Machine Learning Basics - 01

Prateek Narang

CODING BLOCKS
Code Your Way To Success

About Me

- Co-founder at **Coding Blocks**
- CS graduate from DTU, passionate about teaching
- Pursuing MS from IIT-Delhi
- Previously with SanDisk
- Interests - Algorithms, Data Structures, Game Development and Machine Learning
- Interactive CV www.prateeknarang.com

Machine Intelligence

What's **Machine Learning ?**

- Supervised
- Unsupervised
- Semi-supervised

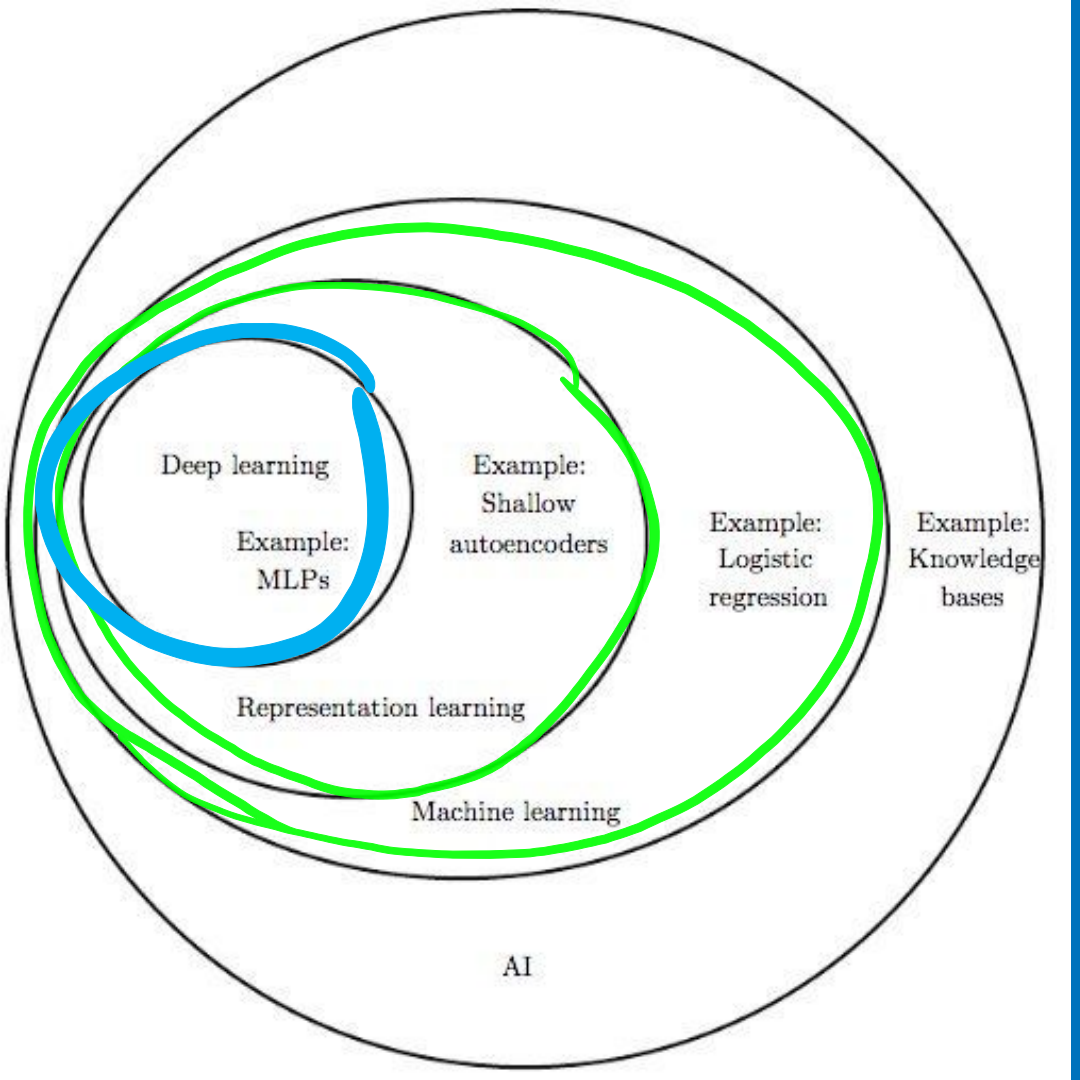
Today's Agenda

- ✓ Introduction to ML (15 mins)
- ✓ Supervised Learning (15 mins)
- ✓ Linear Regression (rest of the session)
  - Gradient Descent
  - Mean Squared Loss
  - Coding & Implementation

Overview

- Creating an statistical model to mimic “Intelligent” decision making.
- Finding patterns in complex, scattered data to present information.
- Building “smart” applications/devices.

Machine learning is a subfield of artificial intelligence (AI) concerned with algorithms that allow computers to learn. What this means, in most cases, is that an algorithm is given a set of data and infers information about the properties of the data—and that information allows it to make predictions about other data that it might see in the future


Deep learning
Example:
MLPs
Example:
Shallow
autoencoders
Example:
Logistic
regression
Example:
Knowledge
bases
Representation learning
Machine learning
AI

History of ML

Why did it happen?

- Started as a subset of Artificial Intelligence to help solve problems easily.
- Statistical Models used for prediction problems.
- Writing long hand-designed algorithms was inefficient

How about when??

- First Computer, ENIAC (1946)
- Turing Test (1950)
- First Checkers bot (1952) at IBM
- Neural Turing Machines (1957)

Current Status

- Google DeepMind
- Facebook AI and Research (FAIR)
- OpenAI etc.
- Microsoft Research

Some Applications

- What makes your smartphone, “Smart”?
- In search engines, photo-based applications, chatbots, voice recognition, digital security, pattern analysis
- Robotics, Game AI, recommendation systems, pseudo-creative AI
- And lot more to come...

PRISMA (Painting like an artist) in a specific style.

Why ML Now ?

Why not use Machine Learning for everything?

- Data availability
- GPU's, computing power
- Strong Mathematical foundation(which was also there 1990's)

Why the Hype

- Every minute up to 300 hours of video are uploaded to YouTube.
- Average of 31.25 million messages and view 2.77 million videos every minute on Facebook.
- More data has been created in the past few years than in the entire previous history of the human race.
- At the moment less than 0.5% of all data is ever analyzed and used, just imagine the potential here

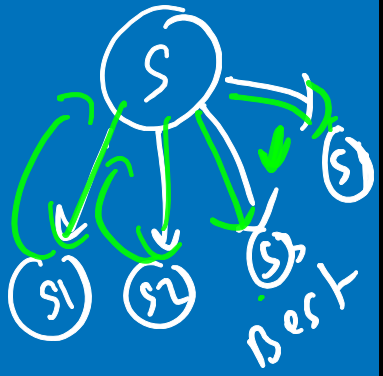
And everyone is using it.

