Intro to ML

- [x] What is AI?
- [x] Some History of AI
- [x] Why is AI Now?
- [x] What is ML?
- [x] ML vs. Traditional prog.
- [x] AI vs ML vs DL vs DS vs BD
- [x] Data Science Lifecycle
- [x] Types of ML
- [x] Regression vs. classification
- [ ] Linear Regression
  - [x] Simple
  - [ ] Multiple
  - [ ] polynomial
  - [x] Cost fn
  - [x] Gradient Descent
  - [x] Normal Eqn

☑ What is AI ?

AI → Artificial Intelligence vs Natural Intelligence

Artificial Intelligence
- create
- produce
- Make
  - Simulation of Natural Intelligence
    - Intelligent Machine

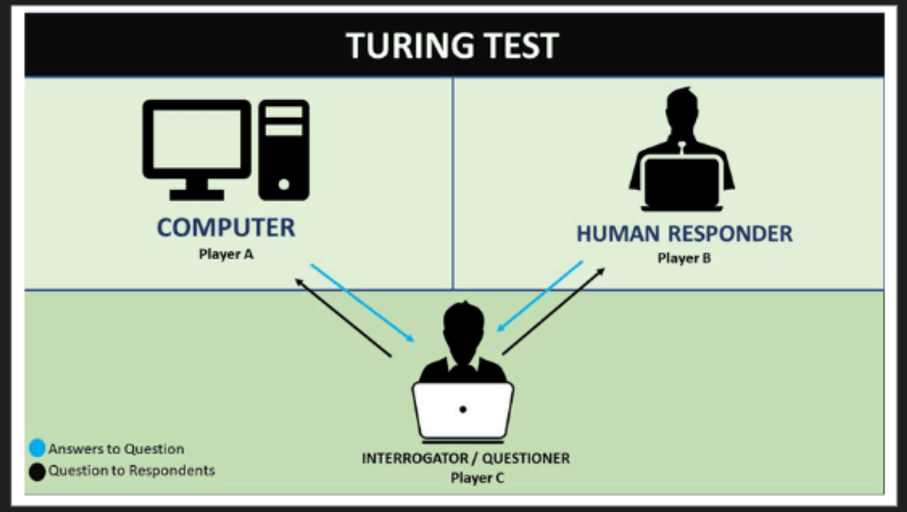
Intelligence
- Ability to:
  - Reason
  - Discover Meaning
  - learn by Experience
  - Generalize

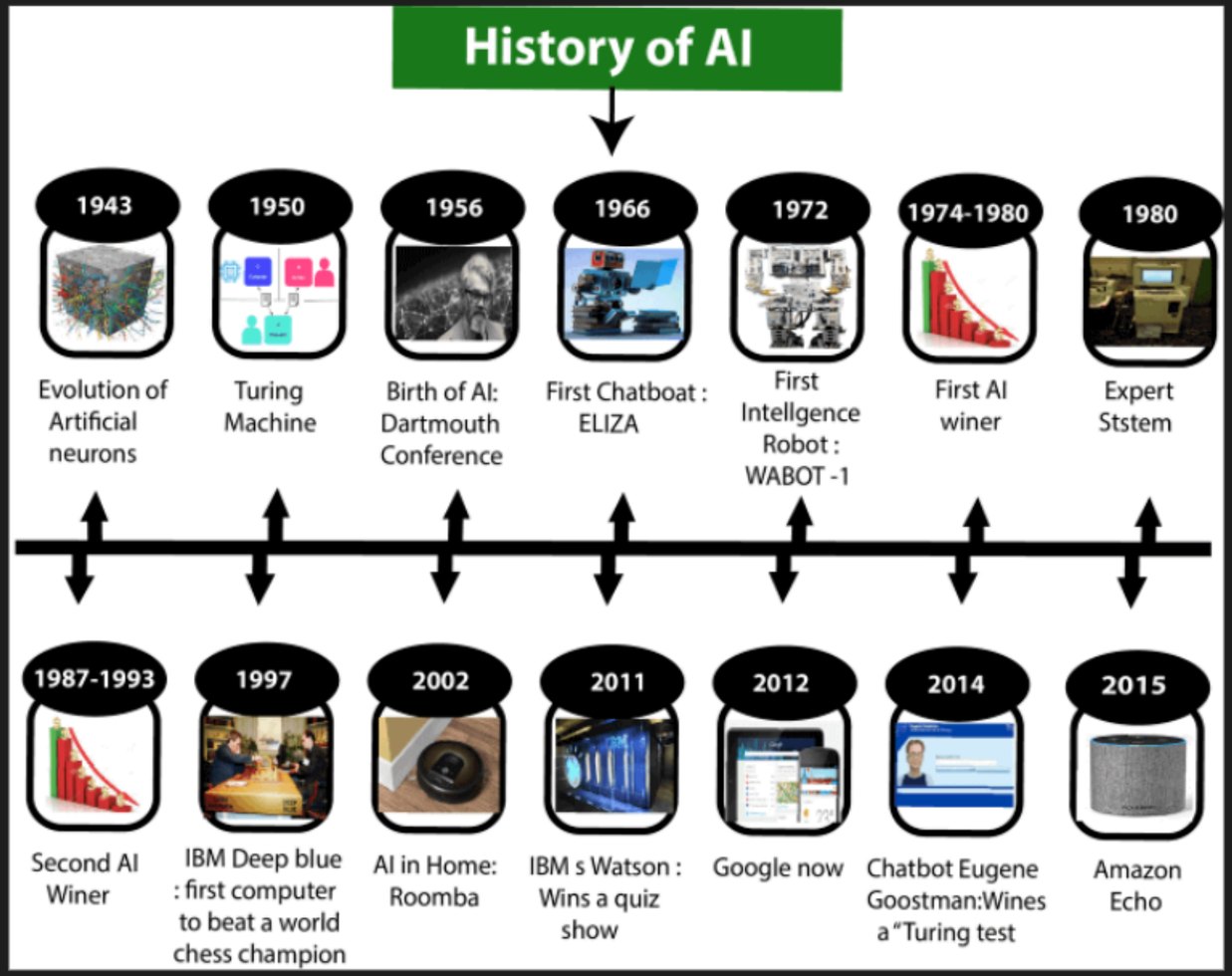
Some History of AI

- Alan Turing : Can Machines think ?

