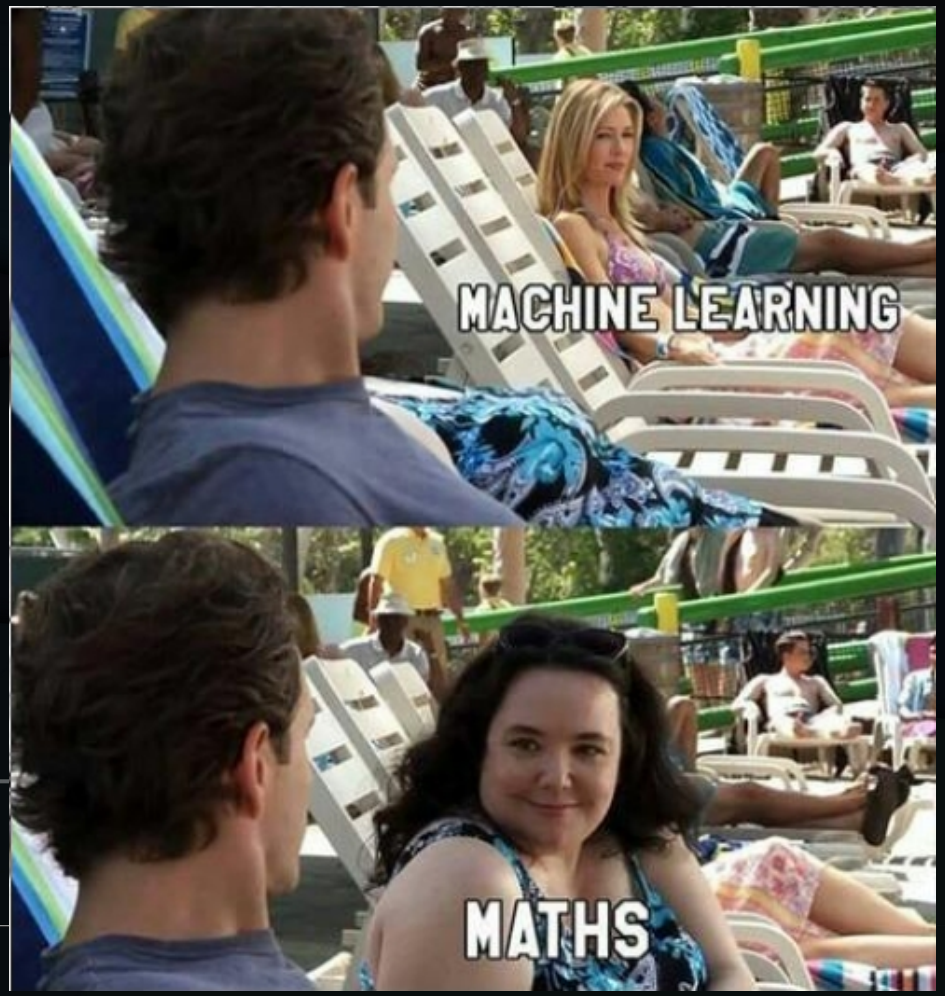
@me – Yash Singh

SME – ML/AI and OpenSauce

contributor to various packages

Now, before I start, I must tell you guys, maths is extremely essential for understanding the concepts and core functioning of Machine Learning. As you'll later learn, without maths, you cannot simply derive the relation between input data and the output obtained.



Artificial Intelligence (AI)

- The hot topic of Tech World
- Decision Dictated Scenario
- Derivation from Quantum Computing
- Use cases

Machine Learning (ML)

- Machine Learning AKA Statistical Learning is a subset of Artificial Intelligence which introduces closer derivation of procedure dictated logistic programming to the masses. It works with algorithms that works on probability inclination to fetch a data vector closest to the desired target.

Types of Machine Learning (ML)

- Machine Learning consitutes of thousands of types. Generically tho, they’re categorized as these three :
- Supervised
- Unsupervised
- Reinforcement

Supervised Learning

- Supervised Learning is one of the most prominent types of Machine Learning implementation used by beginners as well as advance pro Data Scientists.
- It involves training the model on a training dataset to make the model more accurate.