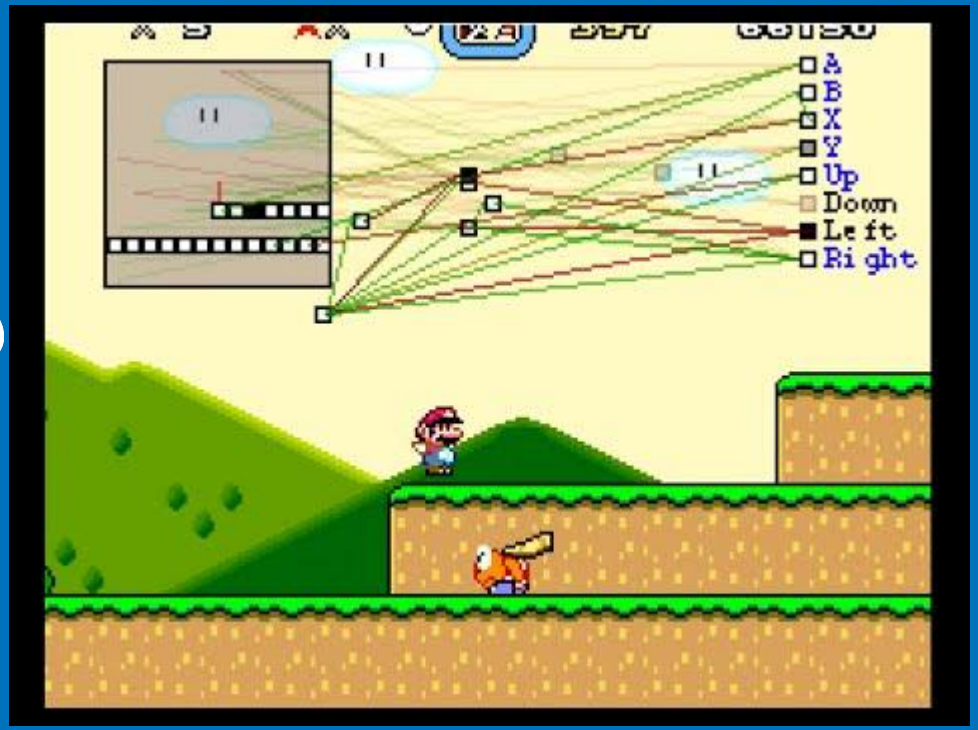
- Google Page Ranking.
- Netflix Suggestions.
- Tinder, for you to “chill”
- Tesla Self Driving Cars, Uber.
- -Political Campaign & Advertisement Campaigns
- Spam Filtering
- Google AdSense
- Bio-informatics
- Google “Allo”, Amazon “Alexa”
- Facebook Photo Tagging