Turing Test

- AI winter I
- AI winter II





ML

learning from Data

- [x] Why is AI Now?
  - [ ] Processing Power 
  - [ ] Data Growth 
  - [ ] Algorithms Development 
  - [ ] Storage Cost 

Data Growth

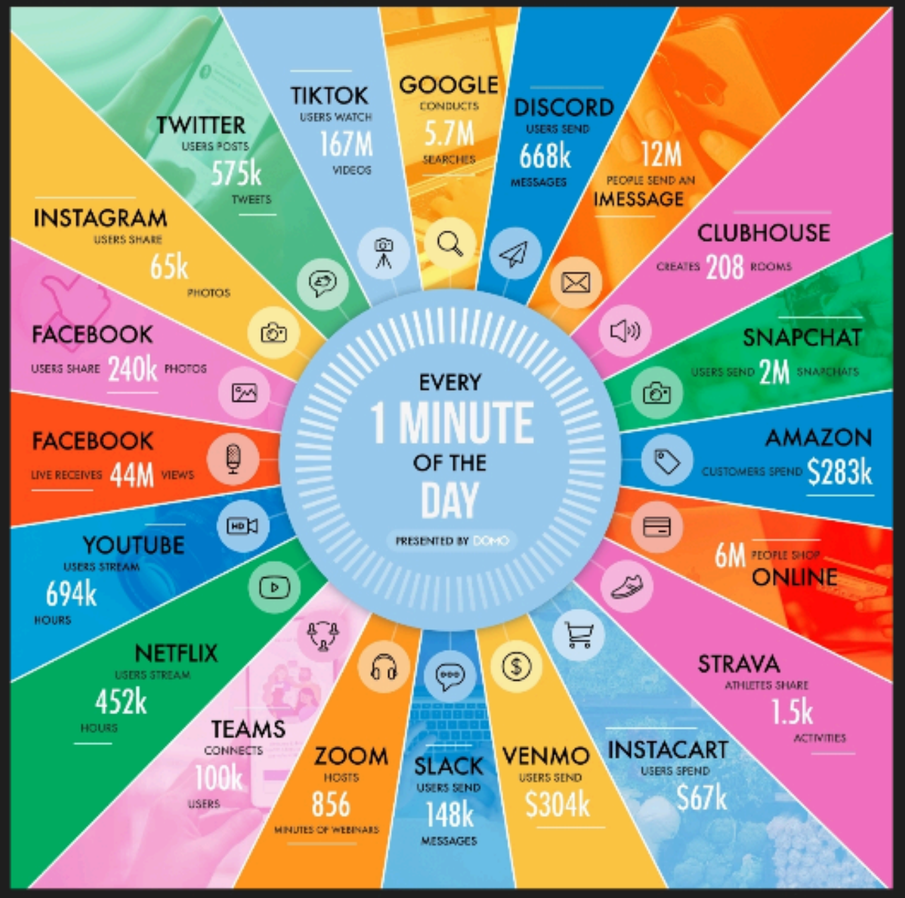
IoT

Internet of things

Data Never Sleeps 9.0

How much data is generated *every minute*?

The 2020 pandemic upended everything, from how we engage with each other to how we engage with brands and the digital world. At the same time, it transformed how we eat, how we work and how we entertain ourselves. Data never sleeps and it shows no signs of slowing down. In our 9th edition of the "Data Never Sleeps" infographic, we bring you a glimpse of how much data is created every digital minute in our increasingly data-driven world.



As of July 2021, the internet reaches 65% of the world's population and now represents 5.17 billion people—a 10% increase from January 2021. Of this total, 92.6 percent accessed the internet via mobile devices. According to Statista, the total amount of data consumed globally in 2021 was 79 zettabytes, an annual number projected to grow to over 180 zettabytes by 2025.

Global Internet Population Growth

(IN BILLIONS)

As the world changes, businesses need to change too—and that requires data. Domo gives you the power to make data-driven decisions at any moment, on any device, so that you can make smart choices in a rapidly changing world. Every click, swipe, share, or like tells you something about your customers and what they want, and Domo is here to help you and your business make sense of all of it.

Learn more at domo.com

SOURCES: LOCALIQ, BUSINESS OF APPS, JUSTIN STOUT, MOOSURE, EXPANDED RAMBLINGS, INTERNET WORLD STATS, STATISTA, CNBC, BRANDWATCH, KILL THE CABLE BILL, YOUTUBE, REVISTA, THE VERGE, MARKETINGTUTT COMMUNICATION, A CASE ANALYSIS APPROACH, INTERNET LIVE STATS, SOCIAL STATISTICS

