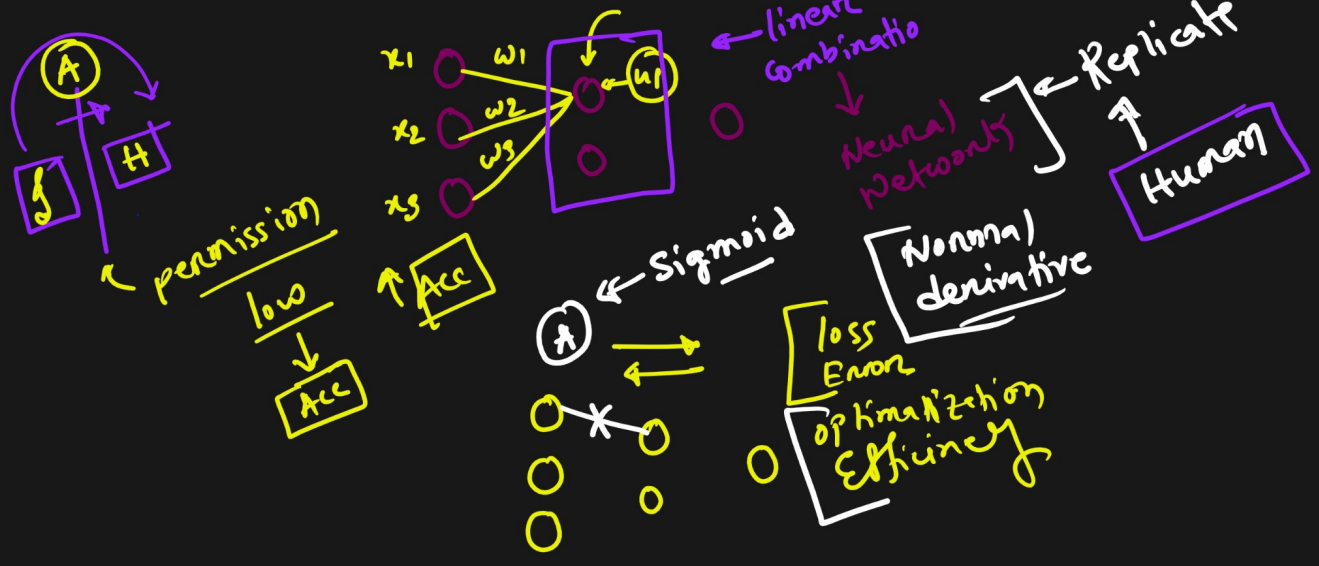
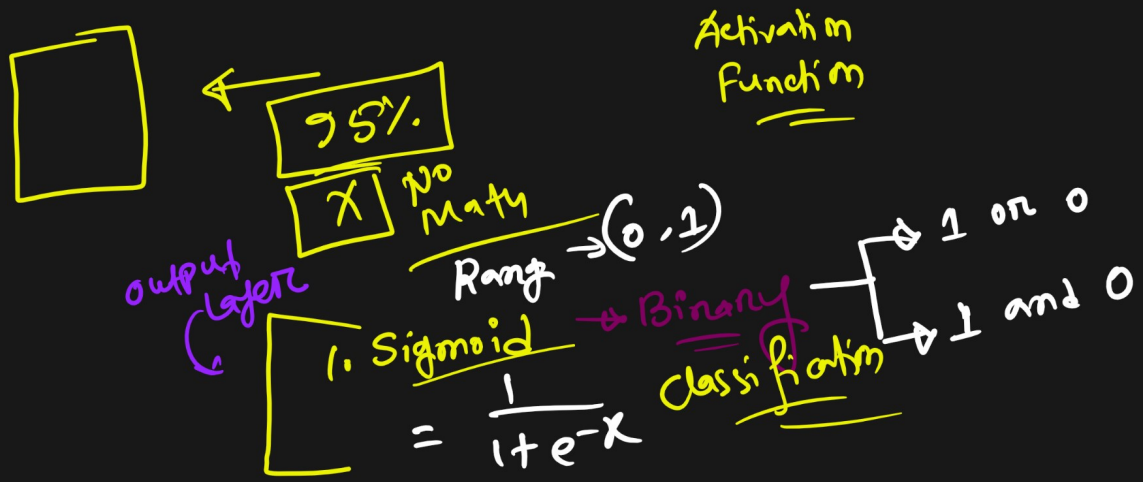
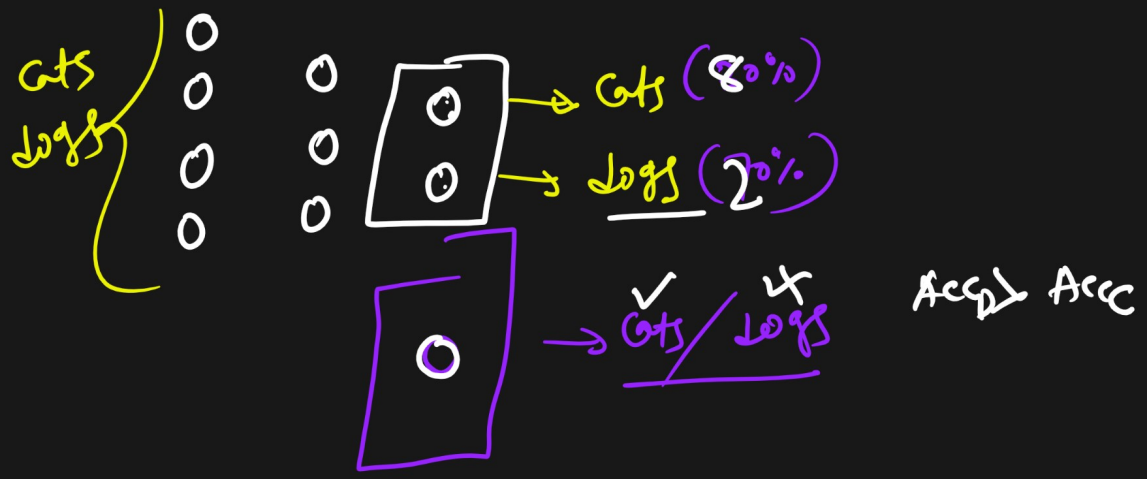