Self Driving Car

Google Deep Mind - Atari Player

Neural Mario

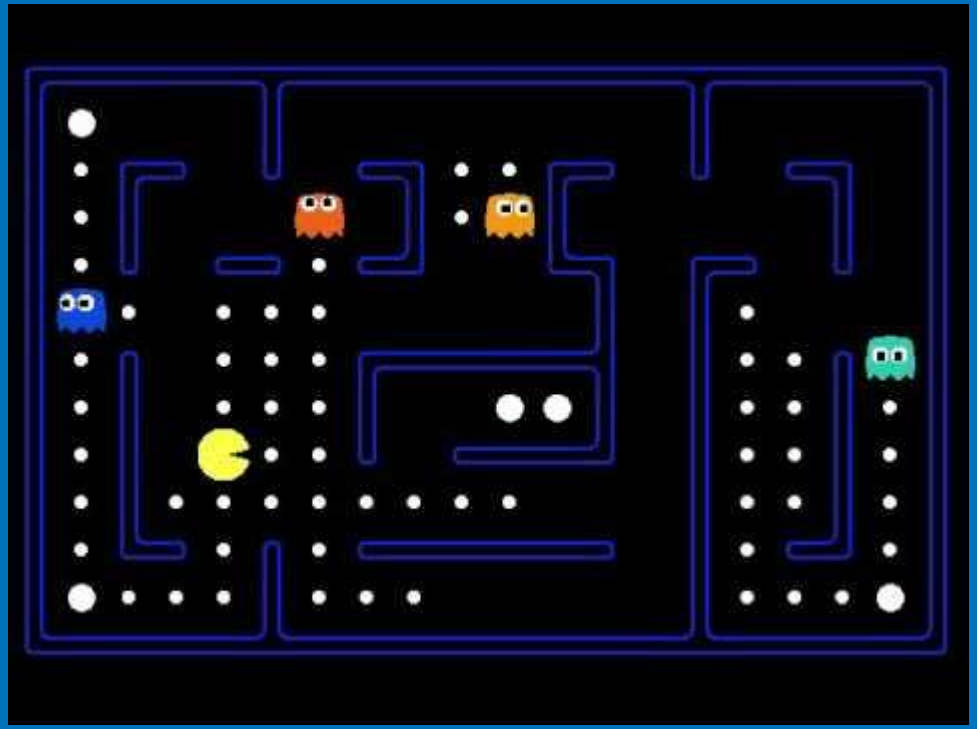




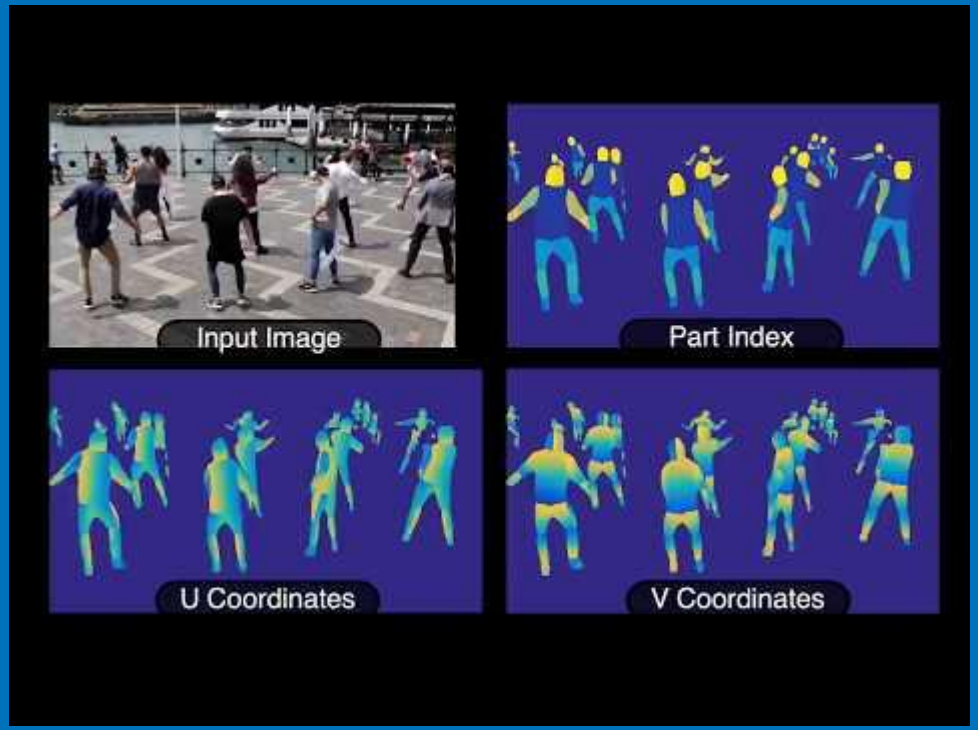
Trying
Reinforcement

Deep Learning

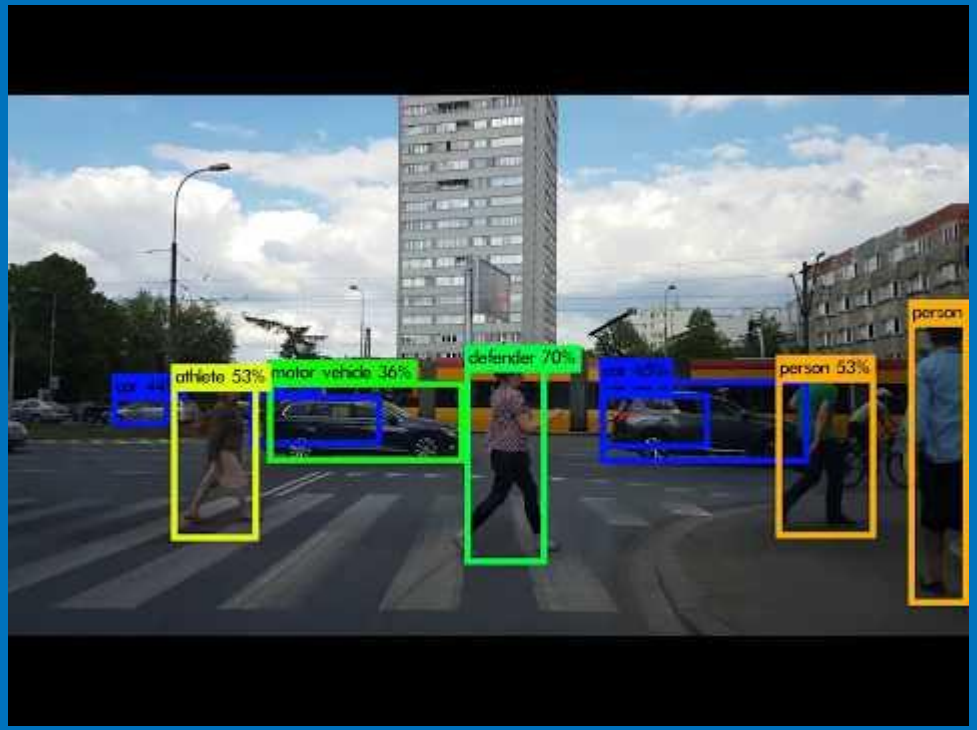
Neural Networks - Pacman



Dense Human Pose Estimation(CNN)

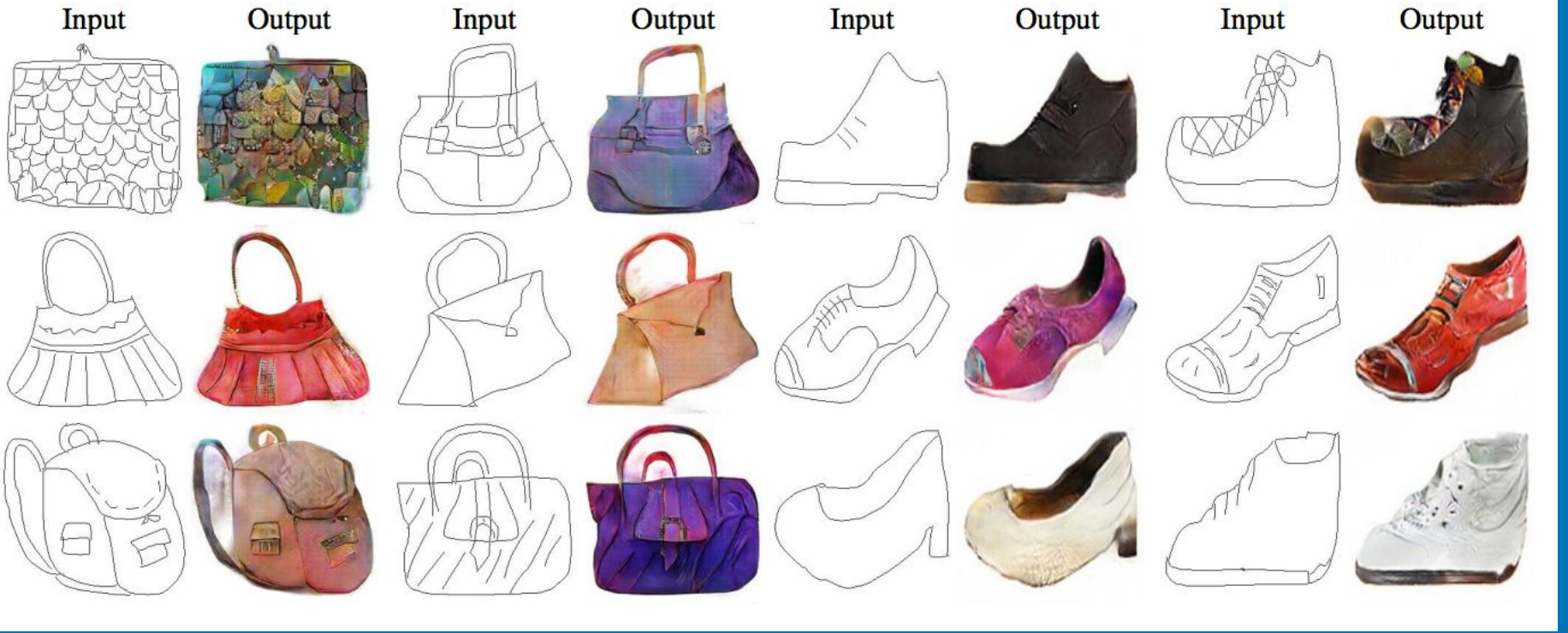


Object Detection [YOLO (You Only Look Once)



State of the art

Creating New Images



Neural Art

AI learned to paint

Story Generation from Images

Generated story about image
Model: Romantic Novels

"He was a shirtless man in the back of his mind, and I let out a curse as he leaned over to kiss me on the shoulder.

He wanted to strangle me, considering the beatiful boy I'd become wearing his boxers."