☑ What is ML

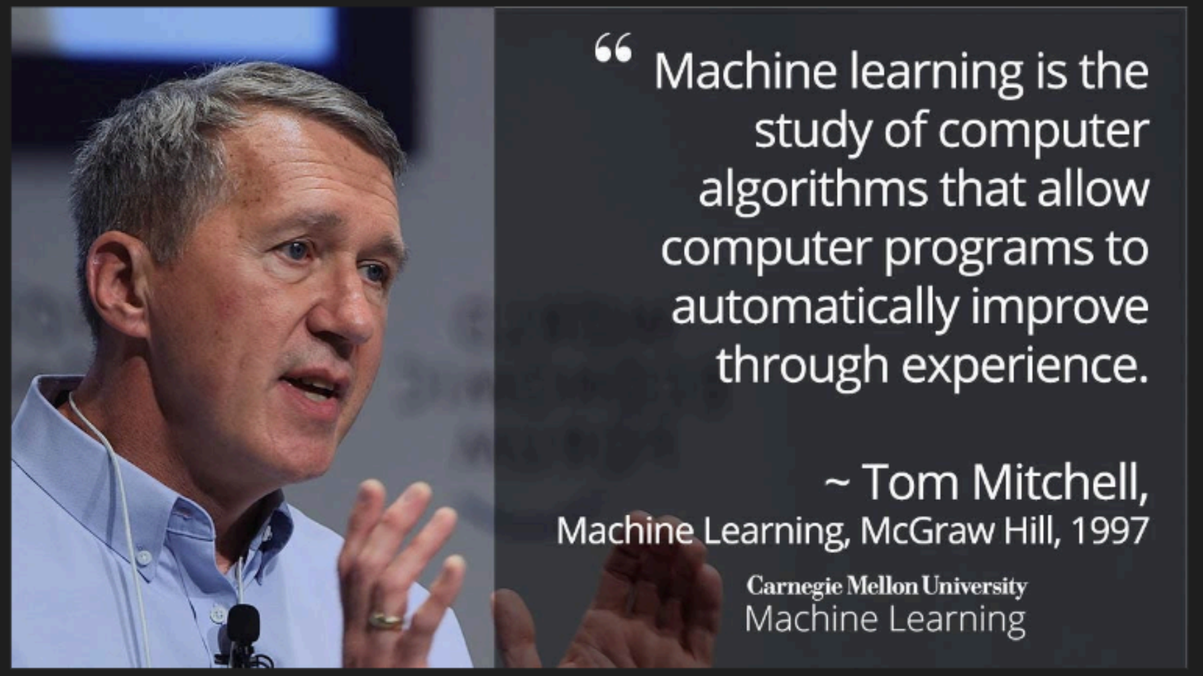
Machine Learning

→ Training the machine to gain Experience

Intelligent
Smart

Arthur Lee Samuel *(1959)*

Machine Learning *the "field of study that gives computers the ability to learn without being explicitly programmed".*



Tom Mitchell's Definition (1997)

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.

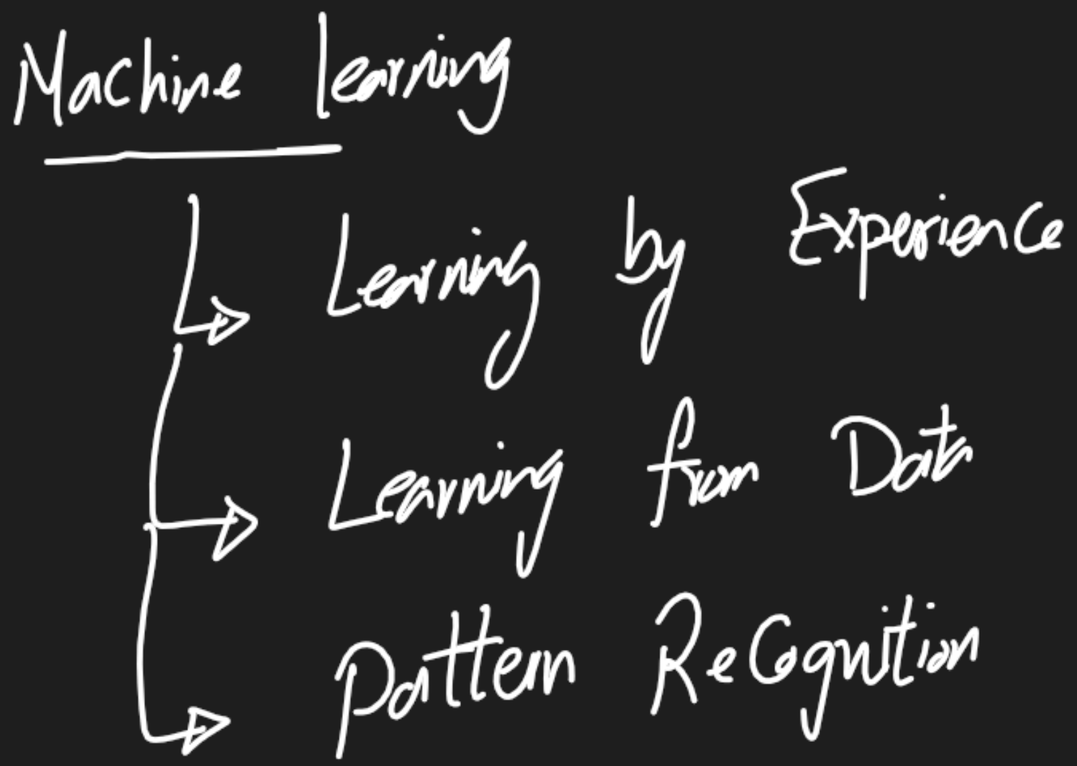
Learning

- Task (T): Estimation of House price
- Performance Measure (P): Error (Actual - Estimated price)
- Experience (E): No. of observations (& features)