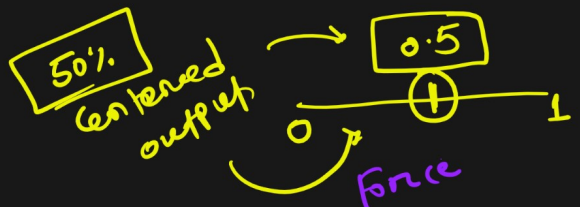
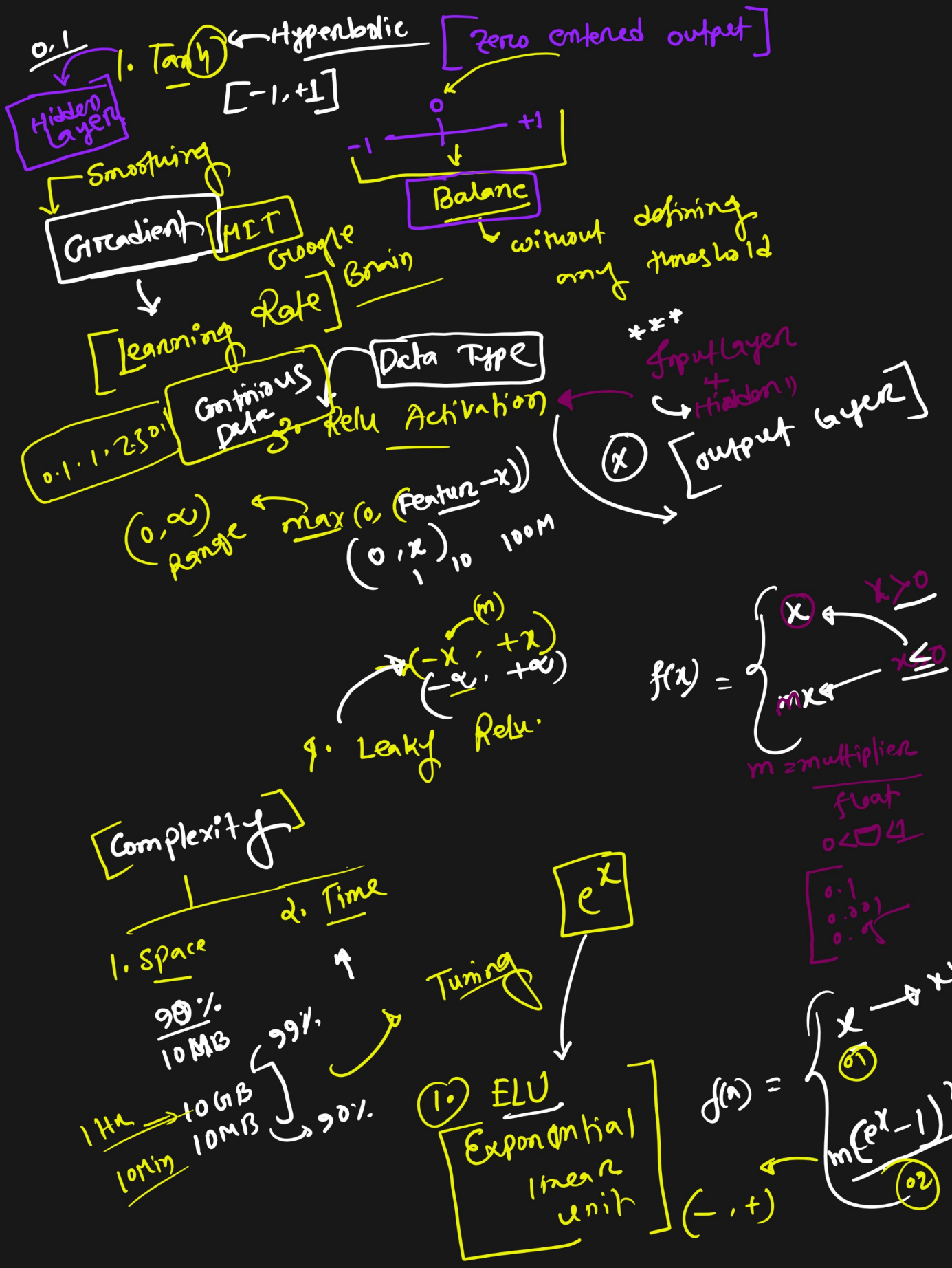

