

Neural Networks - I

Introduction to
Artificial Intelligence

Understanding
Gradient Descent

Introduction to
TensorFlow

Building a Model
using TensorFlow

Day - I

Using Keras with
TensorFlow

Building Classification
Model in Keras

Understanding
Deep Learning

Building Neural
Network using Keras

Day - II



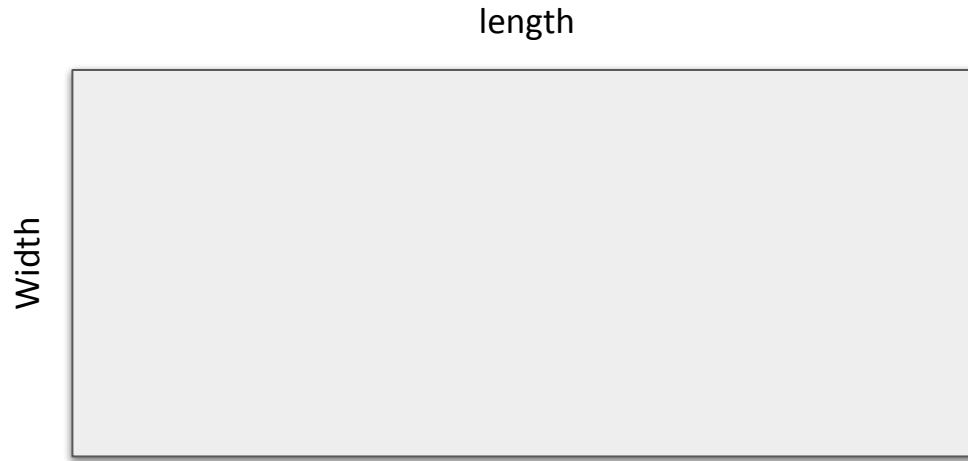
Working with Machines

How do we make Computer do things?

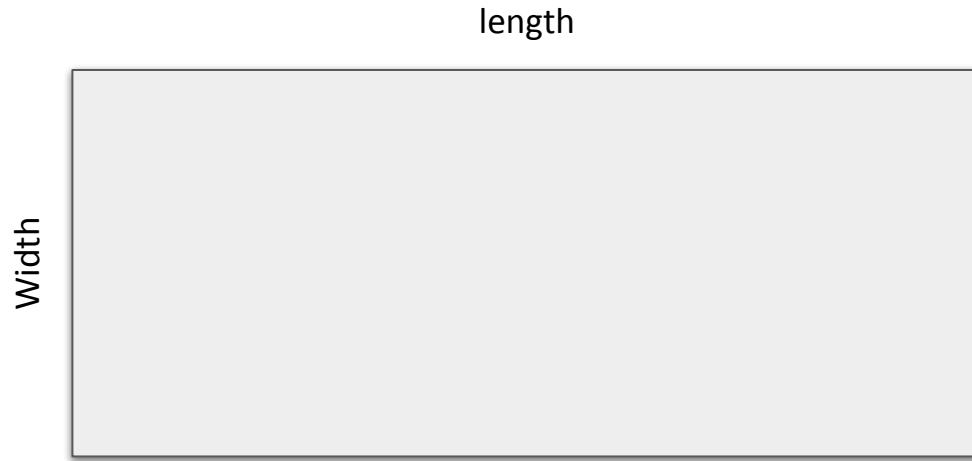
Requirement



Area of a Rectangle



Area of a Rectangle



$$Area = L * W$$

Write down steps (Algorithm)

```
START
    DISPLAY "Enter length of side one: "
    GET side1
    DISPLAY "Enter length of side two: "
    GET side2
    area = side1 * side2
    DISPLAY area
END
```