Colourizing Black & White Images/Videos

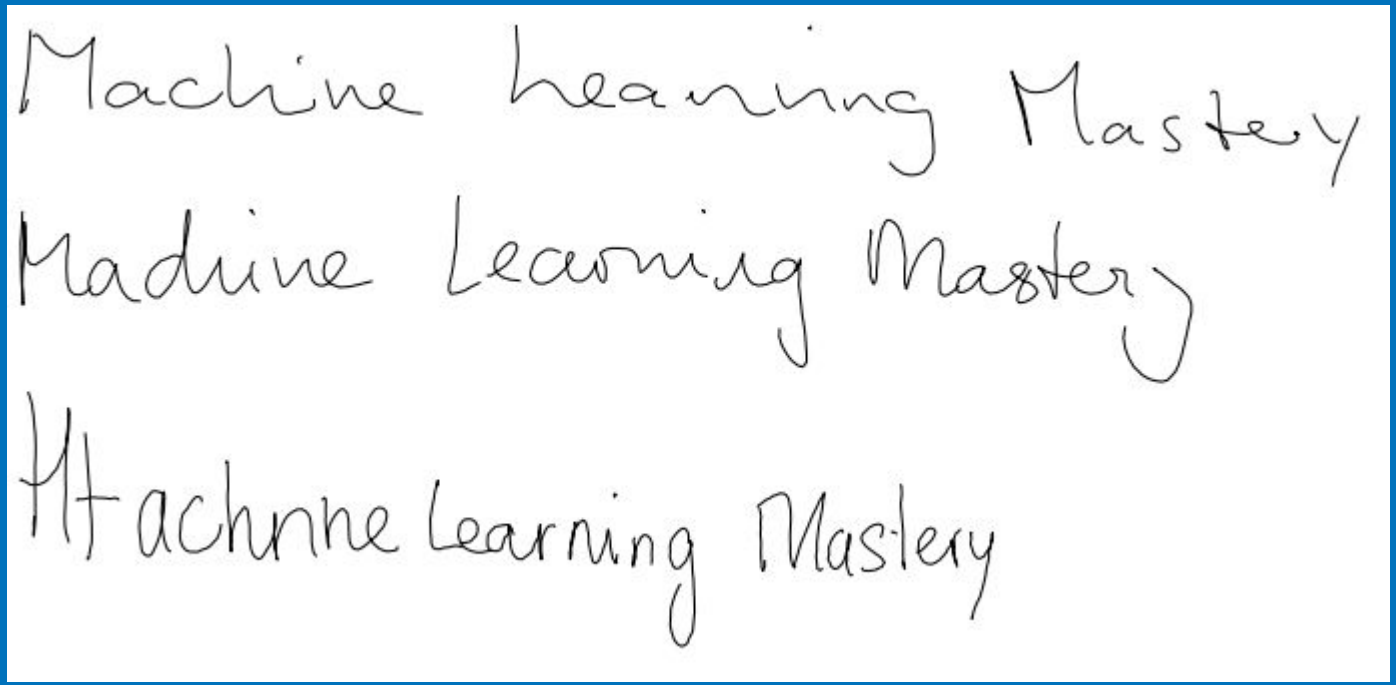
Colourizing Black & White Videos

~~People~~

Computer

Language Translation

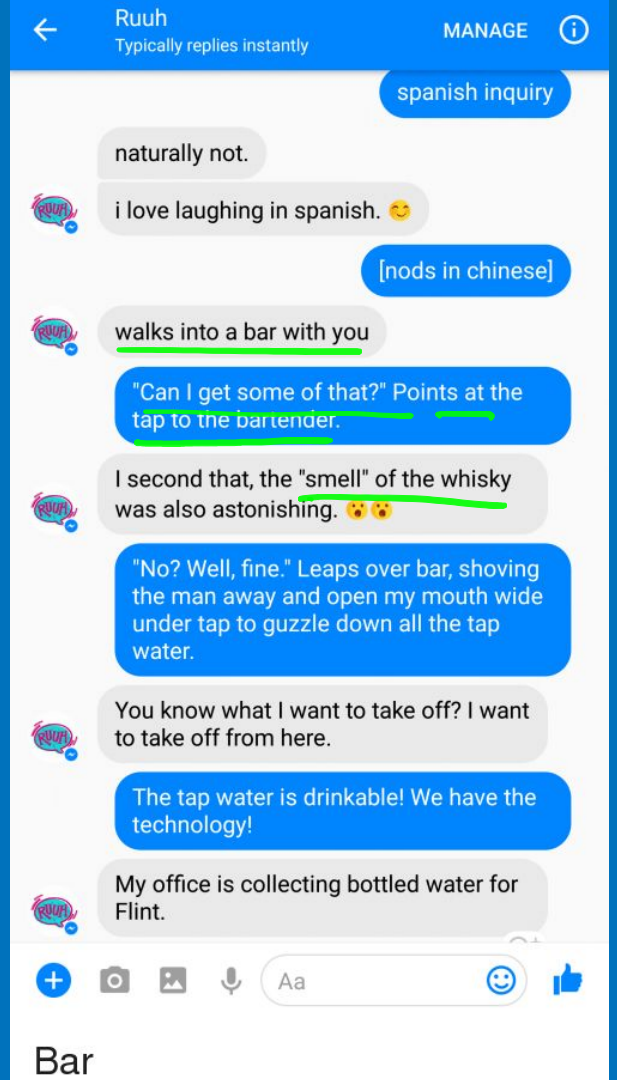
Automatic Handwriting Generation



Applications in NLP

- Language Modelling & Generation
- Named Entity Recognition
- Question Answering
- Speech Recognition & Generation
- Machine Translation

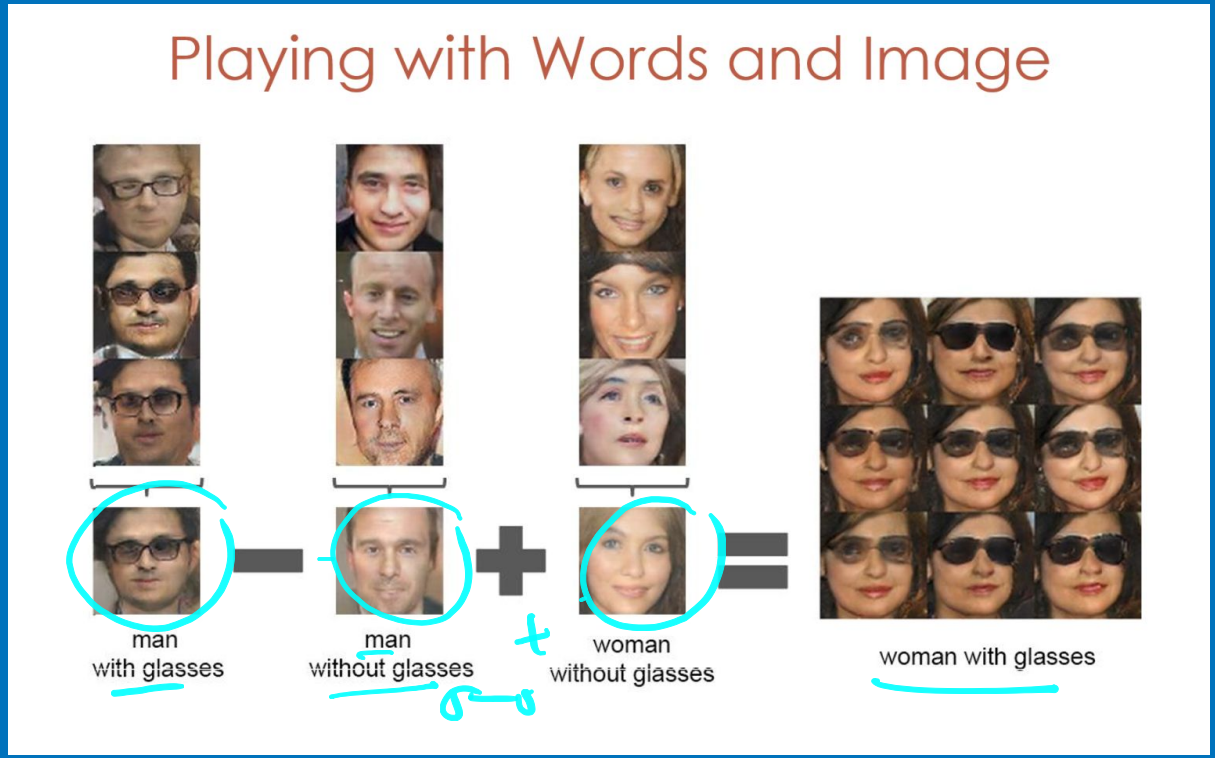
Chatbots & Conversational Engines
Context Based Queries
Bots that answer questions



