

CS 307

FALL 2023

DALPIAZ

WEEK 15

WHAT IS A NEURAL NETWORK? ^{ARTIFICIAL}

- A FUNCTION
 - OF INPUT DATA AND "UNKNOWN" PARAMETERS
 - OFTEN REPRESENTED AS A NETWORK
- PARAMETERS LEARNED FROM DATA

LOGISTIC REGRESSION

SIGMOID FUNCTION

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

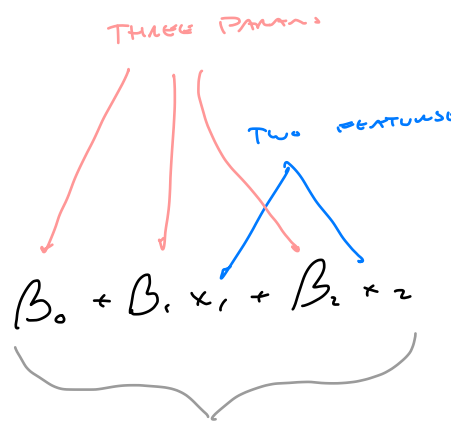
$$\sigma: \mathbb{R} \rightarrow [0, 1]$$

$$p(x) = P[Y=1 | X=x]$$

$$\log\left(\frac{p(x)}{1-p(x)}\right) = \underbrace{\beta_0 + \beta_1 x_1 + \beta_2 x_2}_{\text{"LOGIT"}}$$