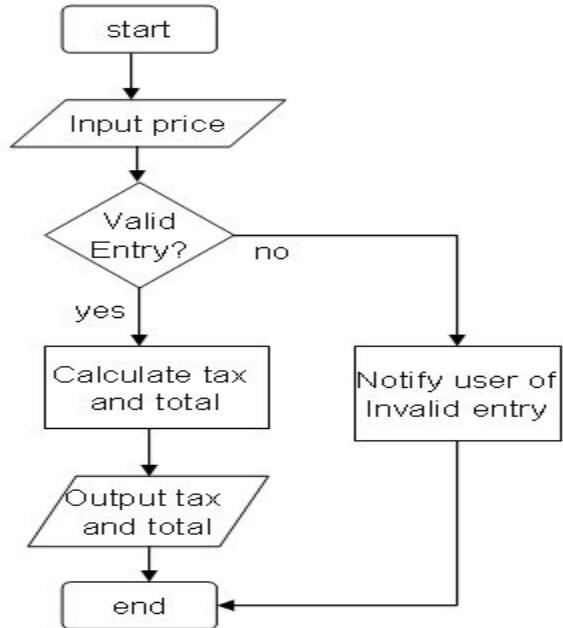
Feed it to the Computer



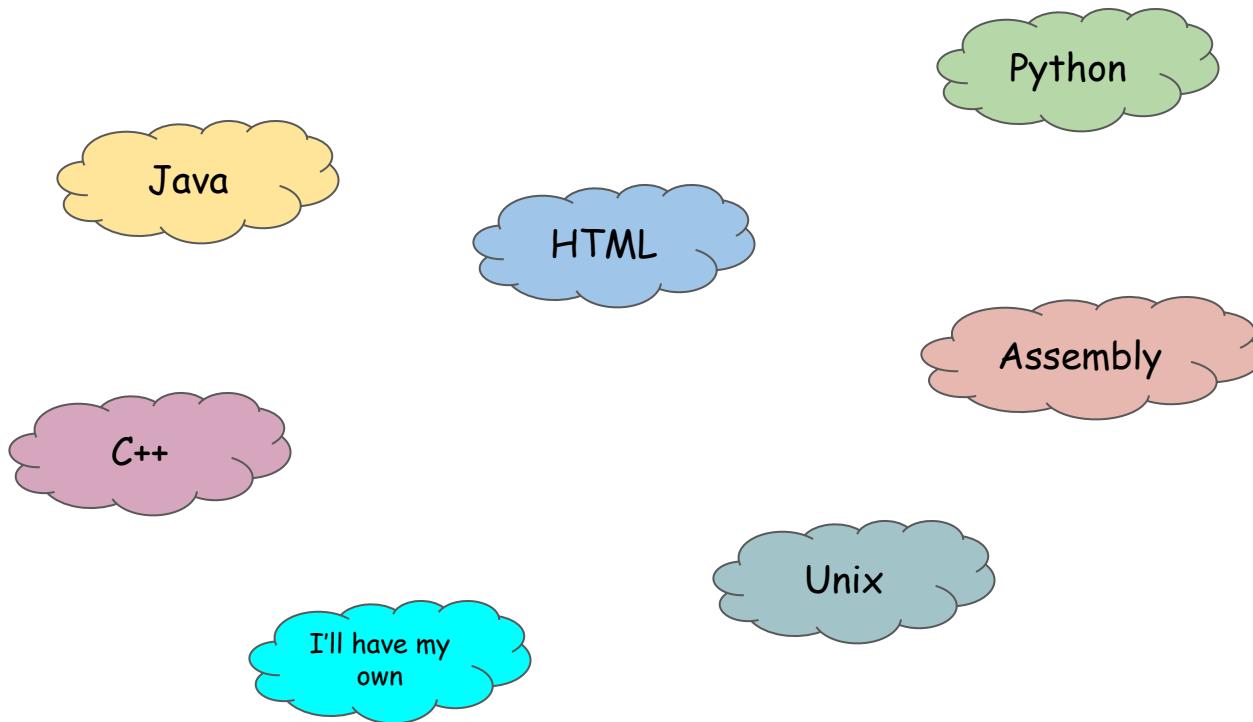


Calculate
Sales Tax

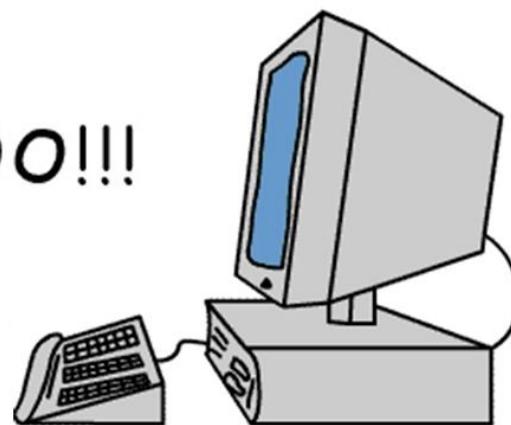
Algorithm



Write code!



WOOHOO!!!



Working with Machines

- ❑ Get the **requirement**

Working with Machines

- ❑ Get the **requirement**
- ❑ Build the **Logic**

Working with Machines

- ❑ Get the **requirement**
- ❑ Build the **Logic**
- ❑ Convert logic into a **Program**

Working with Machines

- ❑ Get the **requirement**
- ❑ Build the **Logic**
- ❑ Convert logic into a **Program**
- ❑ Give it to Computer

Visualizing the World



What do you see?



How about this?



Is it same?



This is very
easy :)

Let's write down the Steps...to recognize Elephant





Get the **requirement**



Build the **Logic**



Convert logic into a **Program**



Give it to Computer



How does our brain understand it?



Can a machine have
Vision?

Case Study: Google Photos



Can a machine understand
Natural Language?

Case Study: Google Translate

Sometime
even having
logic ...