A
H
permission
low
Acc
x1
x2
x3
w1
w2
w3
linear combinatio
Neural Networks
Replicate
Human
Sigmoid
Normal derivative
loss Error
optimization Efficiency


95%
X
no Maty
Activation Function
output layer
Range → (0, 1)
1. Sigmoid
= 1 / 1+e^-x
Binary classification
1 or 0
1 and 0


Cats
Dogs
Cats (80%)
Dogs (20%)
Cats / Dogs
Accpt Accc


50%
Contained output
0.5
1
0
Force


0.1
1. Tanh
Hyperbolic
[Zero centered output]
[-1, +1]
Hidden Layers
0
-1
+1
Smoothing
Balance
Gradient
MIT
Google
Brain
without defining any threshold
[Learning Rate]
***
Input Layer + Hidden
[output Layer]
Data Type
0.1, 1, 2.30
Continious Data
Relu Activation
(0, ∞)
Range
max (0, (Feature - x))
(0, x)
10
100M
(m)
(-x, +x)
(-∞, +∞)
2. Leaky Relu.
f(x) =
x
x ≥ 0
mx
x ≤ 0
m = multiplier
float
0 < m < 1
[Complexity]
1. Space
2. Time
90%
10MB
1Hr → 10GB
10Min 10MB
99%
>0%
Turing
e^x
0.1
0.321
0.9
1. ELU
[Exponential linear unit]
(-, +)
f(x) =
x
x > 0
01
m(e^x - 1)
x ≤ 0
02