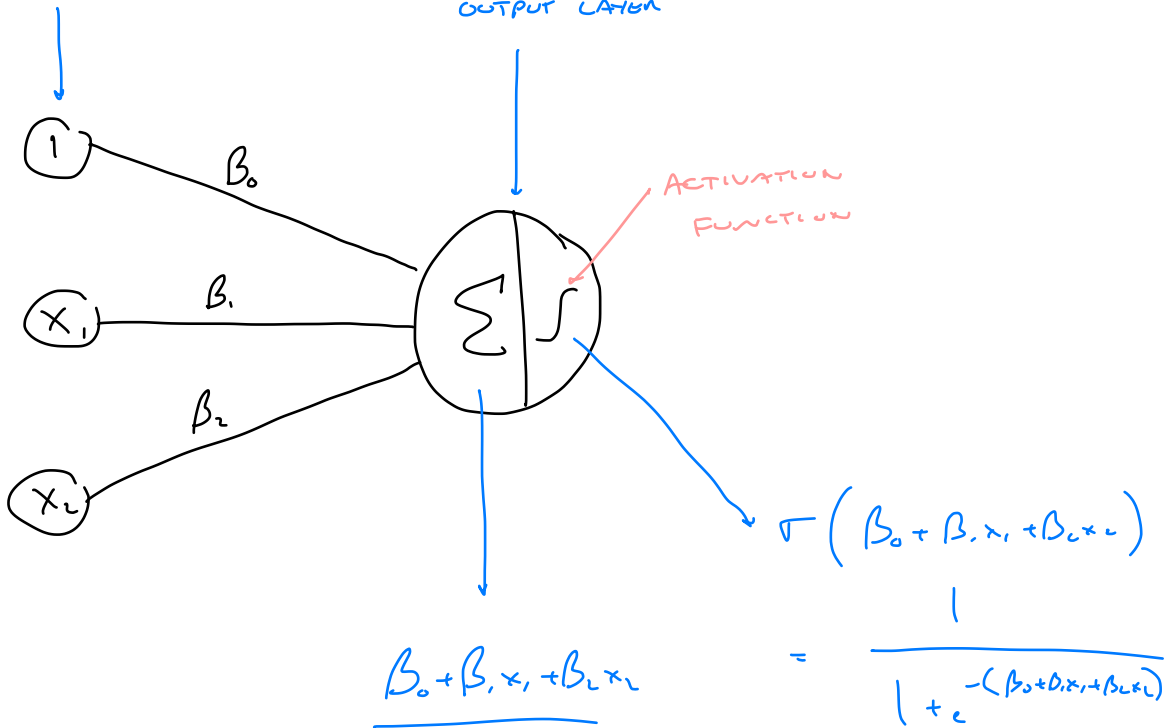
THREE PARAMS

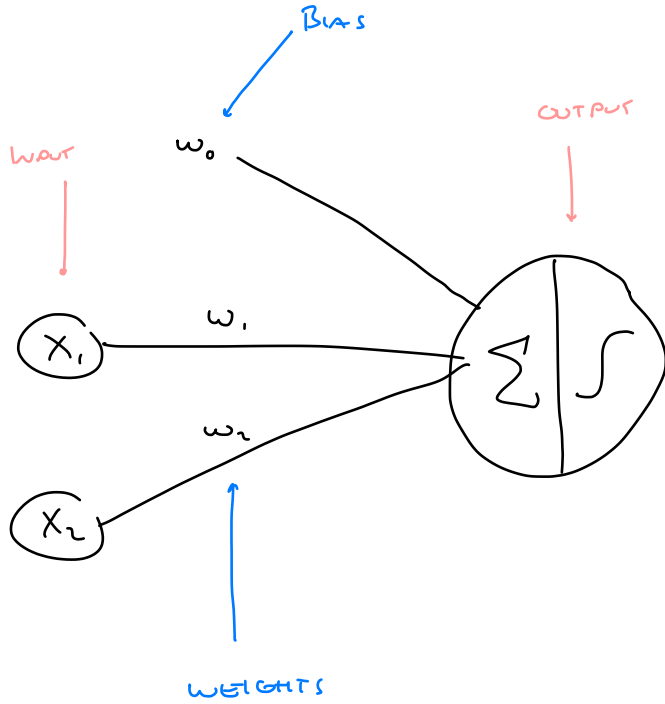
TWO FEATURES

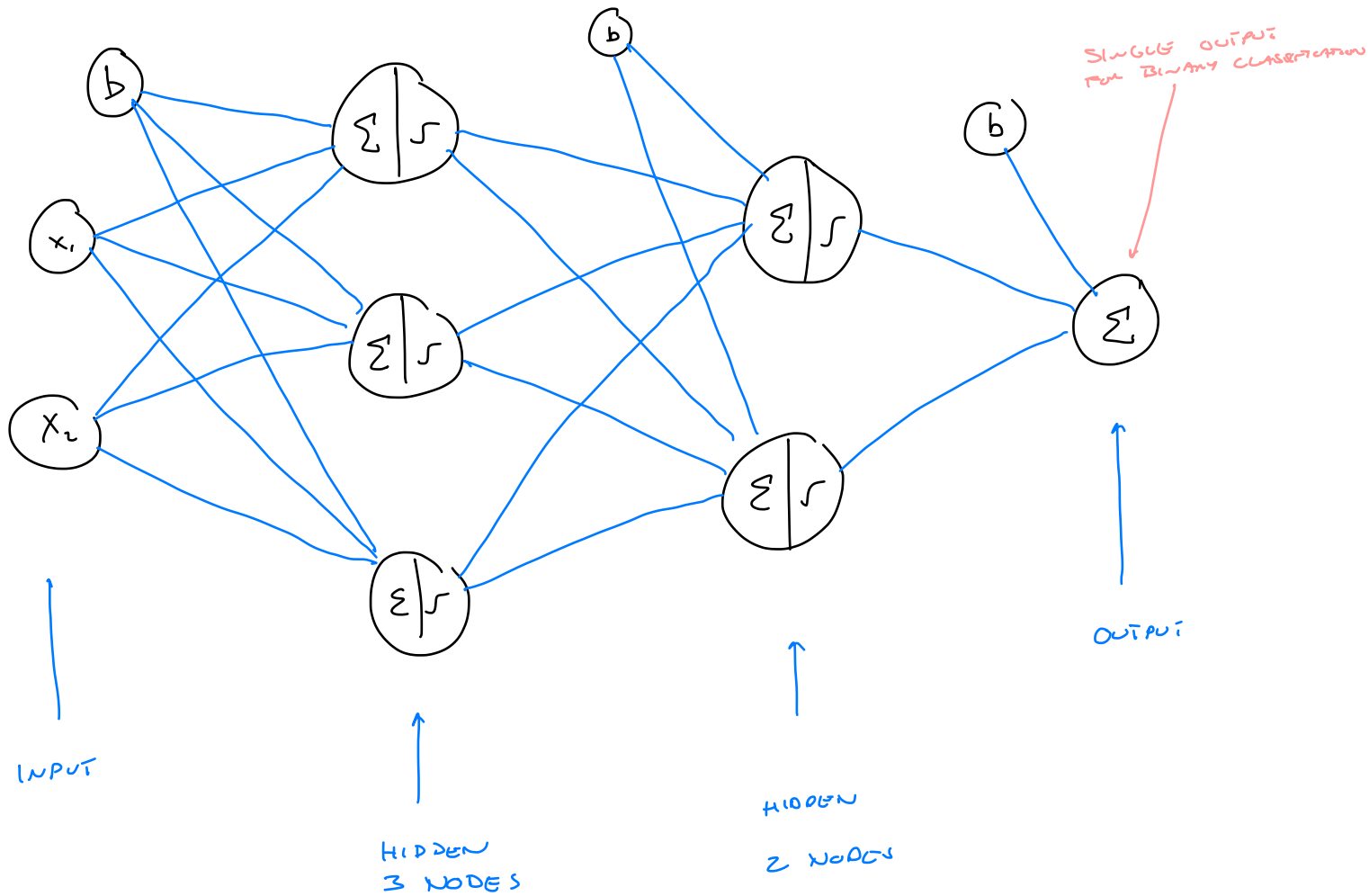


INPUT LAYER

OUTPUT LAYER







WHAT DO WE CONTROL?

- How many HIDDEN LAYERS?
- How many NEURONS IN A LAYER?
- WHICH ACTIVATION FUNCTION?
- How TO OPTIMIZE?
 - METHOD
 - LEARNING RATE

ReLU

WHAT IS DEEP LEARNING?