may not help

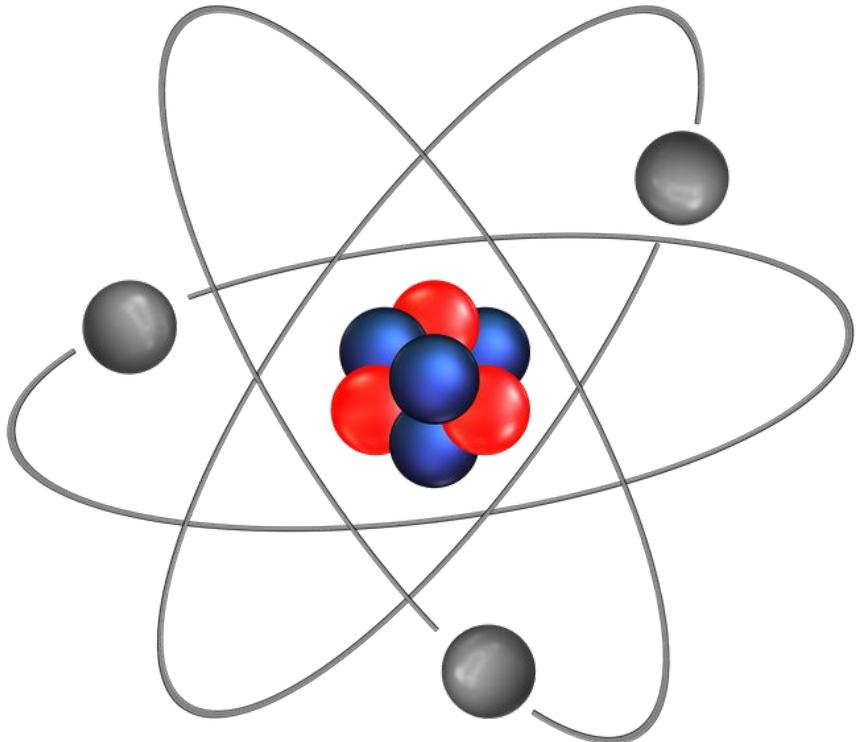






Possible moves in
Chess

10^{123}



Possible atoms in
Observable Universe

10^{80}

In 1997, IBM's Deep Blue machine defeated Gary Kasparov



In 1997, IBM's Deep Blue machine defeated Gary Kasparov



...using huge computing capacity.



Possible moves in Go

10^{360}

Possible moves in Go

10^{360}

We can not apply brute computing capacity :(

In Mar 2016, Google's AlphaGo defeated Lee Sedol



In Mar 2016, Google's AlphaGo defeated Lee Sedol



...by 'learning' to play 'Go'



Recognize objects



Understand Language



Play complex games

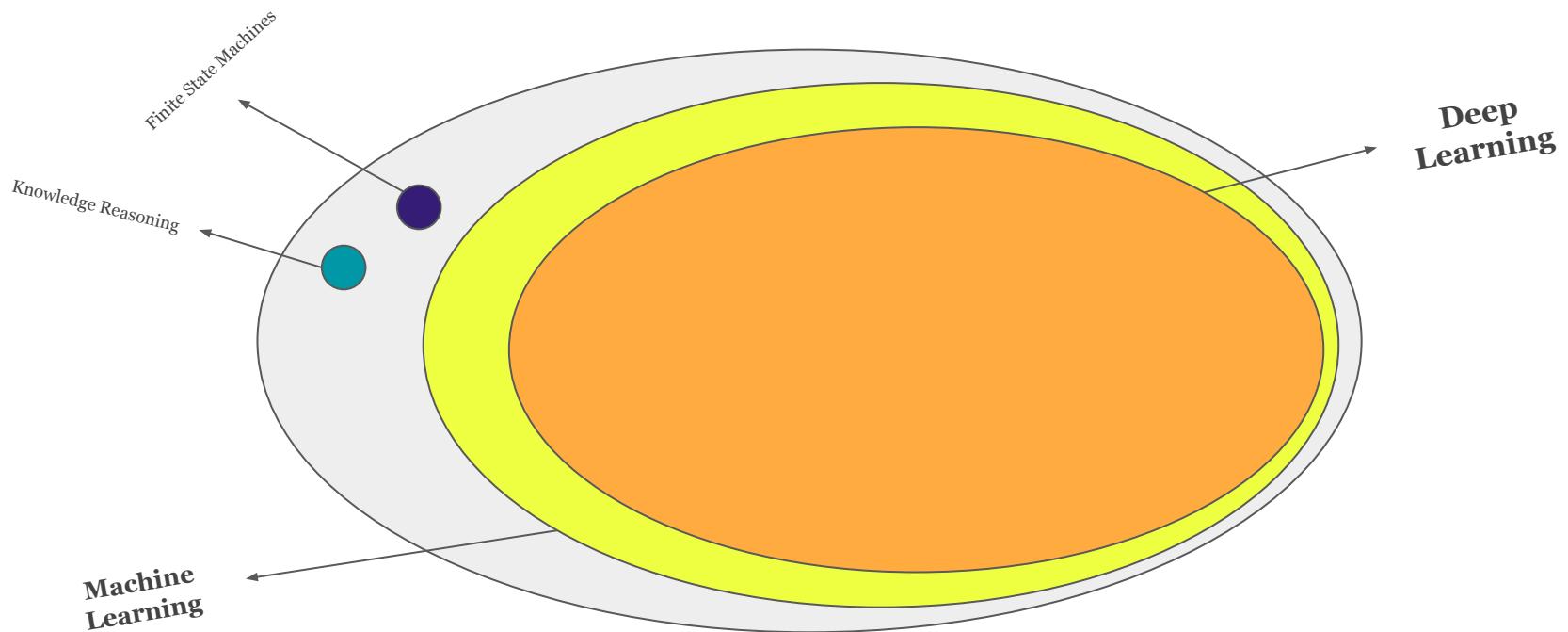
Machine demonstrates an ability to...
Think, Reason

Artificial Intelligence



**How do we build
AI?**

AI World





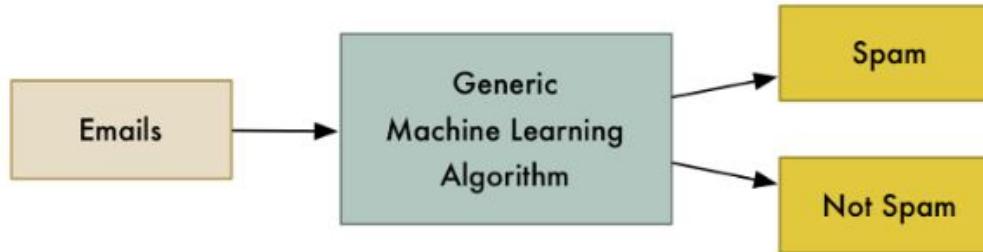
What is
Machine Learning?