Unsupervised Learning

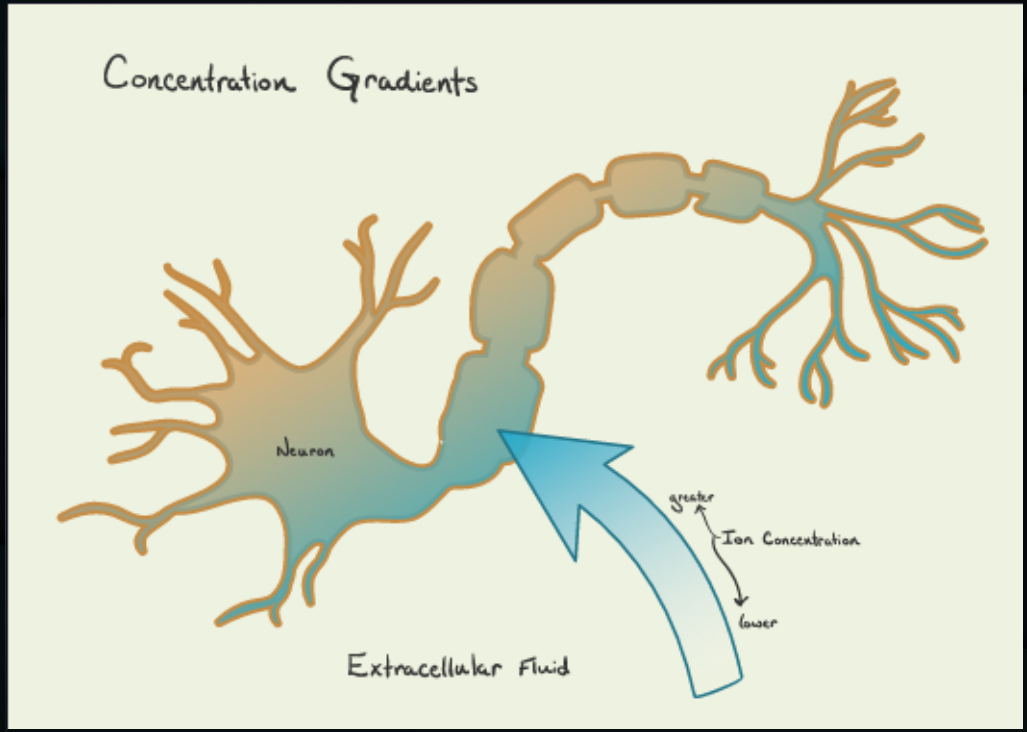
- Unsupervised Learning is established as one of the ways of implementing Machine Learning when the model has to operated without a prediction training prior to application. This type of implementation is purposely selected in classification handling.

Reinforcement Learning

- This type of model is preferred for a job where the model is rewarded with some sort of positive reinforcement everytime it manages to clear an exit node. These models work on the basis of trial and error for higher efficiency.

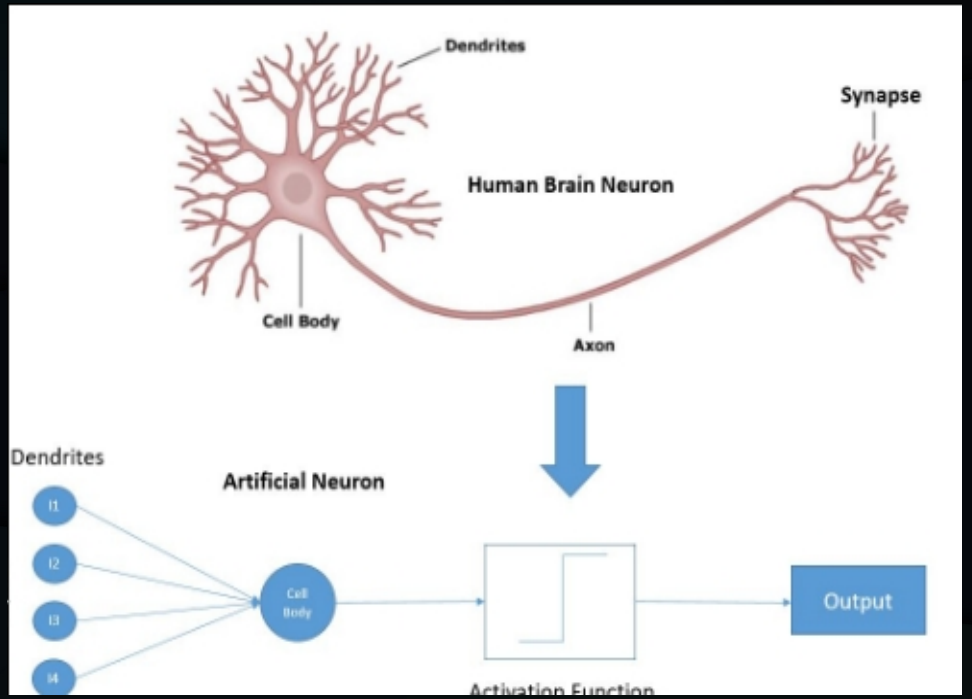
Neural Networks

- Biological aspect of neurons.
- Impulse travel by means of neurotransmitters
- Communication across synaptic cleft