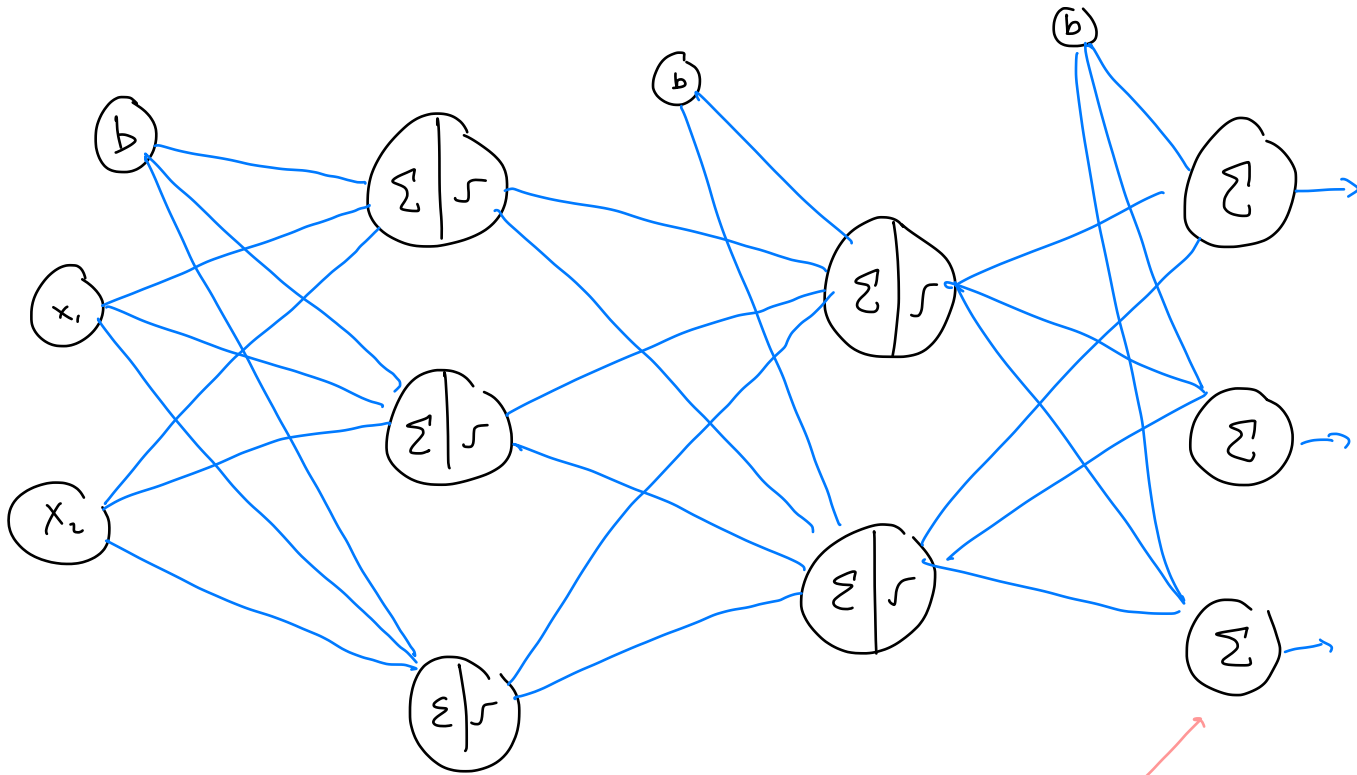
Bio NEURAL NETWORKS

How to TRAIN NNs?

"How to LEARN WEIGHTS FROM DATA?"

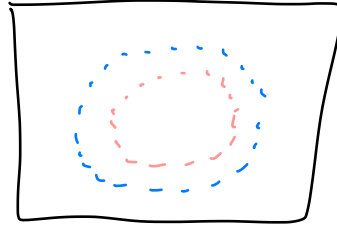
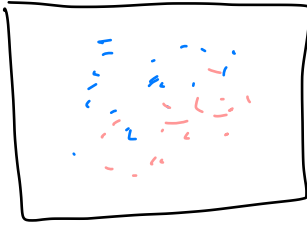
OPTIMIZATION PROBLEM

- LOSS FUNCTIONS
- SGD / ADAM
- CHAIN RULE

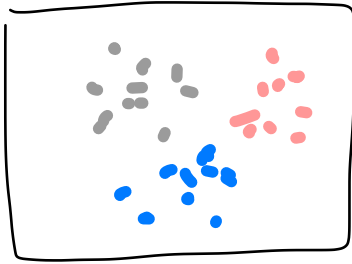


ONE NODE PER CATEGORY
FOR MULTI-CLASS CLASSIFICATION

1. FIT EXAMPLE MODELS TO ^{→ BINARY} TEST DATA



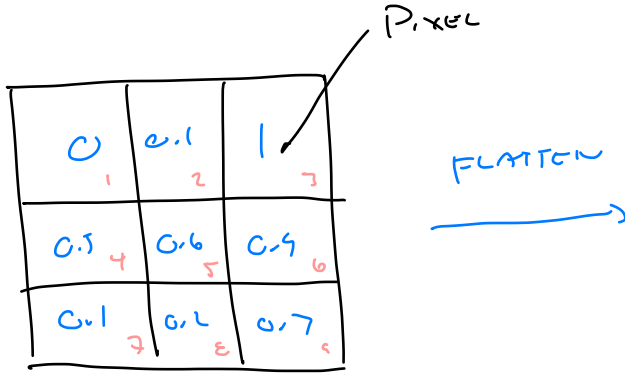
2. FIT EXAMPLE MODEL TO MULTI CLASS DATA