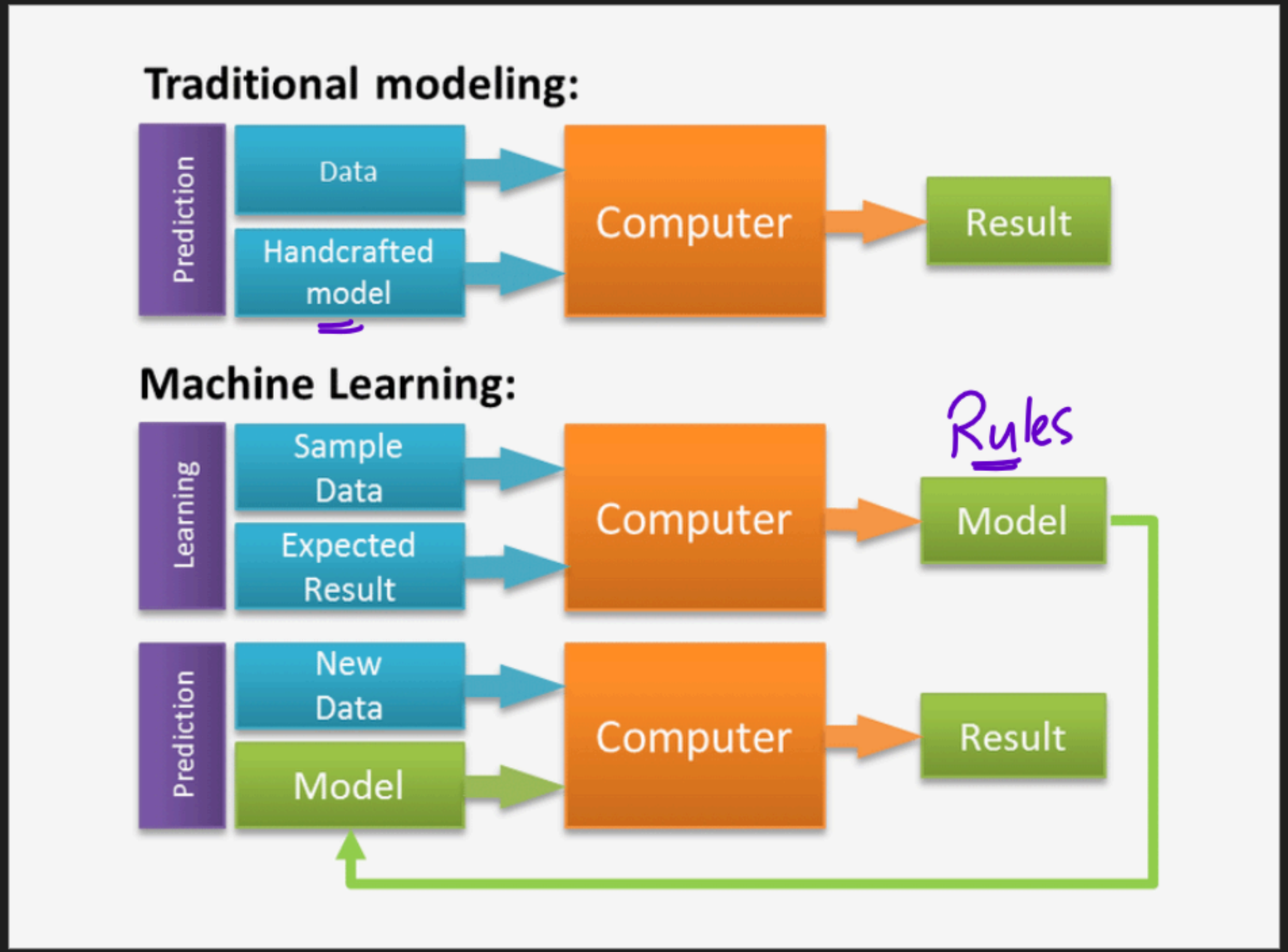
① 100 Houses
Error: 500

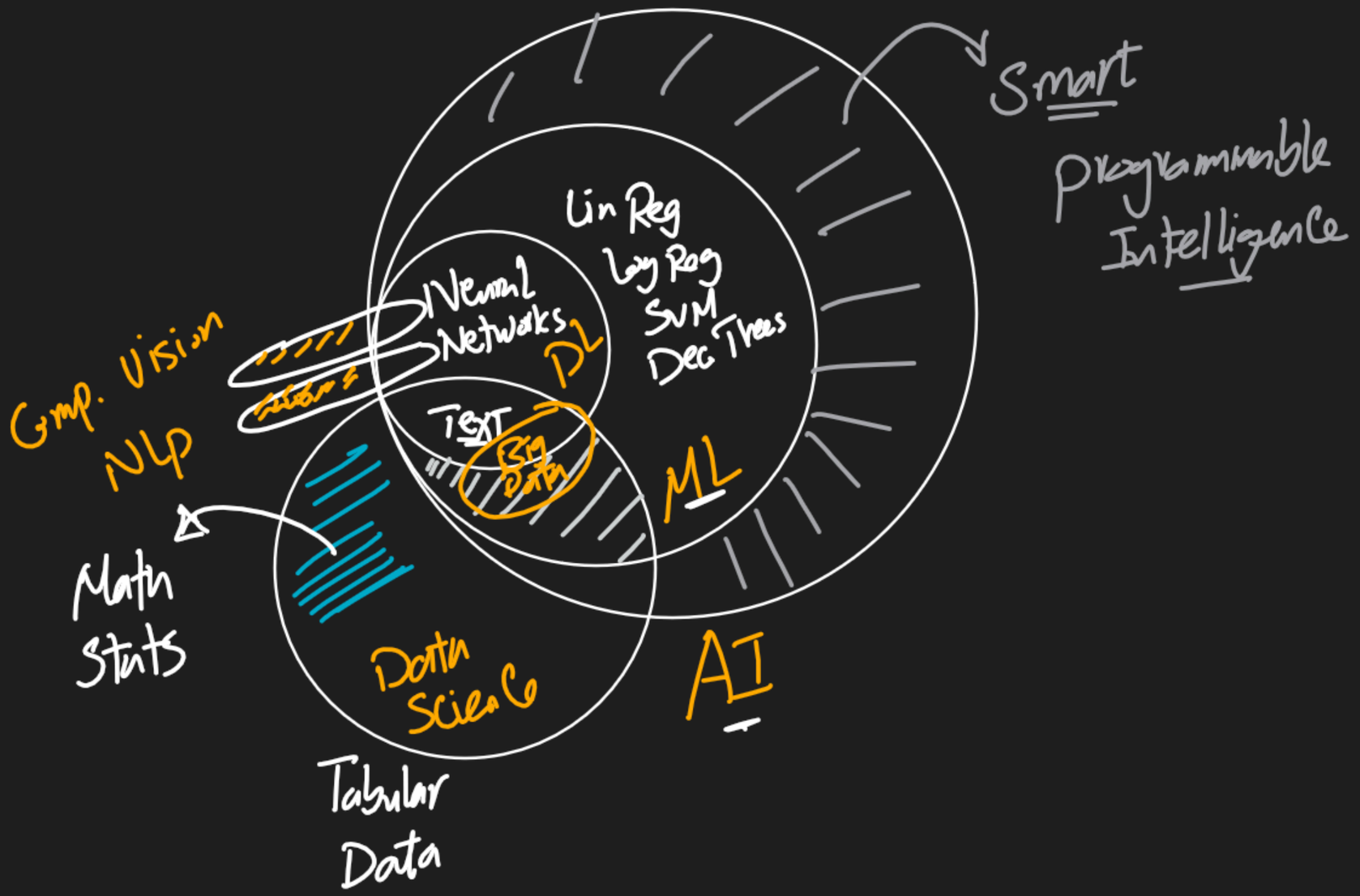
② 1000 Houses
Error: 200

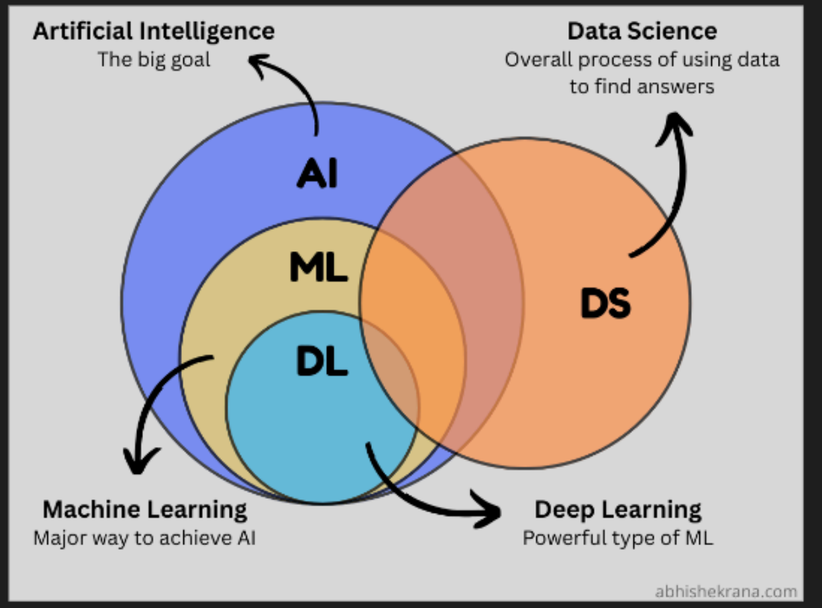
③ 10,000 Houses
Error: 100

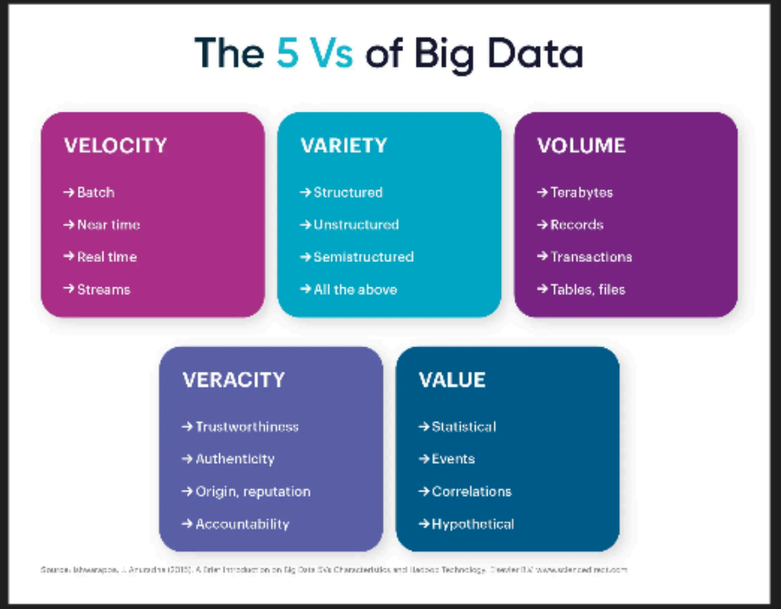

Machine learning
Learning by Experience
Learning from Data
Pattern Recognition

☑ ML vs. Traditional Prog

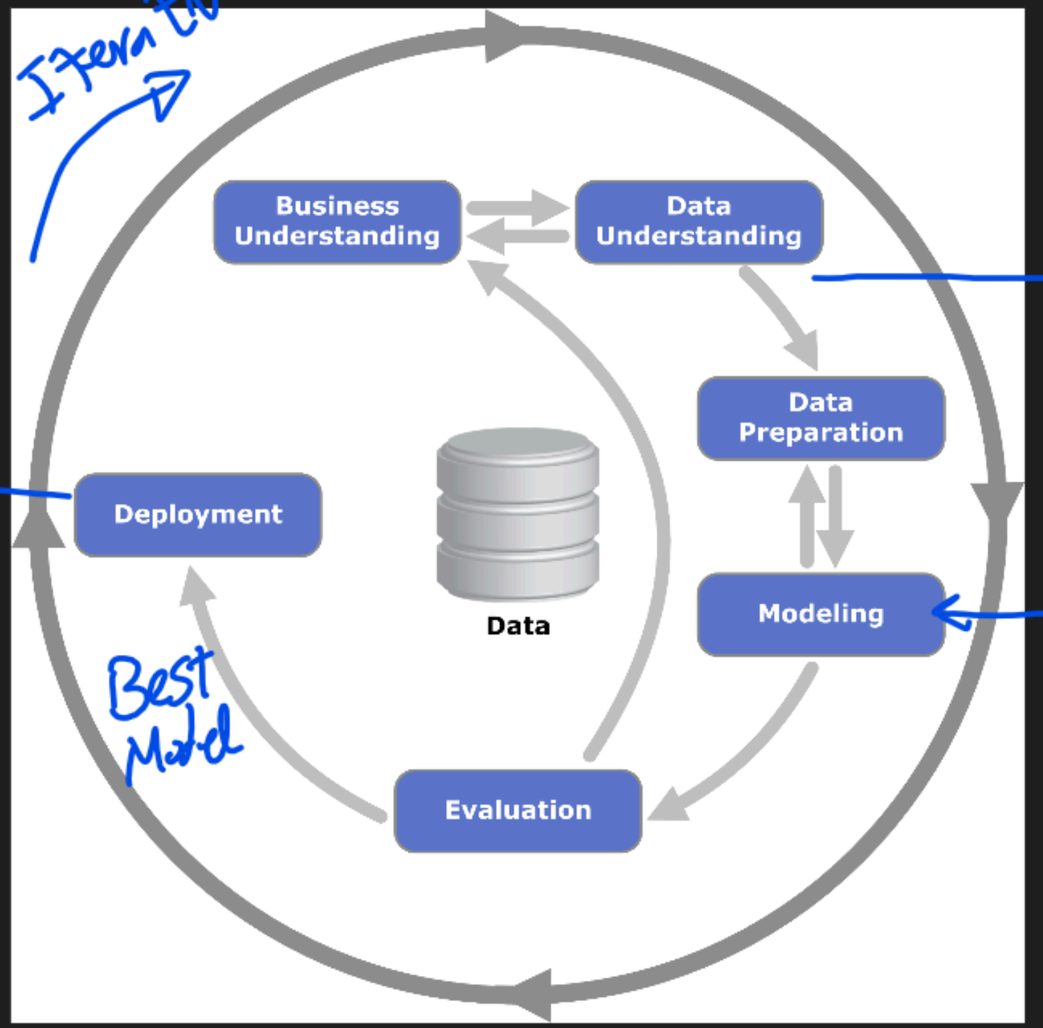



Smart
Programmable
Intelligence
Lin Reg
Log Reg
SVM
Dec Trees
Neural Networks
DL
TEXT
Big Data
ML
AI
Data Science
Tabular Data
Comp. Vision
NLP
Math
Stats


Artificial Intelligence
The big goal
Data Science