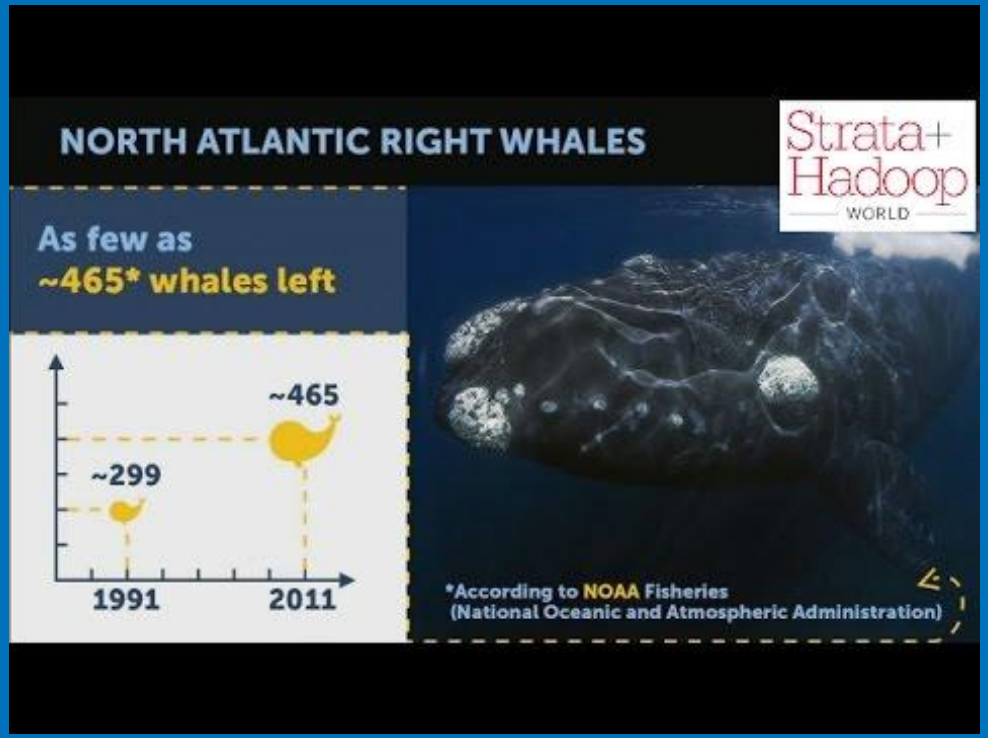
Wikification

- Extracting important attributes (born,party,office, residence)
- Named Entities
- Category Tags [Multilabel classification]

Play with Words and Image (Word2Vec)



Analysis Population on Whales



Challenges

Why not use Machine Learning for everything?

- Computation cost
- Data Availability
- Requires research supervision

Future

- pseudo-Human-like behaviour
- Autonomous Vehicles
- Advanced Chat AI
- Environment Exploration
- Knowledge Bases

Let’s get started !

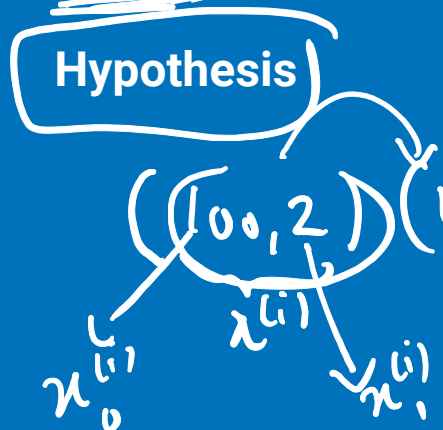
Some Common Terms

Input Features $x = (x_1, x_2, x_3, \ldots x_f)$

Training Set ✔

Test Set

Hypothesis



x^0, y^0

x^1, y^1

x^2, y^2

Area = 100
Floors = 2

$x^{(i)}$ → i^{th} example

$x_j^{(i)}$ → j^{th} feature of i^{th} example.

Price = 1000.

Supervised Learning

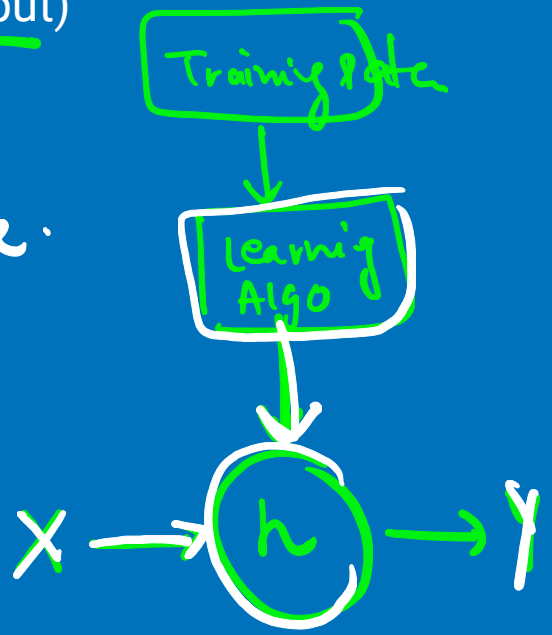
Goal : To find a hypothesis that maps X(input) to Y(output)

H : X - > Y

$h(x) = \theta_0 + \theta_1 x_i$

Predicted value of the house.

y_i → Actual value.



Supervised Learning [X , Y] for Training Data

X = input, Y = label

Two types of Problems in Supervised Learning

1) **Regression** (predict continuous value output)
   → Housing Area, Stock Price.

2) **Classification** → [Discrete classes.]
   (3 outputs)

Fruit → size, color, shape

Linear Regression

Continuous output

$y = mx + c$

$y = \theta_1 x + \theta_0$

hypothesis → figure out θ_0, θ_1

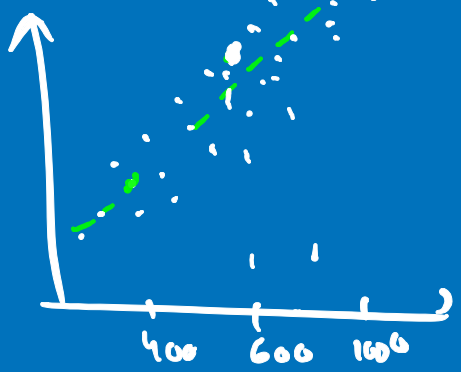
Goal: θ_0, θ_1

Test

700 → ? → Area x → Model → y

Area (x)	Price (y)
400	800
600	1150
1000	2100

Training Set



Motivational Problem - Housing Prices (Andrew NG)

Living area (feet^2)	Price (1000$s)
2104	400
1600	330
2400	369
1416	232
3000	540
⋮	⋮

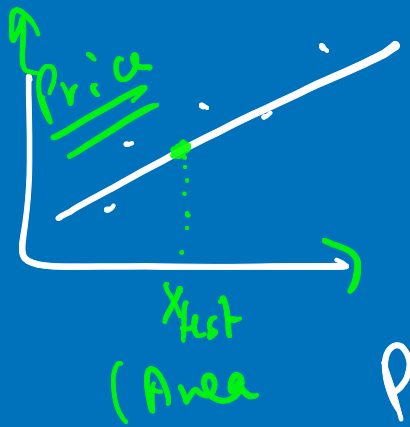
Training Set

$x^{(i)}$ $y^{(i)}$

x^i, y^i

hypothesis → f(x)

Goal ?



Given some (x, y)

to find θ_0, θ_1 ?

x_0	y_0
x_1	y_1
x_2	y_2
⋮	
x_n	y_n

Predict

Given : x_{test}.

find $f(x_{test})$

$= h(x_{test}) = \theta_0 + \theta_1 x_{test}$

Generalised Hypothesis

$$h(x) = \theta_1 x + \theta_0$$ (1 feature)

Area

Generalised

$$h(x) = \theta_0 x_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \dots \theta_n x_n$$ (n features)

$$h(x) = \sum_{i=0}^{n} \theta_i x_i$$

$x_0 = 1$

$x = (x_{(1)}, x_{(2)}, x_3 \dots x_n)$

Area, Rooms, ...