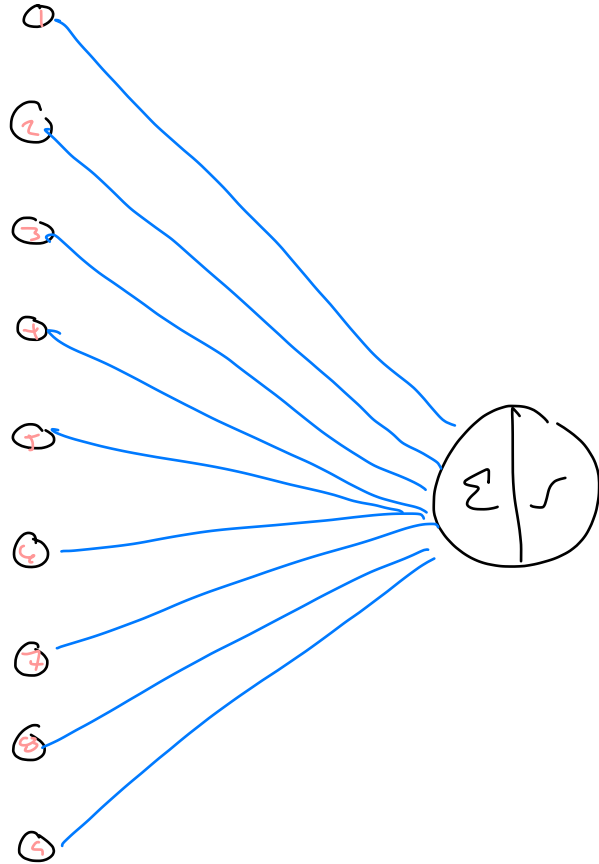
3. FIT CNN TO "CLASSIC" IMAGE DATASET

- KMNIST

IMAGE DATA



3x3 image



CONVOLUTION OF IMAGES

1	2	3
4	5	6
7	8	9

3x3 IMAGE

1	0
0	-1

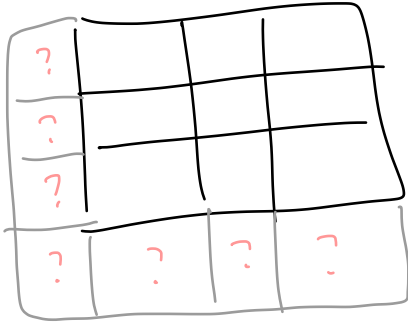
2x2 KERNEL

-4	-4
-4	-4

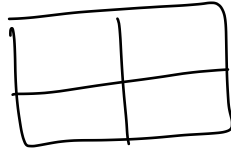
OUTPUT

$$(1 \times 1) + (2 \times 0) + (4 \times 0) + (5 \times -1) = -4$$

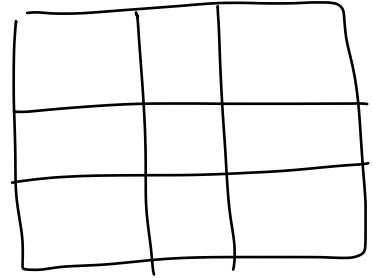
PADDING



3x3 image
padding of 1



2x2 kernel



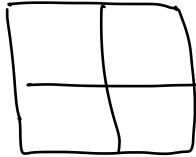
3x3 output

MAX Pooling

→ OR AVE

1	2	5	6
3	4	7	8
9	10	0	1
11	12	-1	4

4x4 image



2x2 pooling

4	8
12	4

2x2 output