Machine learning (ML) allows computers to do things....

without our logic

Computer program themselves

Instead of writing code, we feed the data to the Generic algorithms

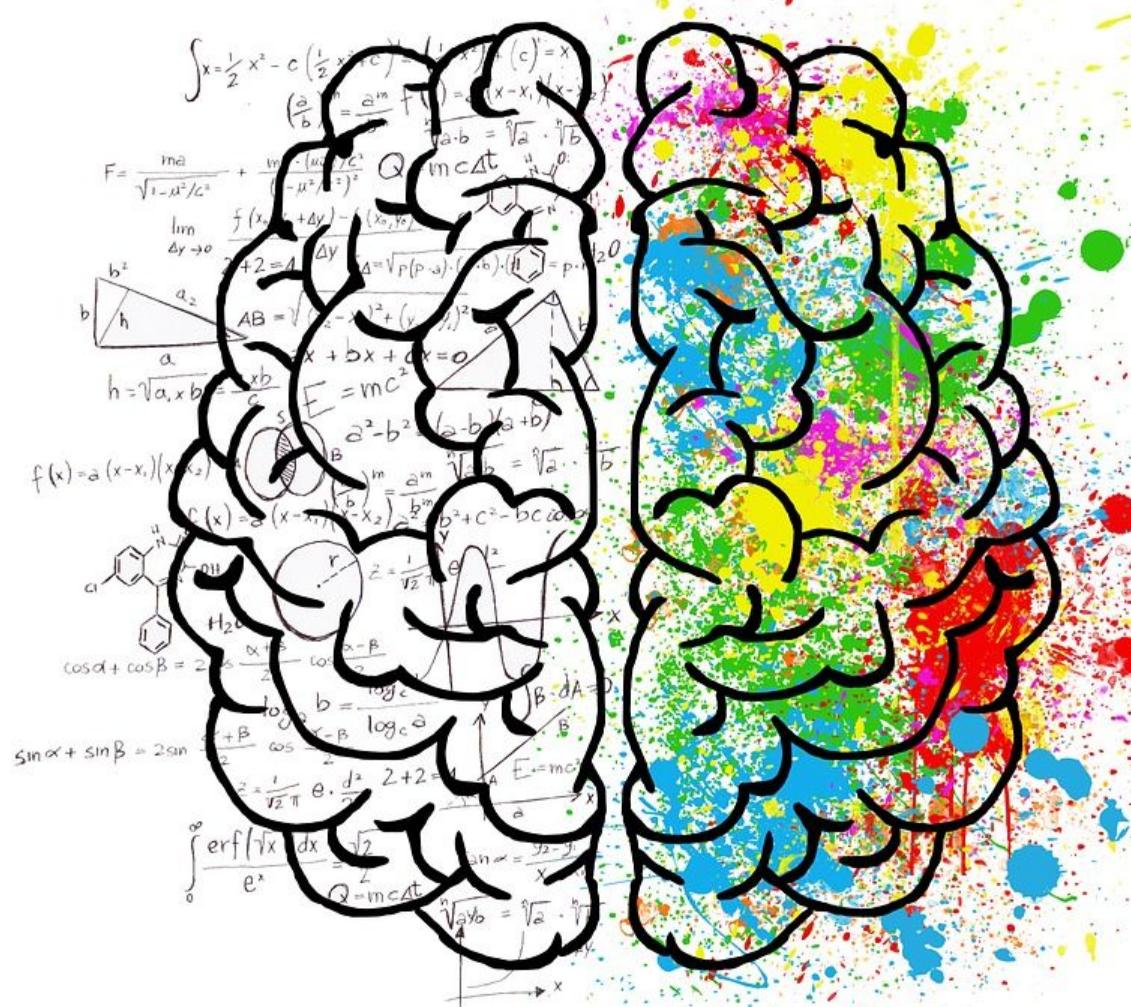


...and it builds its own logic based on the data.



How did we 'humans'
learn to recognize
Elephants?

How Machine builds logic



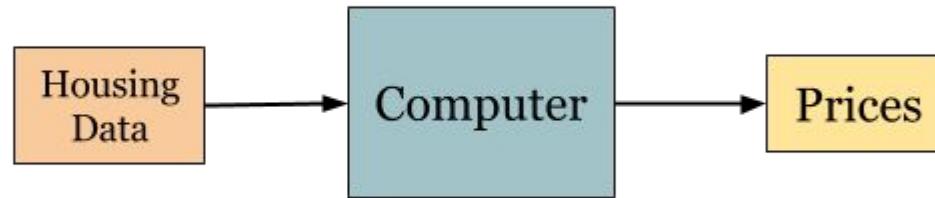
Sq. Ft	Neighbourhood	Bedrooms	Price ('ooo)
2000	Gachibowli	3	180
1750	Jubilee Hills	3	210
1100	Kukatpally	2	55
900	Gachibowli	2	72
1245	KPHB	3	60
...
...

