Revision - Week-1

ML is the study of algorithms that improves automatically through experience & by the use of data.

→ Weather Prediction.
→ Car rate " " etc.

Why ?
→ Scale / Speed / cost of human labor.

Data

	feature 1	feature 2	feature 3	Label
→ O_1	a	b	c	∈ α
→ O_2	d	e	f	∈ β
→ O_3	g	h	i	∈ γ

Model :-
Mathematical simplification of reality.

Types of ML Algorithms

→ Supervised & Unsupervised.

* Supervised learning :- Label is present.

Regression

→ label is continuous
→ Curve fitting

$$loss = \frac{1}{n} \sum (f(x^i) - y^i)^2$$

$$f(x) = w^T x + b$$

If f & g are 2 different model,
loss[f] < loss[g] then
f is better model than g.

Classification

→ label is discrete ∈ (-1, 1) or (0, 1)

$$loss = \frac{1}{n} \sum_{i=1}^{n} 1(f(x^i) \neq y^i)$$

$$f(x) = sign\ (w^T x + b)$$

Evaluating learning models :-

→ Test data
→ Training data
→ Validation data.

Unsupervised :- label is not present in the data.

→ Understanding data.
→ Build models that compress, explain & group data.

→ Dimentional Reduction

→ Compression & Simplification.

→ Encoder & Decoder

Encoder → Compress
Decoder → Undo

Goal :-

$$g(f(x^i)) \simeq x^i$$

$$\text{loss} = \frac{1}{n}\sum_{i}^{n} \| g(f(x^i)) - x^i \|^2$$

Density estimation :-

A density estimation model takes in several samples from a random source, & outputs a model that assigns a probability score to every possible instance.

$$\text{loss} = \frac{1}{n}\sum_{i=1}^{n} -\log_{10}(P(x^i))$$

Ex- To increase the no of observation