Overall process of using data to find answers
AI
ML
DL
DS
Machine Learning
Major way to achieve AI
Deep Learning
Powerful type of ML
abhishekrana.com


The 5 Vs of Big Data
VELOCITY
→ Batch
→ Near time
→ Real time
→ Streams
VARIETY
→ Structured
→ Unstructured
→ Semistructured
→ All the above
VOLUME
→ Terabytes
→ Records
→ Transactions
→ Tables, files
VERACITY
→ Trustworthiness
→ Authenticity
→ Origin, reputation
→ Accountability
VALUE
→ Statistical
→ Events
→ Correlations
→ Hypothetical

☑ Data Science lifecycle



Data Analysis

Current Interest

App

UI

Monitoring & feedback

Data Drift

MLOps

☑ Types of ML

① Supervised ML : Machine is supervised by Human

↳ Labeling

- Questions
- Right Answers

features (X)

Target (label) (Y)

↳ Labeled data

$y = f(x)$

??

Numeric → Regression

Categorical → Classification

- 2 classes → Binary clf
- > 2 classes → Multi class

☑ Regression vs. Classification

Temp 27.2°C

Sunny
Cloudy
Rainy

② Unsupervised ML

→ Unlabeled Data (No target) (features only)

→ Grouping using Similarity → Clustering

→ Anomaly Detection

→ Dimension Reduction

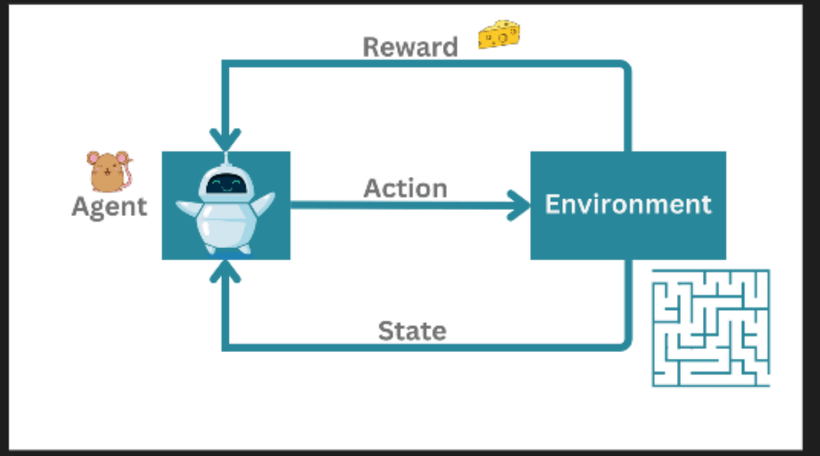
③ Reinforcement Learning

→ No features, No target (At start)

