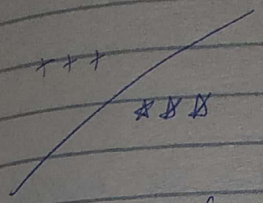
29-11-2021 Monday

Week #14 Lecture #17

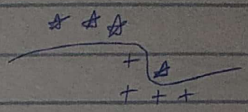
Paper Discussion

(Q#1)



Linear classifier

Wx



Non-linear classifier.

$W_3(W_2(W_1X)) = WX$

as a linear classifier.

λ = Regularization strength

Hyperparameter

L_2 regularization also prefers smaller weights & also prefers diffused weights

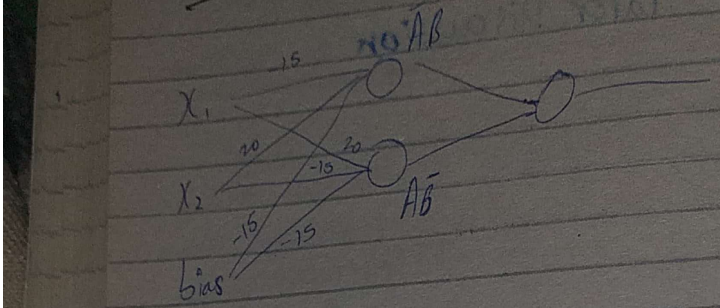
(Q#1) Non-linear Mapping function characteristics:

a) Don't saturate output

b) easy to compute

c) ~~Abs~~ Zero centered.

Q19) XOR:-



we have to show what could be the weights in this NN

X_1	X_2	
A	B	XOR
0	0	0
0	1	1
1	0	1
1	1	0

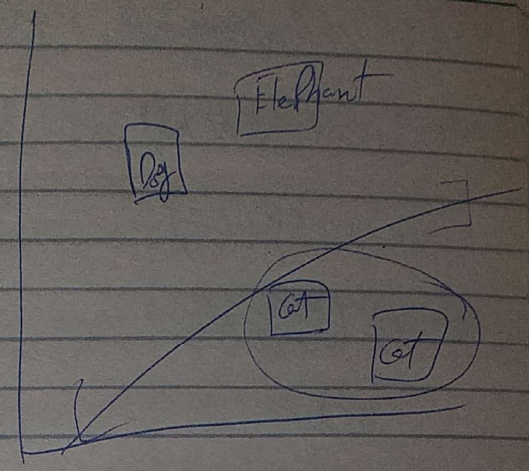
$XOR = \bar{A}B + A\bar{B}$

Q20) Interpretations of rows in W:

① Generic image of each class

② Each row is a equation of line which seperates this class from other classes.

③ weights +ve or -ve.



Q2) sigmoid + loss is

$$\sum_{y} -\log(P_y)$$

for multi-label classification.

Q2/(3) Code.

```
x = np.array([2, -1, 3])
w = np.array([[0.1, 0.2, 0.3], [-1, 1, 1], [2, 4, 0.5]])

O1 = w.dot(x)        # output of 3 neurons

def sig(x):
    return 1/(1+np.exp(-z))

# h1 = sig(O1)
```

w2 = np.array([[2, 0.5, 0.4]])

o2 = w2.dot(h1)

h2 = sig(o2)

[illegible]

t1 = y * np.log(h2)

t2 = (1-y) * np.log(1-h2)

Loss = t1+t2

Q2/ Gradient code.

Assignment #3:-

MNIST:

No Batch normalization

Optimizer:

Not gradient decent

use Adam

dropout

Thursday

from scratch.