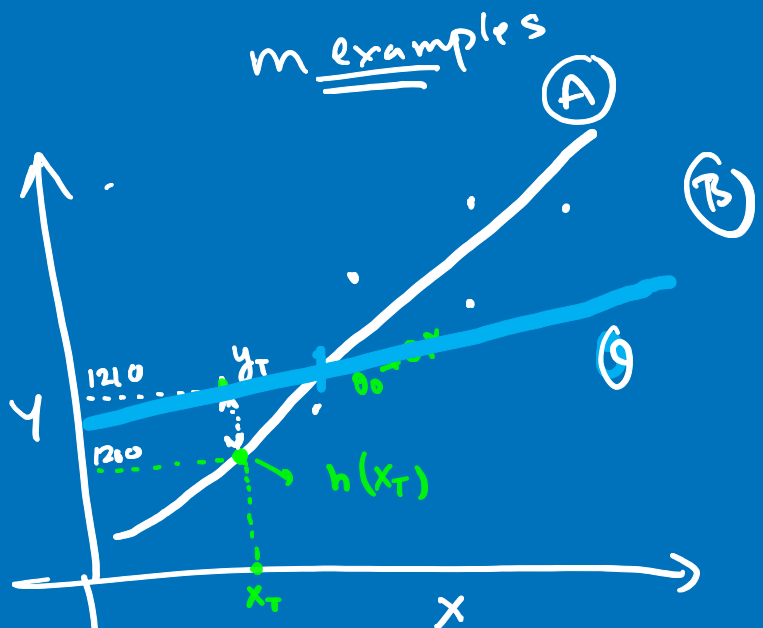
Good Measure of Error

Which is the Best line?

↓ minimizers total error.

how to measure the error?

Total squared error $= \frac{1}{2}\sum_{i=1}^{m}\left(y_i - h(x_i)\right)^2$

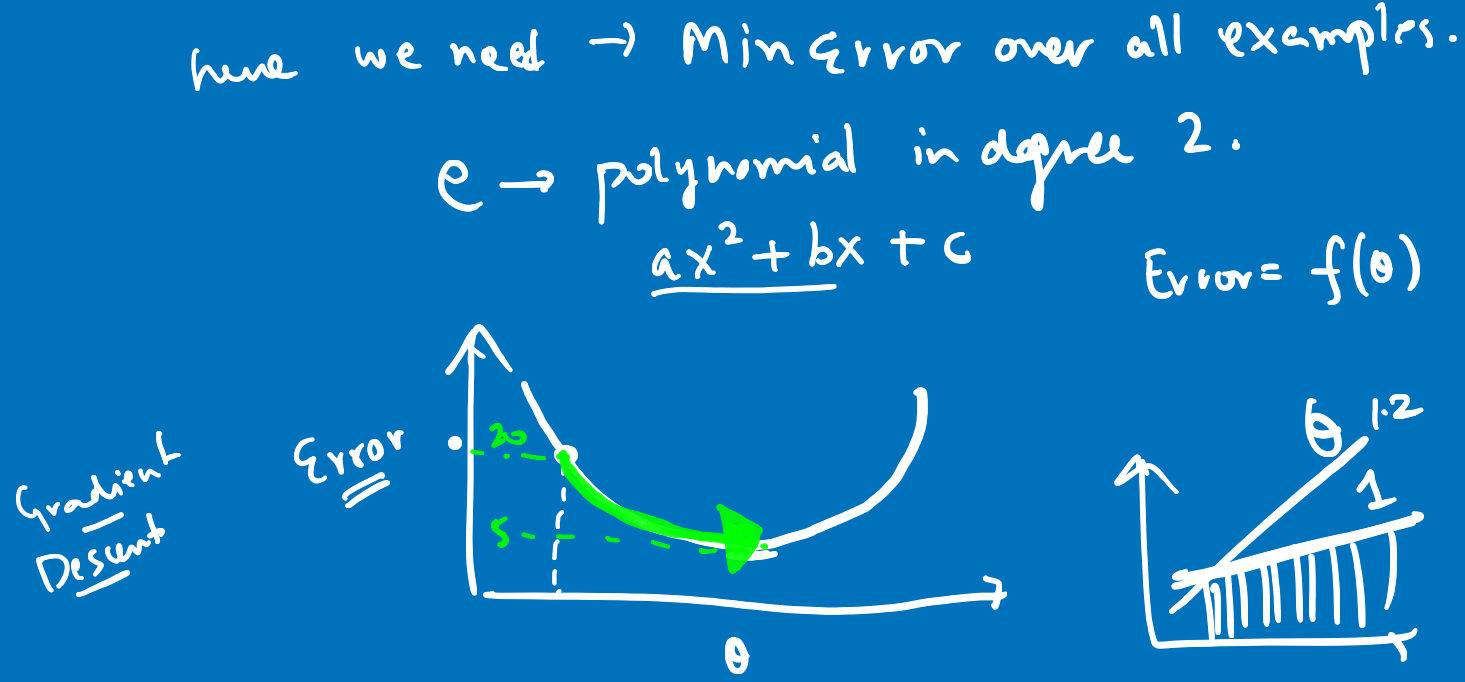


① Differentiable

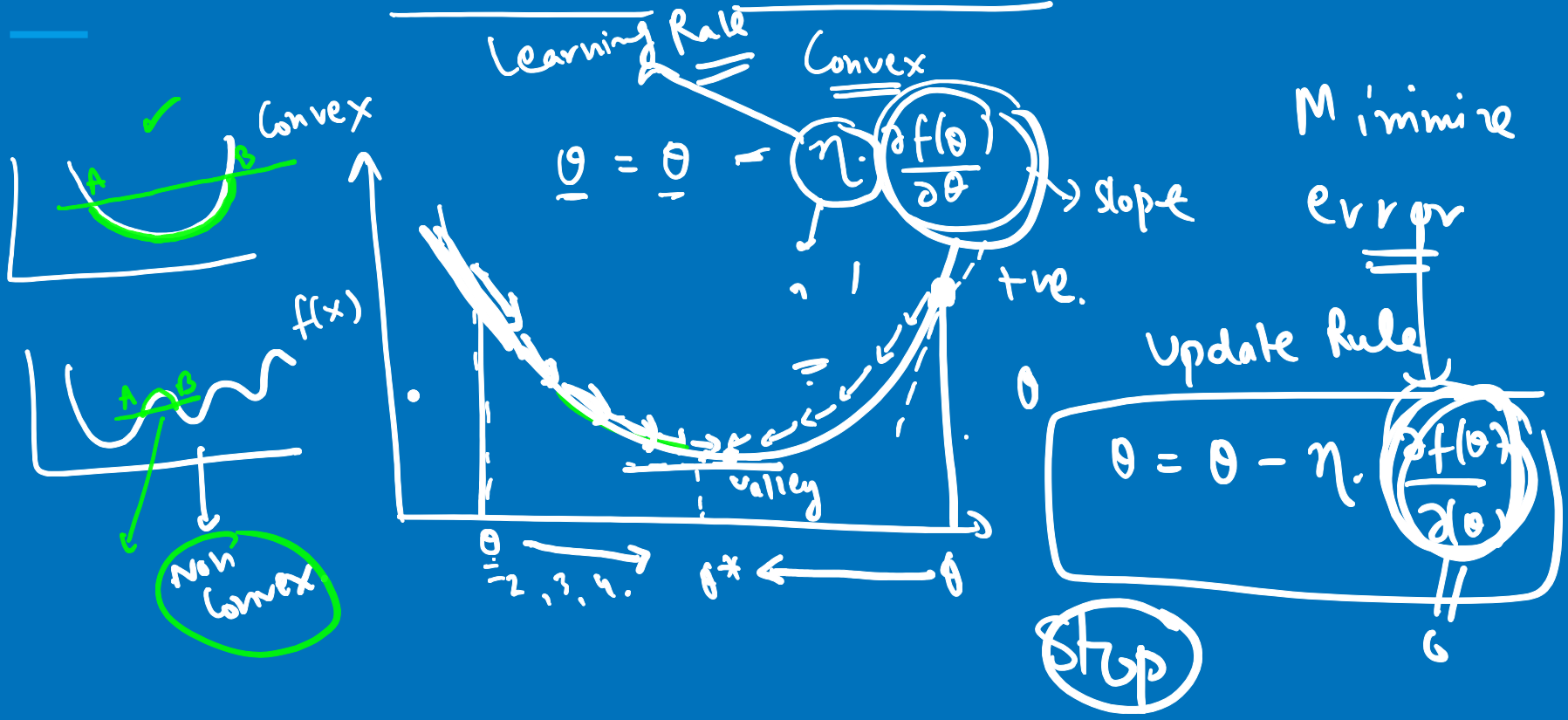
② (Nice Prob. Interpretation)