→ Trial & Error

learning (Build your own Data) within → Environment, policy

Agent

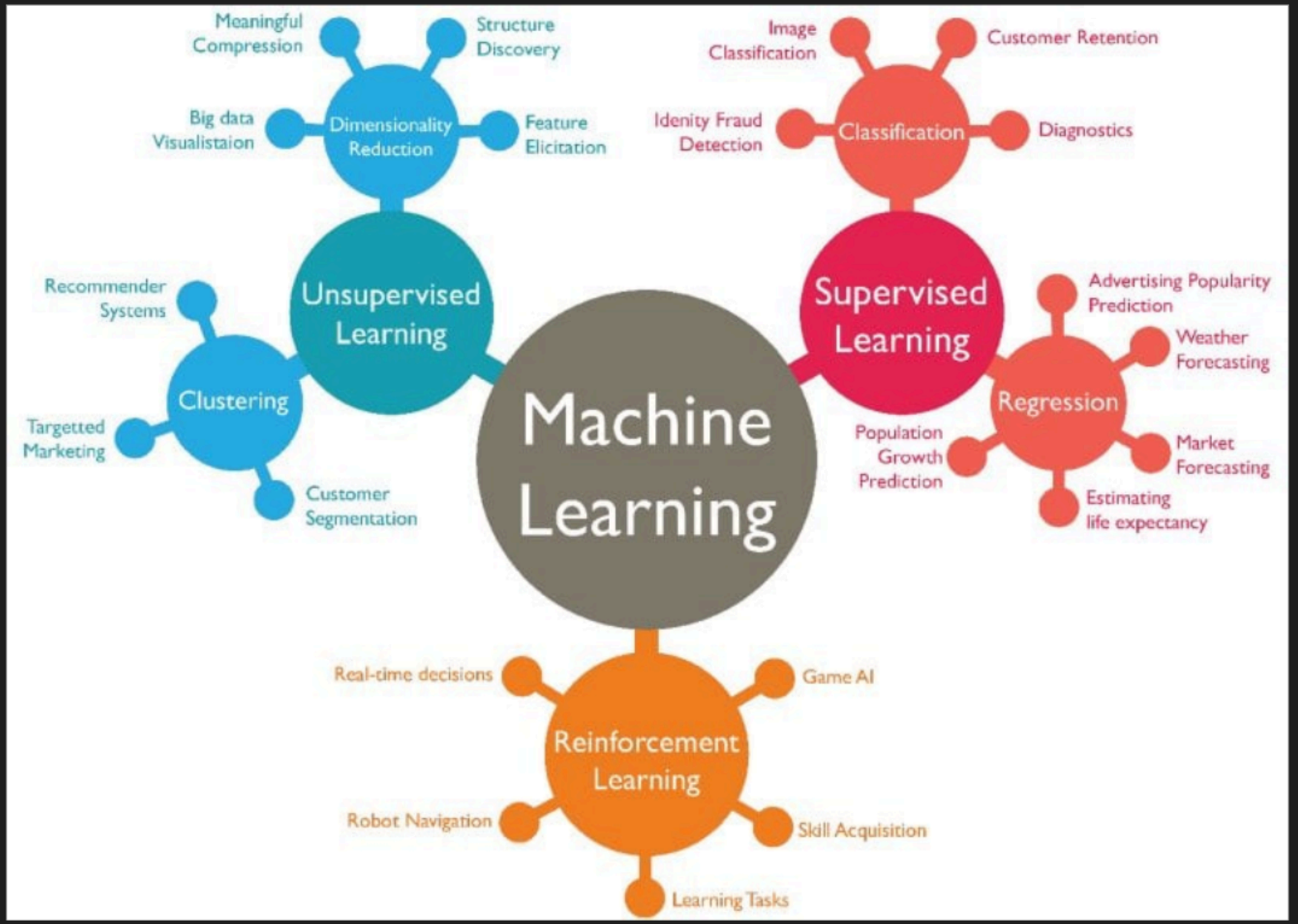


→ Action & feedback

Reward Penalty

State A → Action → State B → feedback

- Exploration vs. Exploitation


Machine Learning
Unsupervised Learning
Dimensionality Reduction
Meaningful Compression
Structure Discovery
Big data Visualistaion
Feature Elicitation
Clustering
Recommender Systems
Targetted Marketing
Customer Segmentation
Supervised Learning
Classification
Image Classification
Customer Retention
Idenity Fraud Detection
Diagnostics
Regression
Advertising Popularity Prediction
Weather Forecasting
Market Forecasting
Estimating life expectancy
Population Growth Prediction
Reinforcement Learning
Real-time decisions
Game AI
Robot Navigation
Skill Acquisition
Learning Tasks