Neural Networks

- Technical aspect of Neurons
- Hidden neurons process tree
- Neural Network & it’s functioning



Types of Supervised Learning

- Classification
- Classification deals with output data format comprising of types and/or categories
- Regression
- Regression deals with output data format in numerical nature

Machine Learning Elements :

- Machine Learning Elements can be encapsulated in these 4 subtypes :
- Data
- Model
- Objective Function
- Optimization Algorithm

Linear Model Illustration with Exemplification

- Let’s consider a simple intercept function given as
  - $y = mx + c$
- Where, y is the output, m is the slope and c is the intercept point/range.

Translation to Machine Learning

- The same function when translated to Machine Learning is given as
  - $f(x) = xw + b$
- where w is the weight and b is the bias

- Weight
- Weights control the signal (or the strength of the connection) between two neurons. In other words, a weight decides how much influence the input will have on the output.

- Bias
- Biases are offset values acting as an additional input into the next layer that will always have the value of 1.

Example to tackle Weight-Bias Problem

- Let’s say we have to predict the cost of an apartment based on the input – it’s area expressed in sq. Feet.
- Using the model expression $y = xw + b$
- Assumption – the weight is dealing as the cost per sq. Feet, depending on the locality, whether the apartment is fully furnished or not, and similar factors.
- Bias talks about the sole quality of the input parameter, whether the apartment is close to a beach, has a nice glass view of the ocean, etc.

- CASE I:
- For an apartment of size x = 763 sq feet,
- Let the model decide a weight of w = 300 rupees per sq feet, and the bias b = -1000 (discount for not so good look of the apartment).
- Hence, cost of apartment = f(x) = xw + b = 227900 rupees.

- CASE II:
- For another flat in Goa of the same size x = 763 sq feet, we have a weight balance of w = 550 rupees per sq feet, and the bias is given as b = +5000 rupees because the apartment is new, and looks great.
- Now, the cost of the apartment = f(x) = xw + b = 424650 rupees.

N-input Model

- For higher number of parameters fed to the model, like the price, stories, services, etc,
- the cost is given as
  - $y = x_1w_1 + x_2w_2 + ... + x_nw_n + b_1 + b_2 + ... + b_n$

Objective Function

- Objective Function is the evaluation handler of how good or bad a model is.
- Works as two constraints :
-

  Loss function Reward Function

L2 - norm

- Squared loss function used for regression.
- Given as
  - $\text{l2-norm} = \Sigma_i (y_i - t_i)^2$
- It is the summation of square of difference between target vector and
- output vector.

Cross Entropy

- Loss function for classification Supervised Models
- Given as

$$L(y, t) = -\Sigma_i t_i \ln y_i$$

- Hence, it is the summation of product of target vector and natural log of output vector

Example to Cross-Entropy

- Let's say a supervised model trained over alphabet sets A, B and C.
- An image of B is processed by the model,
  Hence, target vector = t_i = [0, 1, 0], where 1 implies true for B.
- Output vector obtained from model
  y_i = [0.2, 0.8, 0.1]
- Hence, cross entropy is given as
  L(y, t) = -0 x ln 0.2 – 1 x ln 0.8 – 0 x ln 0.1 = 0.22314

Optimization Algorithm

- Algorithm that handles the accuracy of the model output by means of variance in the bias and weights.
- Most commonly used Algo – SGD (Stochastic Gradient Descent) that works over probability distribution.
- SGD is calculated as :

$$x_{(i+1)} = x_i - \eta f'(x_i)$$

Functioning of SGD

- Let’s say the user performs over a linear model relation given as
  - $y = 5x^2 + 3x – 4$
- By formula, First derivative $y’ = 10x + 3$
  Pick arbitrary number for $x0 = 4$,
  Calculating $x1 = 4 – \eta[10 * 4 + 3] = 4 – \eta 43$
- Where η is the learning rate of the model