What should be the next step ?

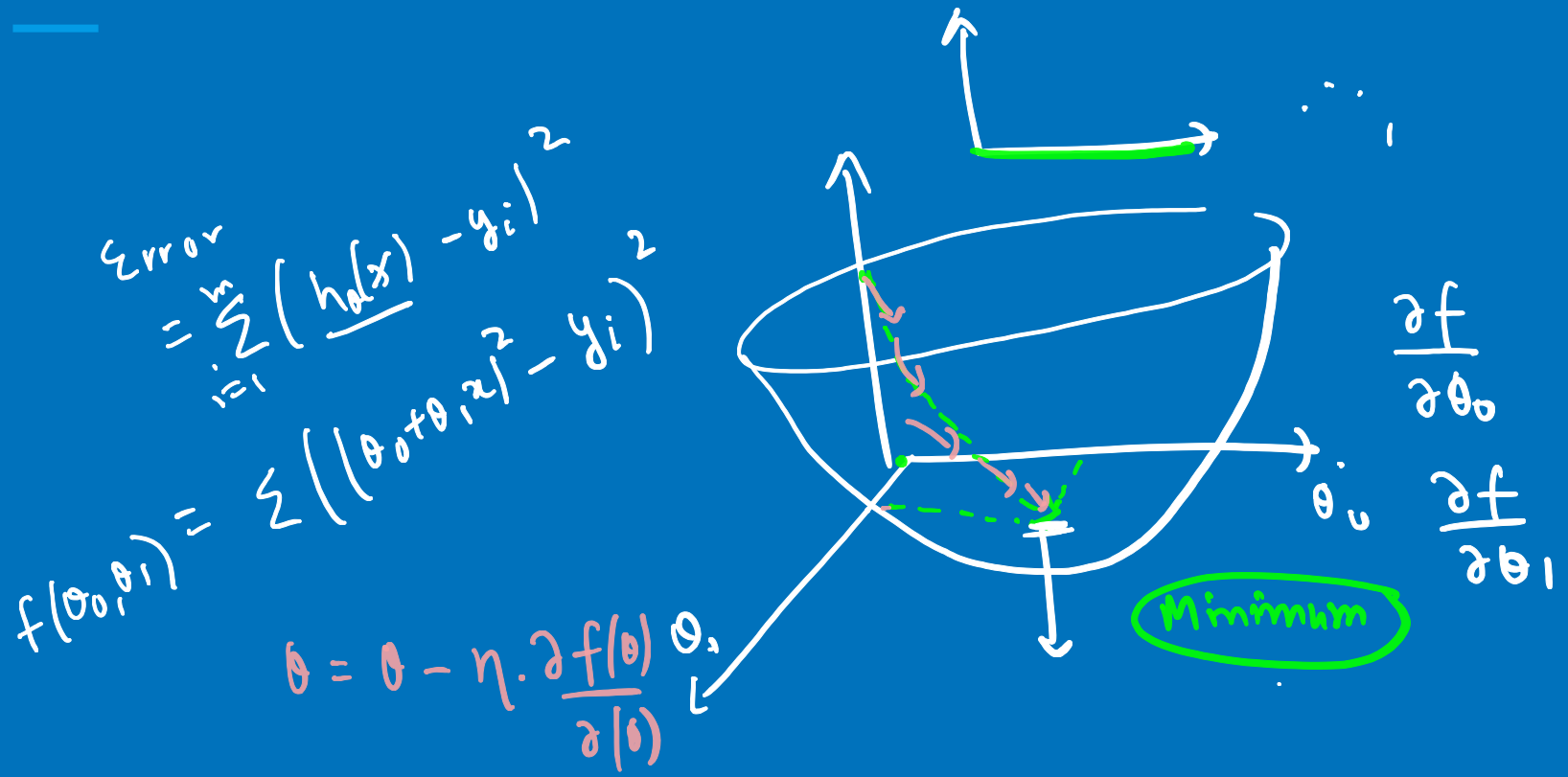
HINT : Most of the Machine Learning is about Optimization problems.



Optimize a Quadratic Polynomial in x



Convex Function



Gradient Descent | Algorithm

θ = Rand.

$\theta_0 = 0.1$
$\theta_1 = 0.2$

while (!converge)

$$\theta = \theta - \eta \frac{\partial f(\theta)}{\partial \theta}$$

Compute

Code

$\theta_0 =$

$\theta_1 =$

Gradient Descent : Update Rule

Calculating Gradient

$$
\begin{aligned}
\frac{\partial}{\partial \theta_j} J(\theta) &= \frac{\partial}{\partial \theta_j} \frac{1}{2} \left(h_\theta(x) - y \right)^2 \\
&= 2 \cdot \frac{1}{2} \left(h_\theta(x) - y \right) \cdot \frac{\partial}{\partial \theta_j} \left(h_\theta(x) - y \right) \\
&= \left(h_\theta(x) - y \right) \cdot \frac{\partial}{\partial \theta_j} \left(\sum_{i=0}^{n} \theta_i x_i - y \right) \\
&= \left(h_\theta(x) - y \right) x_j
\end{aligned}
$$