Linear Regression

- Simple Linear Regression (uni-var)
  - one feature

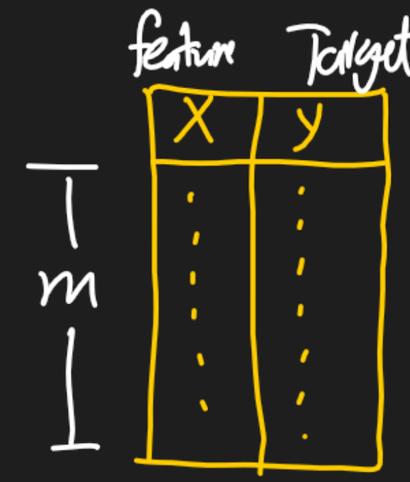
Linear Eqn → $y = c + mx$

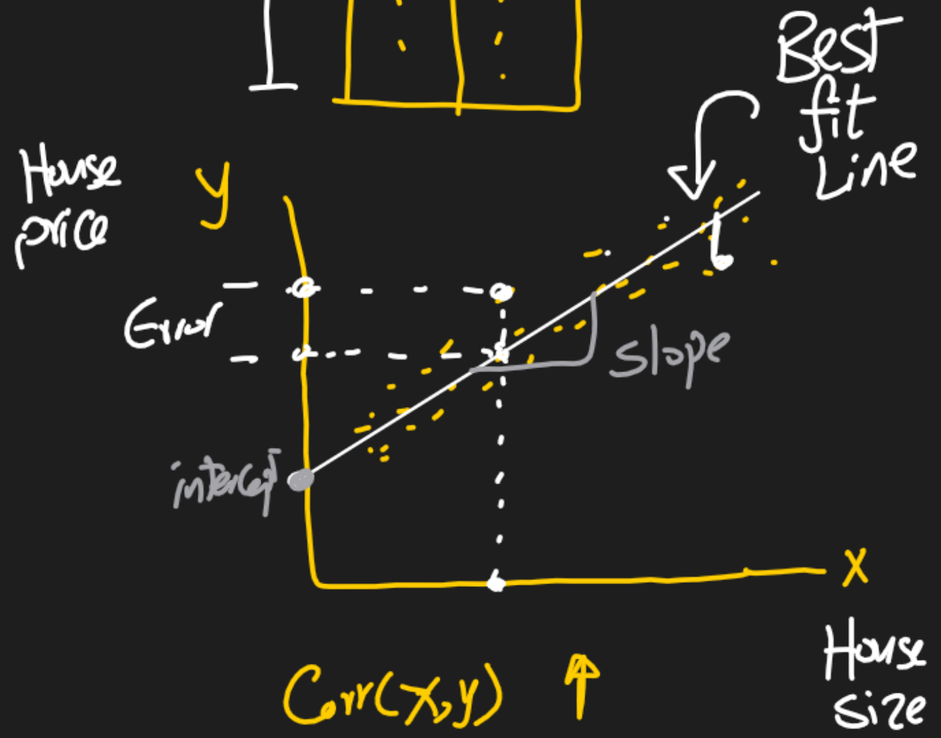
$y = f(x)$

$y = w_0 + w_1 x$

$W = f(x, y)$

Best fit Line
↓
Best weights





Get Best Weights

- Normal Eqn (OLS)
- Gradient Descent

* Cost Function → Measurement of Error
  - y_{actual}
  - y_{pred}

$$J = \frac{1}{m} \sum_{i=1}^{m} \left(y_{actual}^{(i)} - y_{pred}^{(i)} \right)^2 \rightarrow \text{MSE}$$

objective

Min Error

$w_0, w_1 \rightarrow$ Line 1 $\rightarrow J_1 = 50$

$w_0', w_1 \rightarrow$ Line 2 $\rightarrow J_2 = 20$

$w_0'', w_1 \rightarrow$ Line 3 $\rightarrow J_3 = 40$

} → Best fit line → Line 2

Min Error

$$J(w_0, w_1) = \frac{1}{m} \sum_{i=1}^{m} \left(y_{actual}^{(i)} - w_0 - w_1 x^{(i)} \right)^2$$

unvectorized

$$\frac{\partial J}{\partial w_0} = 0 \quad \Rightarrow \quad w_0 = \checkmark$$

$$\frac{\partial J}{\partial w_1} = 0 \quad \Rightarrow \quad w_1 = \checkmark$$

$\frac{\partial J}{\partial w} = 0$ → OLS (vectorized) → Normal eqn

$$W = (X^T X)^{-1} X^T y$$

$$[W] = [w_0 \; w_1]$$

$$[X] = \begin{bmatrix} 1 & x \\ 1 & x_0 \end{bmatrix}$$

Closed Solution

Iterative → Gradient Descent

① Random initialization of weights

② Calculate $\partial J / \partial w$ (direction)

③ update weights : $w^{new} = w^{old} - \eta \frac{\partial J}{\partial w}$

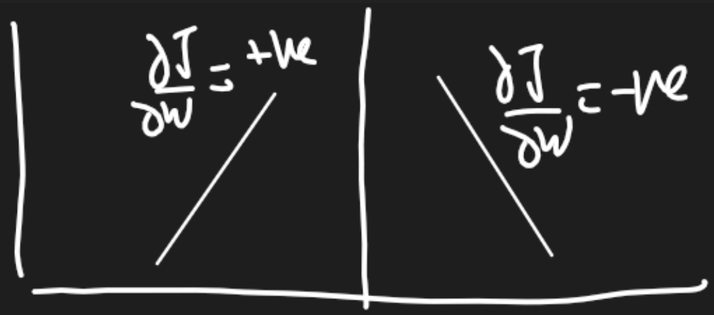
↳ learning Rate (control)

Till $\frac{\partial J}{\partial w} \approx 0$ (loop)

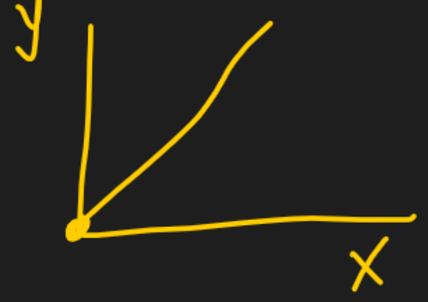
↳ Best Weights

Gradient Descent

$$J(w) = \frac{1}{m} \sum_{i=1}^{m} \left(y_{actual}^{(i)} - w^T x^{(i)} \right)^2$$



$w_0 = 0 \Rightarrow \boxed{y = w_1 x}$



$$w^{new} = w^{old} - \eta \frac{\partial J}{\partial w}$$

Min Error

- $-ve$ → $w^{new} > w^{old}$
- Zero → $w^{new} = w^{old}$
- $+ve$ → $w^{new} < w^{old}$