0 - 1
-1 +1
-∞ , +∞

(2) SELU

Scaled → Normalization

Mean
Variance
↓
Stable
(0 ≠ 1)
→ Std
↳ optimal

$$f(x) = \lambda \begin{cases} x & x > 0 \\ \alpha(e^x - 1) & x \le 0 \end{cases}$$

Swish

$= f(x) =$ $x \cdot \sigma(x)$

smooth ← Swish

(0 - 1)
↑
0 - 1

Real

(+)

Mish

(0, ∞)

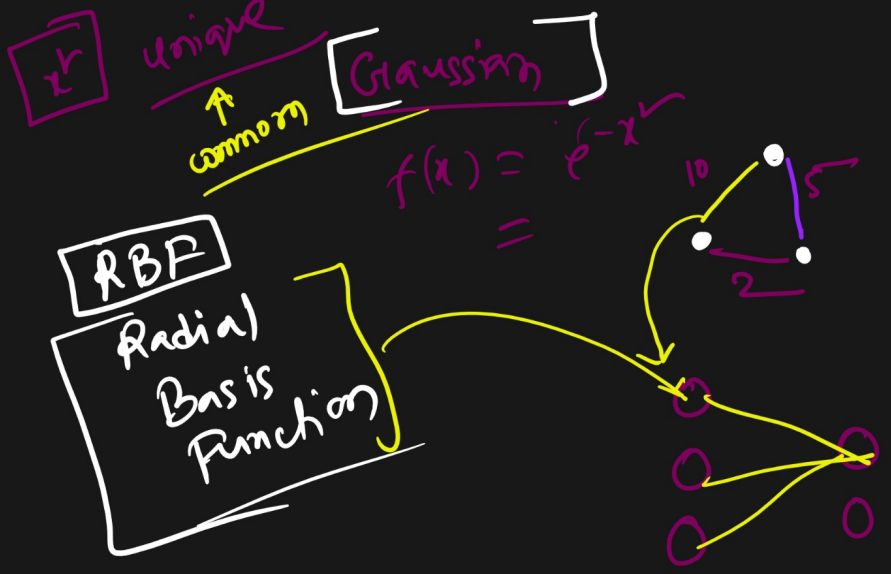
-1, +1

Gradient Smoothing

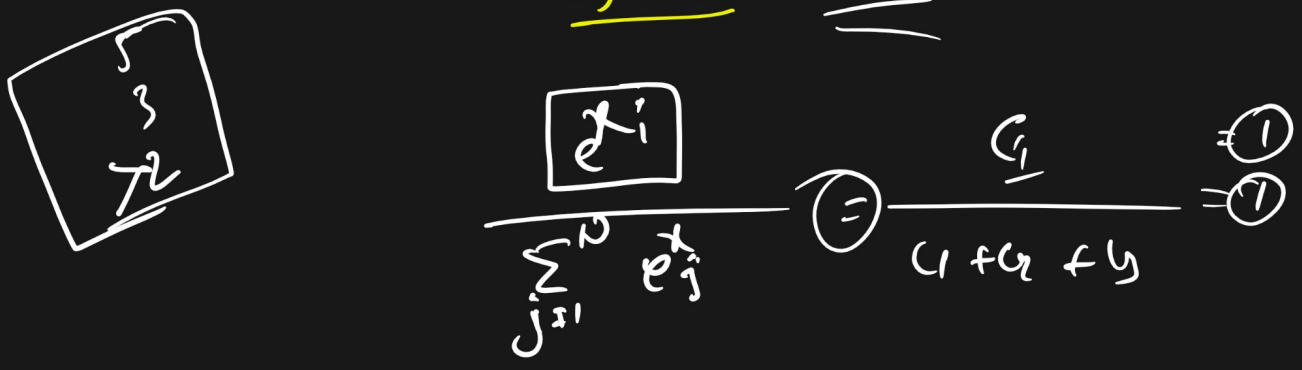
Computer vision

Relu + Swish

$f(x) = x \cdot \tanh\left(\ln(1+e^x)\right)$

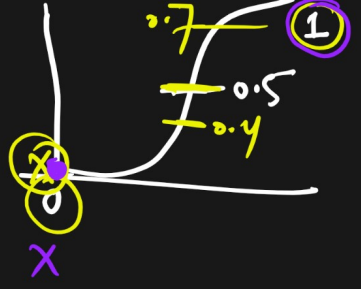


Softmax Multiclass



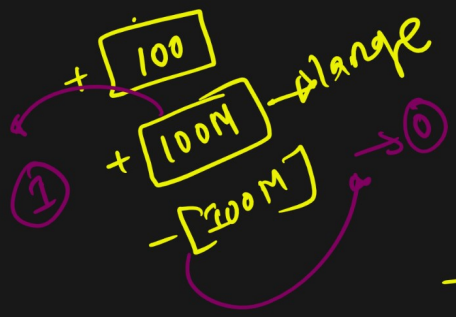
1. Sigmoid

$$f(a) = \frac{1}{1+e^{-x}}$$

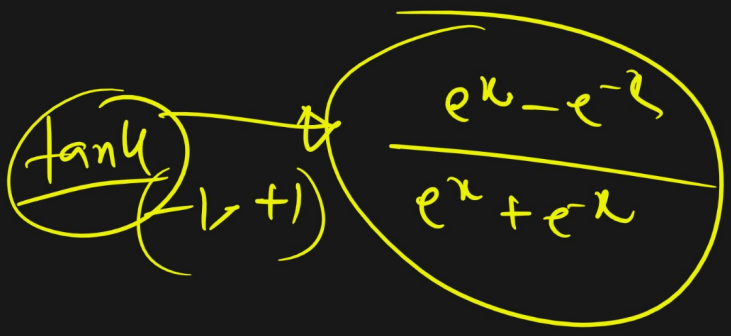


Real:
1. Smooth Curve
2. Non-linear
3. Symmetry