Real-Estate Data

Requirement

Sq. Ft	Neighbourhood	Bedrooms	Price ('ooo)
1250	Gachibowli	3	???

Machine to build its own program



What will Machine try to build?

$$\text{Price} = W_1 * \text{Sq. Ft.} + W_2 * \text{Bedrooms} + W_3 * \text{Neighbourhood}$$

$$\text{Price} = W_1 * \text{Sq. Ft.} + W_2 * \text{Bedrooms} + W_3 * \text{Neighbourhood}$$

$$y = W_1 * x_1 + W_2 * x_2 + W_3 * x_3$$

ML Terminology

- ❑ Weights : W_1, W_2, W_3
- ❑ Features : x_1, x_2, x_3
- ❑ Target / Prediction / Output : y

Handling bias in Data

$$y = W_1 x_1 + W_2 x_2 + W_3 x_3 + b$$



How do I choose
Weights and bias?

$$y = W_1 x_1 + W_2 x_2 + W_3 x_3 + b$$

Pick random values

Sq. Ft	Bedrooms	Neighbourhood	Actual Price	Machine Price	Loss
2000	3	Gachibowli	160	118	42
1750	3	Jubilee Hills	230	269	-39
1100	2	Kukatpally	55	67	-12
900	2	Gachibowli	92	83	9
1245	3	KPHB Colony	60	60	0
Total					0

Is machine's logic good?

Another way to calculate Loss

$$Loss = \frac{\sum_{i=1}^m (Prediction_i - Actual_i)^2}{2 \times m}$$

$m \rightarrow$ number of examples

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

More complex way :)

Sq. Ft	Bedrooms	Neighbourhood	Actual Price	Machine Price	Loss
2000	3	Gachibowli	160	118	(42) ²
1750	3	Jubilee Hills	230	269	(-39) ²
1100	2	Kukatpally	55	67	(-12) ²
900	2	Gachibowli	92	83	(9) ²
1245	3	KPHB Colony	60	60	(0) ²
Total					3510

$$Loss = 3510 / (2 * 5) = 351$$



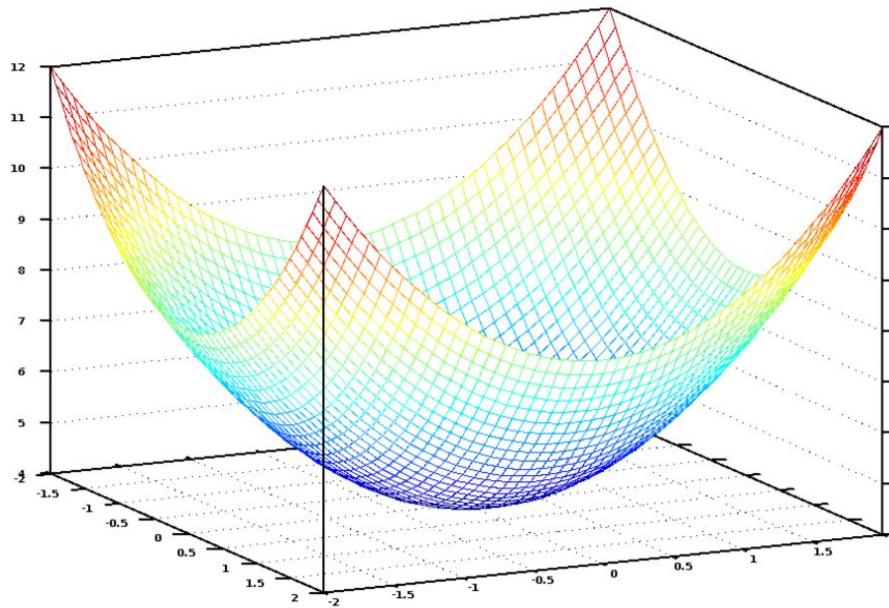
What is the machine's goal?

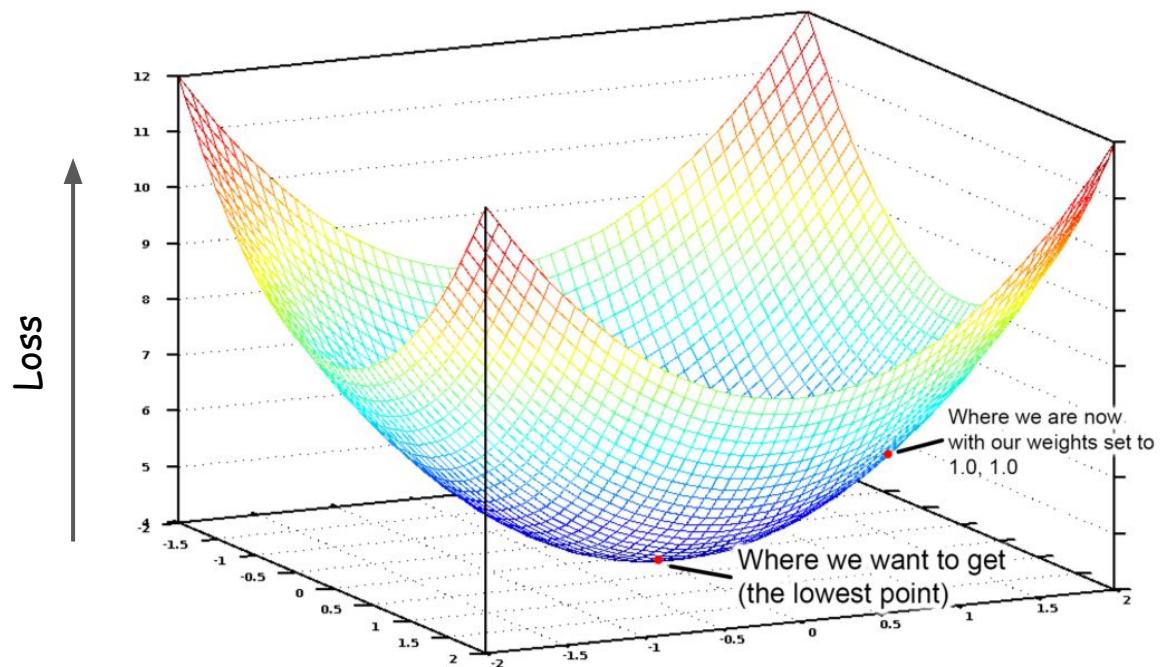
Find weights and bias to minimize ‘Loss’