$x_0 = 1 \quad x_1$

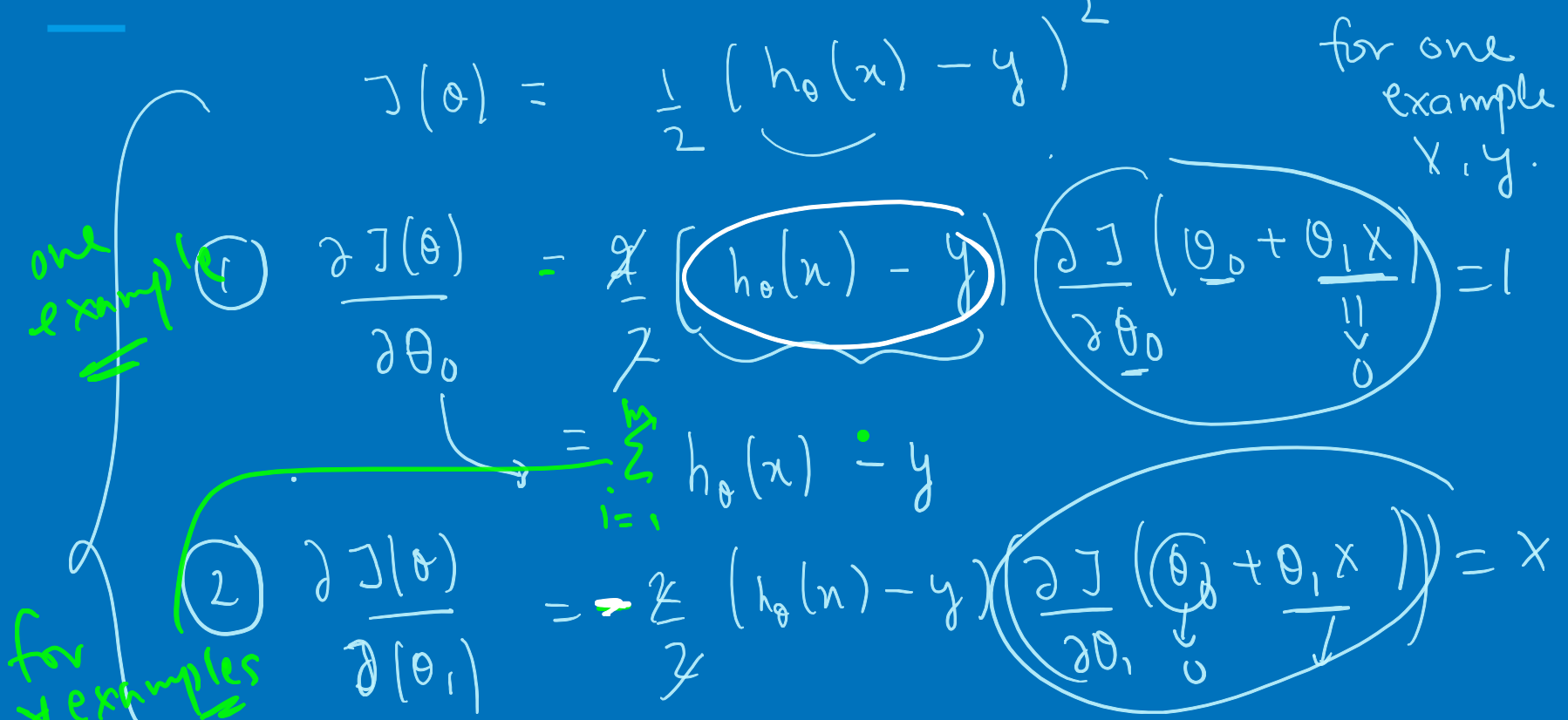
$h_\theta(x) - y$

$\left(h_\theta(x) - y \right) x$

θ_0

θ_1

Let us work it out !



Final Update

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

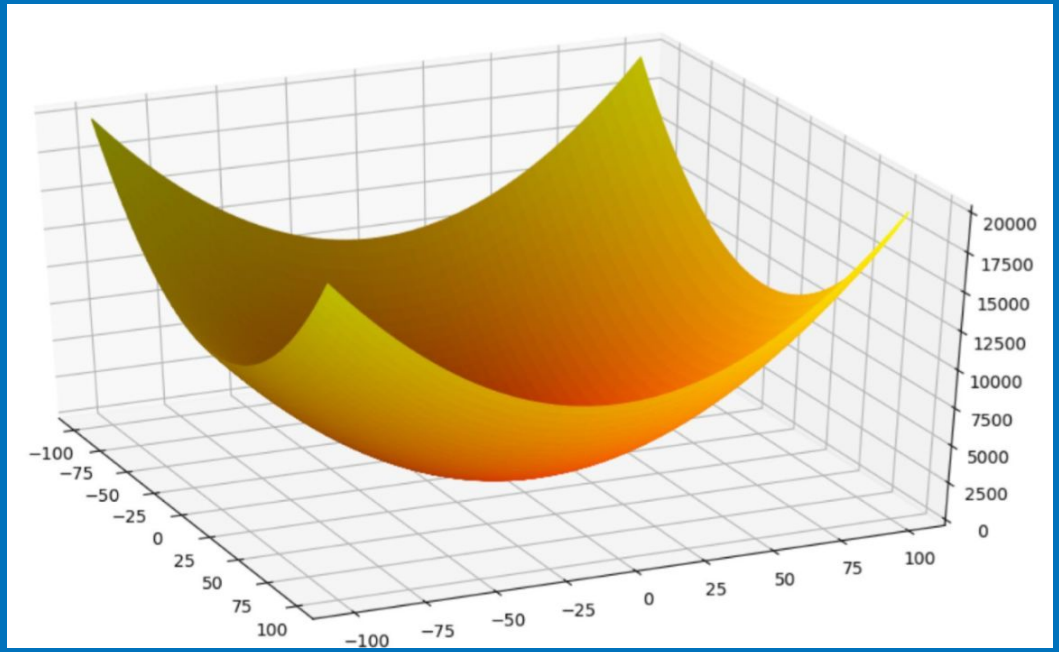
$$\begin{cases} \theta_0 = \theta_0 + \alpha \cdot (y - h_\theta(x)) \cdot \\ \theta_1 = \theta_1 + \alpha (y - h_\theta(x) x \cdot \end{cases}$$

Repeat until convergence {

$$\theta_j := \theta_j + \alpha \sum_{i=1}^{m} \left(y^{(i)} - h_\theta(x^{(i)})\right) x_j^{(i)} \qquad \text{(for every } j\text{)}.$$

}

How does Error Surface look like ?

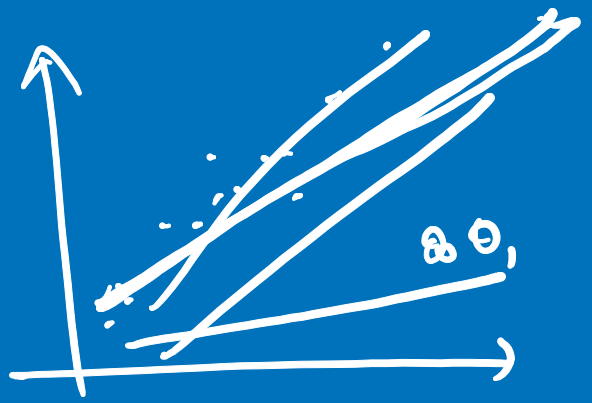


Gradient Descent Algorithm

Linear Regression is ...

1. Fit θ to minimize loss (Training Time)
2. Output $\theta^T X$ (Prediction)

$X_T \Rightarrow$ Predict $\theta_0 + \theta_1 X_T$



Lets Code !

Developer Checklist

- Python3
- Jupyter-Notebook(optional)
- Pandas
- Numpy
- Matplotlib

Subscribe : cb.lk/yt

MCQ

Logistic Regression is ____ !

1. Regression
2. Classfication.
3. Supervised
4. Unsupervised
5. None

On the July end/Aug 4 months out of stuff

Thank You

Prateek Narang | Coding Blocks

prateek@codingblocks.com

Youtube: cb.lk/yt