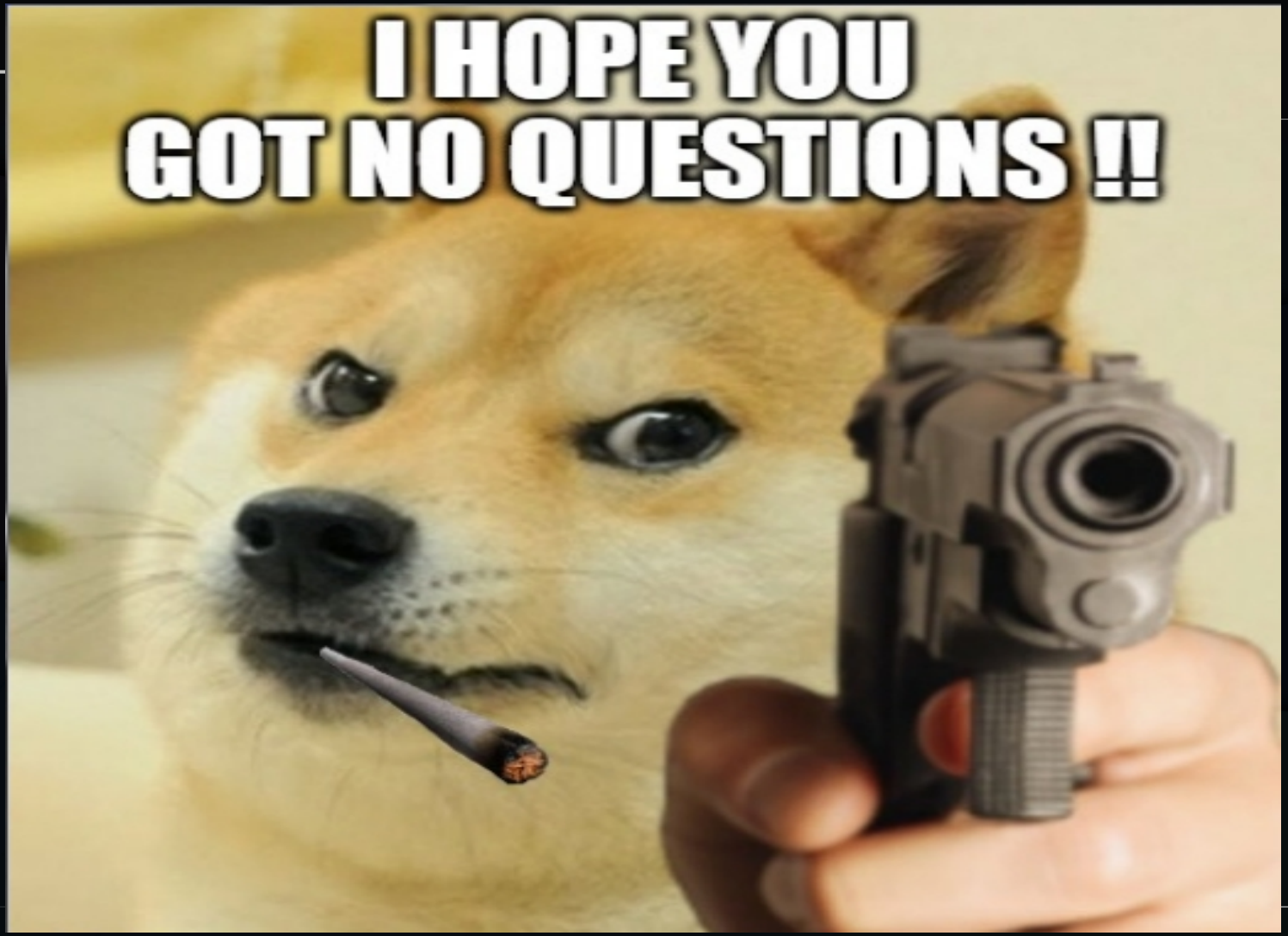
- Learning Rate η = NULL/0
- Learning Rate should be iterative difference
- Oscillation of Minimum
- SGD step limit – 0.001
- Relevance to Rayleigh’s Power Method

$$x_{(i+1)} - x_i > \text{step}$$

- Underfitting – The model has not captured the underlying logic of the data
- Overfitting – The model has trained over a particular dataset so much that it has “missed the point”
- Bias-variance tradeoff – The aggregate balance for model stability

Pursuing ML/AI

- Elasticity to relevant ideas and their scope
- A clearer overview and wider perception for handling problems
- Obvious requirements – calculus, linear algebra, probability distribution, well versed in atleast one HLL (ofc implementation on any platform is possible, but for beginners, it’s advisable to continue with languages that offer predef and well supported community libraries, like Java, Python, GoLang, JS)


I HOPE YOU
GOT NO QUESTIONS !!

THANK YOU!

- Visit me

  https://probe2k.github.io

- Contact

  - +91-8210832018
  - kyash054@gmail.com