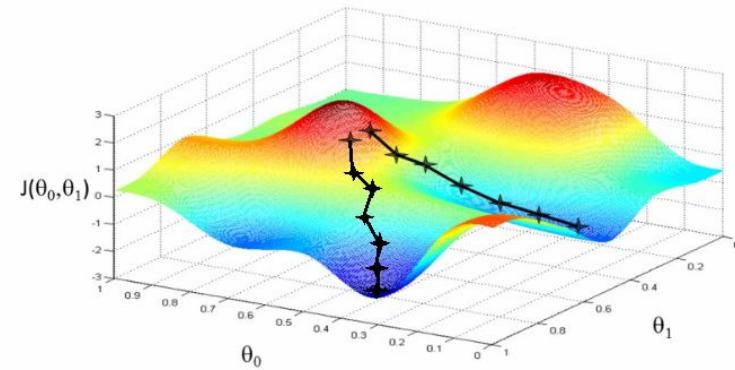
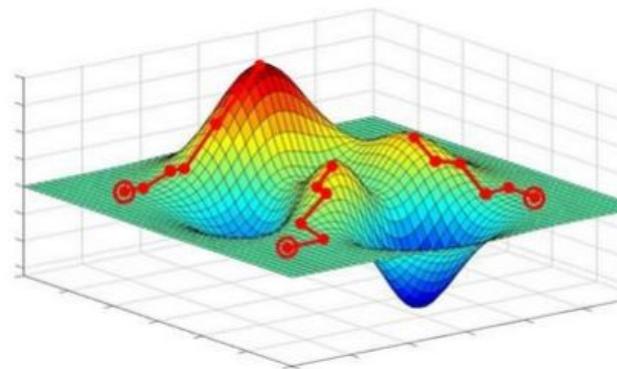
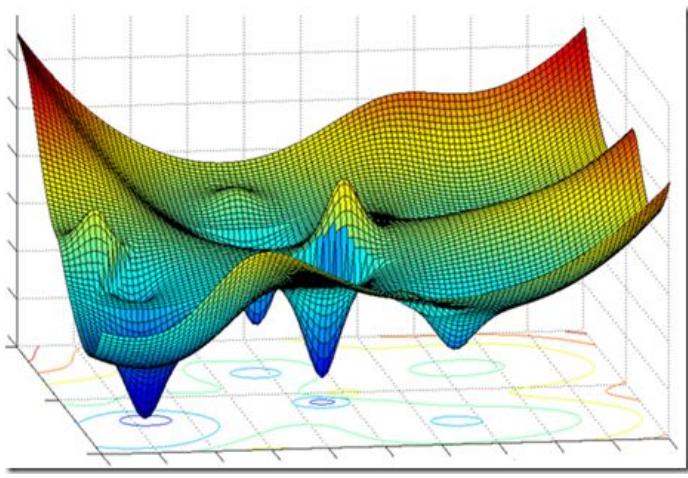
How do I decide on next set of weights?



The world of Gradient Descent







Usually it's not pretty

What is Gradient?

$$\frac{d}{dx} f(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

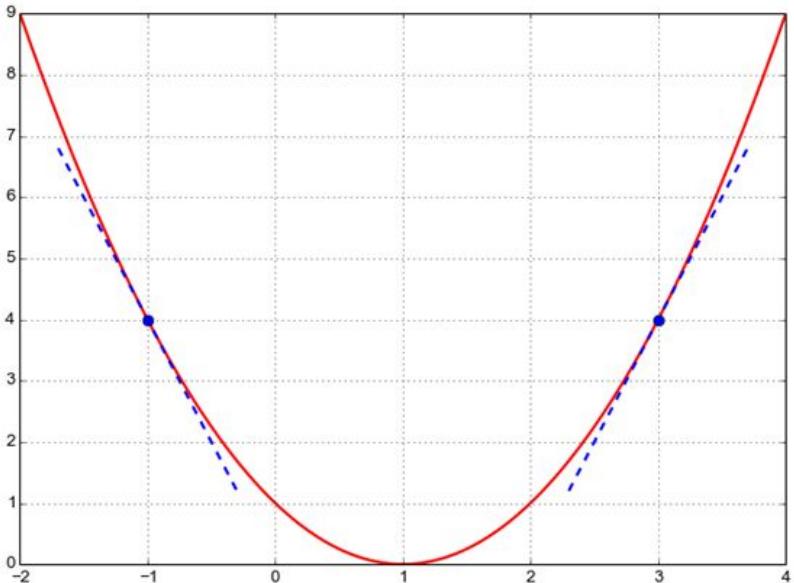
Derivative of a function

$$Loss = \frac{\sum_{i=1}^m (Prediction_i - Actual_i)^2}{2 \times m}$$

What function we want to minimize?

