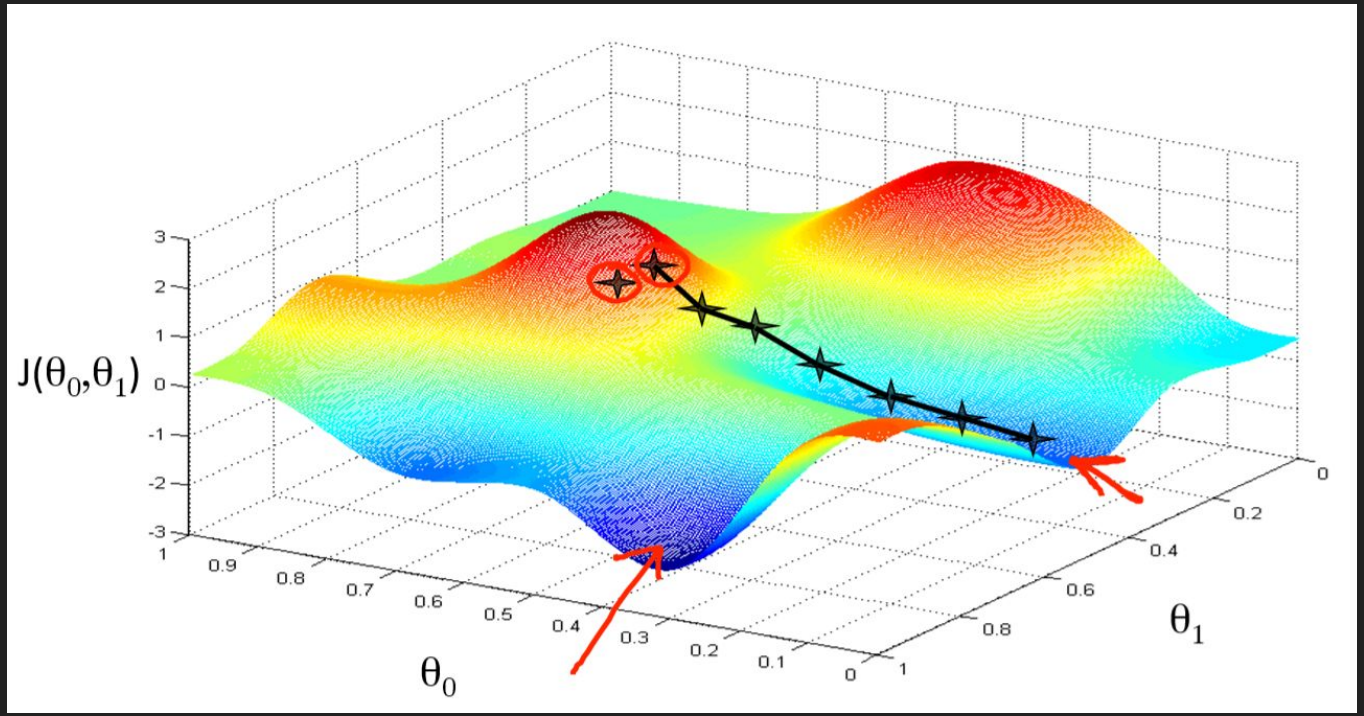
$$\}$$

Learning Rate

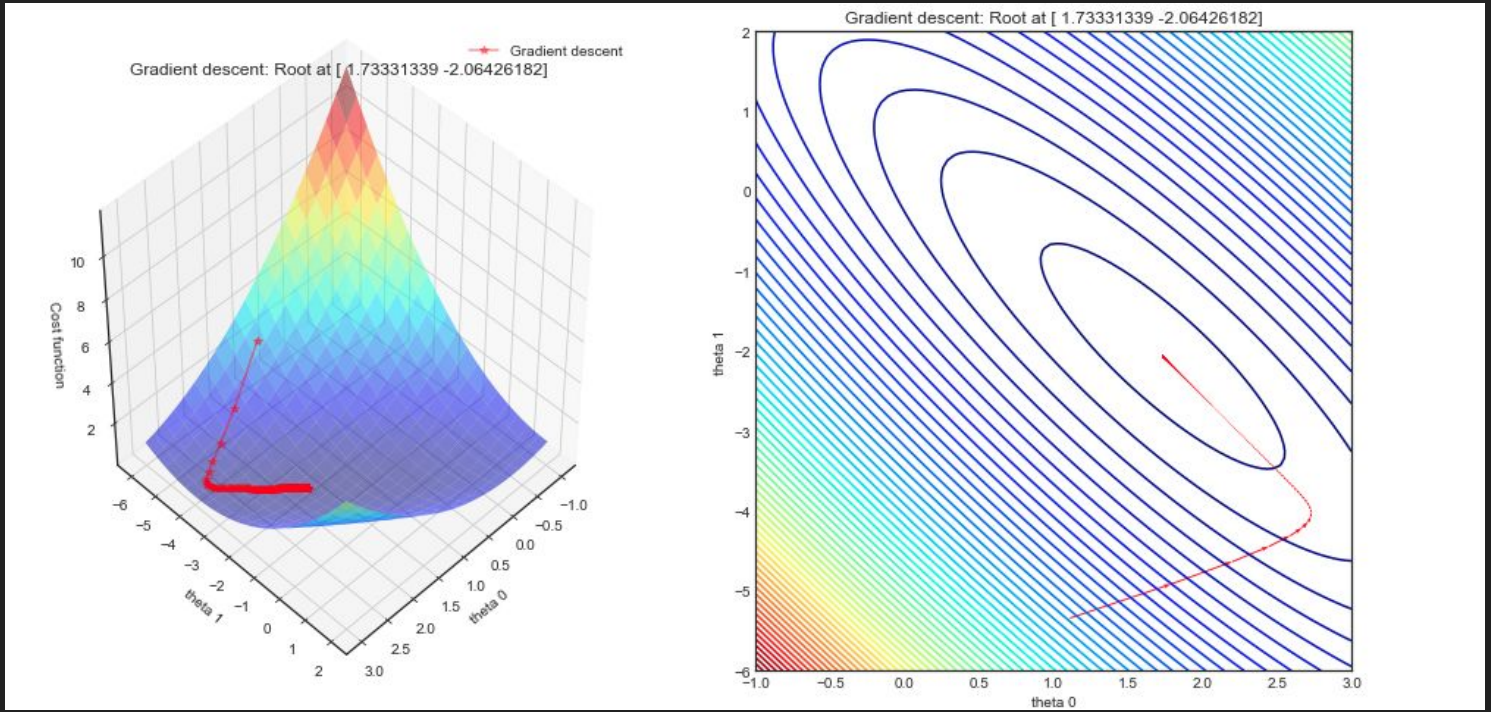
Gradient Descent Algorithm



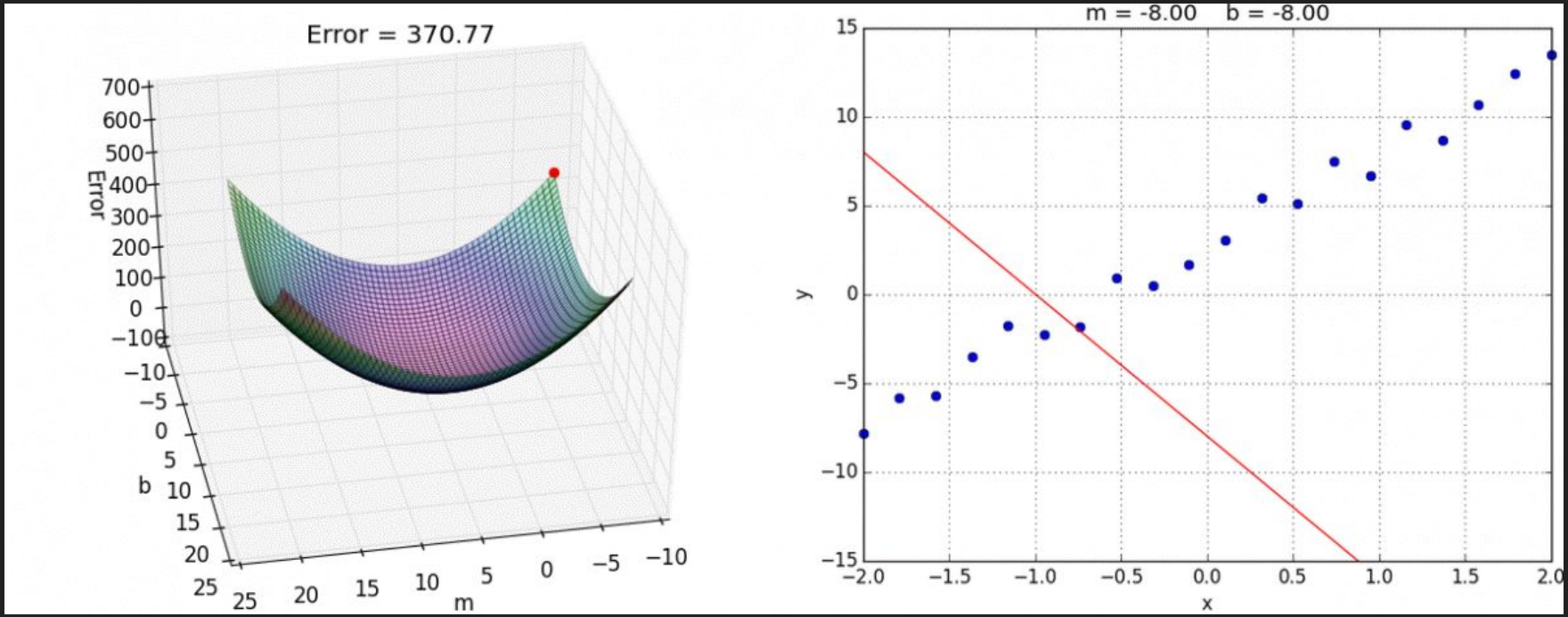
Gradient Descent Algorithm



Gradient Descent Algorithm



Gradient Descent Algorithm



Q & A

Gradient Descent Algorithm Hands On

- Visualization
  - https://bl.ocks.org/EmilienDupont/aaf429be5705b219aaaf8d691e27ca87
- Example
  - https://github.com/arjun-kava/linear-regression-with-gradient-descent
- Notebook
  - https://colab.research.google.com/drive/1Dxs5fu3ELxF3B4P_Our6hO-pi7UQbLOH?usp=sharing