Never Cross 0 value

problem
1. Vanishing Gradient Problem
2. Not zero Centered
3. Saturation was too much slow



$$f(x) = f(x)\,(1 - \sigma(x))$$



Chain Rule

$$e^x = \frac{1}{1+e^{-x}}$$

$$= \boxed{(1+e^{-x})}^{-1}$$

m

$$m = (1+e^{-x})$$

$$\sigma(x) = m^{-1}$$

$$= m^{-1}$$

$$= -1 \times m^{-1-1}$$

$$= -m^{-2}$$

$$\frac{d}{dx}\sigma(x) = -m^{-2} \times \frac{d}{dx}(1+e^{-x})$$

$$\frac{d}{dx}(1) + \frac{d}{dx}(e^{-x}) \qquad = -e^{-x}$$

$\rightarrow 0$

$$\frac{1}{1+e^{-x}}$$

$$= (-1)(1+e^{-x})^{-2} \times (-e^{-x})$$

$$= e^{-x}(1+e^{-x})^{-2}$$

$$\sigma(x) = \frac{1}{1+e^{-x}}$$

$$1-\sigma(x) = \frac{e^{-x}}{1+e^{-x}}$$

$$\frac{e^{-x}}{(1+e^{-x})}$$

$$\sigma'(x) = \sigma(x)(1-\sigma(x))$$