Visualizing a simple function

$$f(x) = (x - 1)^2$$



Derivative when $X = 3$

$$\begin{aligned}\frac{d}{dx}f(3) &= \lim_{h \rightarrow 0} \frac{f(3+h) - f(3)}{h} \\ &= \lim_{h \rightarrow 0} \frac{(3+h-1)^2 - (3-1)^2}{h} \\ &= \lim_{h \rightarrow 0} \frac{h^2 + 4h}{h} \\ &= \lim_{h \rightarrow 0} h + 4 = 4\end{aligned}$$

Derivative when X = 3

$$\begin{aligned}\frac{d}{dx}f(3) &= \lim_{h \rightarrow 0} \frac{f(3+h) - f(3)}{h} \\&= \lim_{h \rightarrow 0} \frac{(3+h-1)^2 - (3-1)^2}{h} \\&= \lim_{h \rightarrow 0} \frac{h^2 + 4h}{h} \\&= \lim_{h \rightarrow 0} h + 4 = 4\end{aligned}$$

A **positive** derivative indicates that function $f(x)$ is increasing.

Derivative when X = -1

$$\begin{aligned}\frac{d}{dx}f(-1) &= \lim_{h \rightarrow 0} \frac{f(-1+h) - f(-1)}{h} \\ &= \lim_{h \rightarrow 0} \frac{(-1+h-1)^2 - (-1-1)^2}{h} \\ &= \lim_{h \rightarrow 0} \frac{h^2 - 4h}{h} \\ &= \lim_{h \rightarrow 0} h - 4 = -4\end{aligned}$$

A **negative** derivative indicates that function $f(x)$ is decreasing.

New value of X

$$x_{new} = x_{old} - \eta \frac{d}{dx} f(x_{old})$$

New value of weight

$$w_{new} = w_{old} - \eta \frac{d}{dw} J(w_{old})$$

Loss function

Individual weights

$$w_{1(new)} = w_{1(old)} - \eta \frac{d}{dw_1} J(w_{1(old)})$$

$$w_{2(new)} = w_{2(old)} - \eta \frac{d}{dw_2} J(w_{2(old)})$$

$$w_{3(new)} = w_{3(old)} - \eta \frac{d}{dw_3} J(w_{3(old)})$$

What about bias?

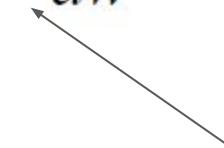
$$b_{new} = b_{old} - \eta \frac{d}{db} J(b_{old})$$



Life is Gradient Descent

Learning Rate

$$w_{new} = w_{old} - \eta \frac{d}{dw} J(w_{old})$$

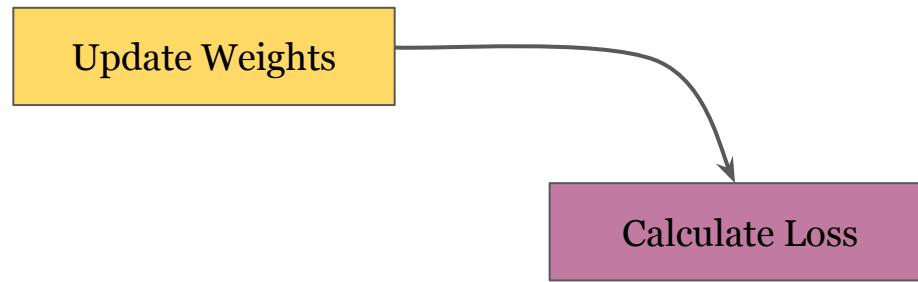


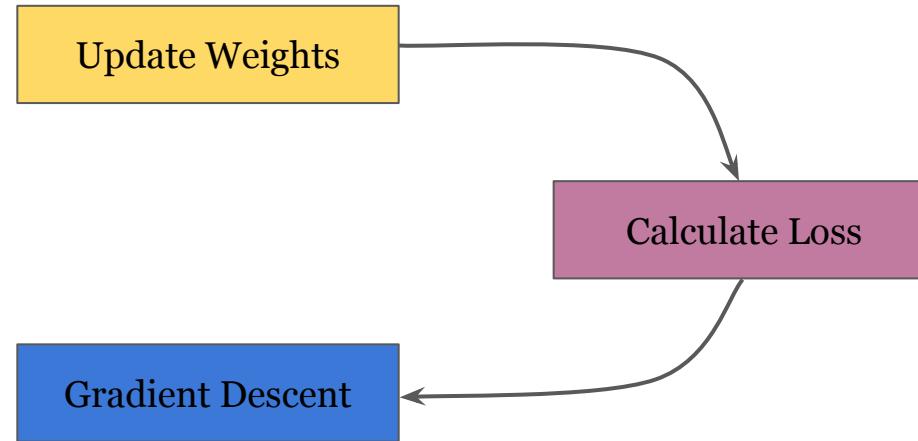
What should be value of
learning rate?

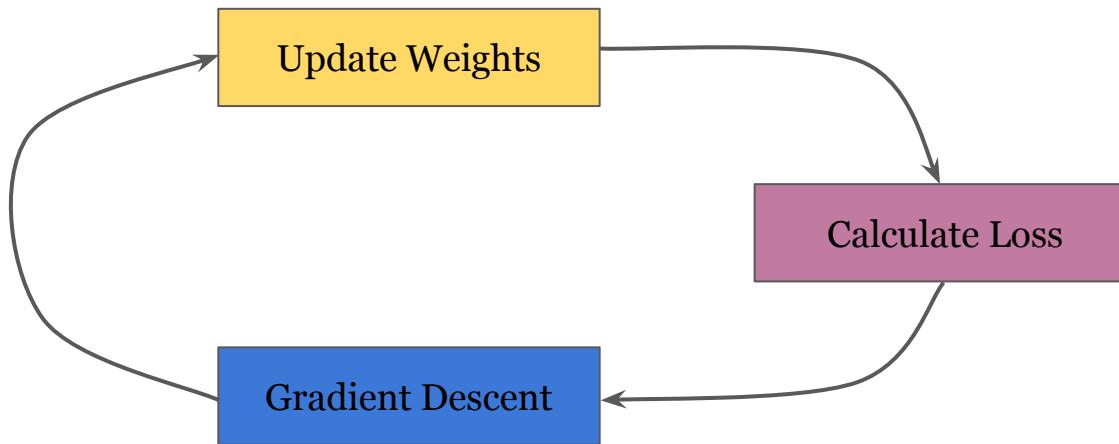


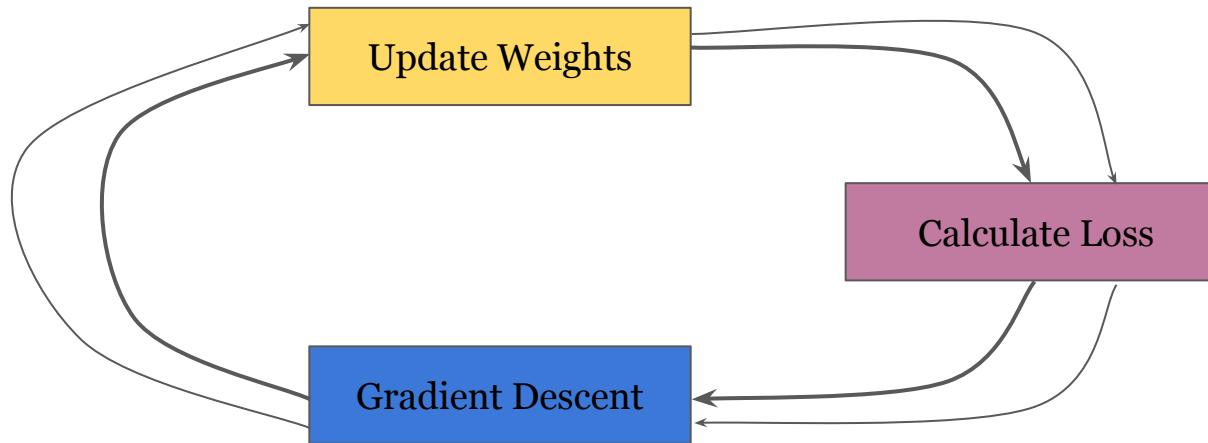
What's Next?

Update Weights

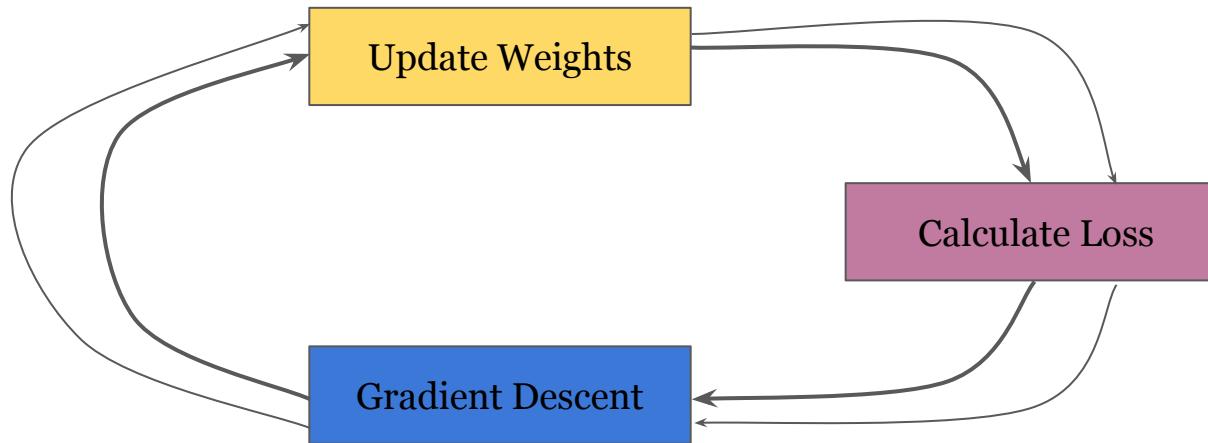








Do this multiple times...



Do this multiple times...

After lot of iterations...

$$y = 8369.175 * x_1 + 1.84732 * x_2 - 23.65543 * x_3 + 22.121977$$