Amit Nikam - Midterm

Q1. The task T is to predict weather, performance measure P is the probability of predicting the weather correctly. The learning algorithm can guess the weather and learn from its mistakes.

Q2. The correct options are: ~~A & D.~~ A, C & D.

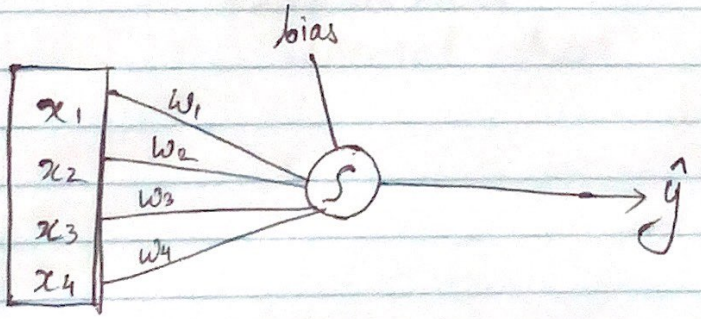
Q3. The correct options are: B.,

Q4. The correct answer: D

Q5. Answer: C

Q6. Answer: C

Q7.



inputs — x_1, x_2, x_3, x_4

output — $\hat{y}$

S = sigmoid activation

weights - w_1, w_2, w_3, w_4

$z = \text{bias} + \sum \text{weights} * \text{input}$

$\text{sigmoid} = \frac{1}{1+e^{-z}}$

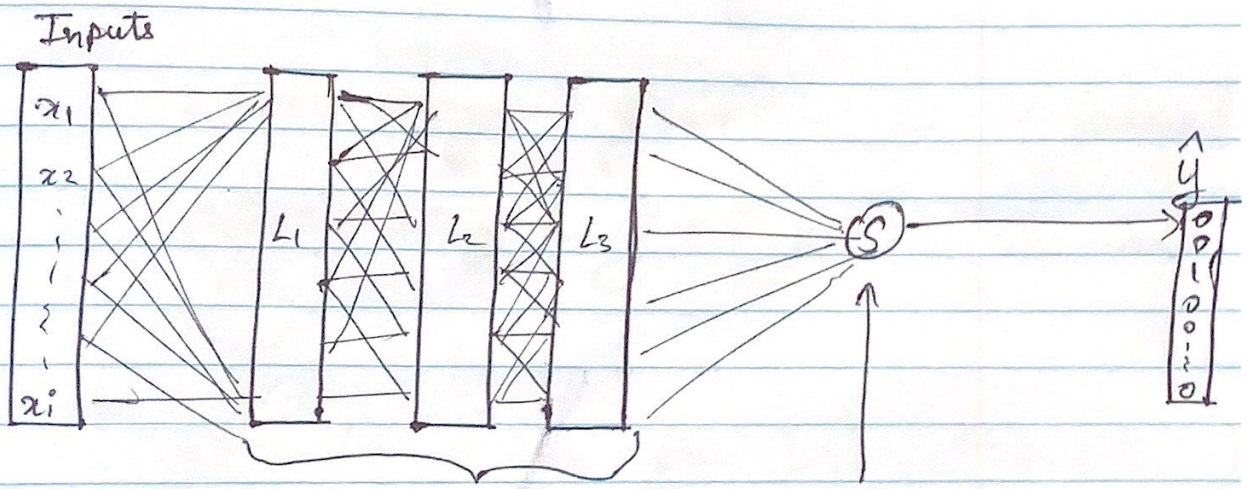
8. $z = 1 + 1\times1 + 2\times1 + -1\times1 + 3\times1$

$= 1 + 1 + 2 + -1 + 3$

$= 6$

$\text{Sigmoid} = \frac{1}{1+e^{-6}} = 0.997527$

9.



Fully connected neural network

Softmax (because multi-class classification)

Q10. The chain rule provides a technique for finding the derivative of composite functions in neural network. This is the foundation of backpropagation.

For, $\partial J(\theta) / \partial \theta_{11}^{(1)}$:

$$\partial J(\theta) / \partial \theta_{Z_1}^{(3)} = g'(z_1^{(3)}) \quad [\text{... derivative}]$$

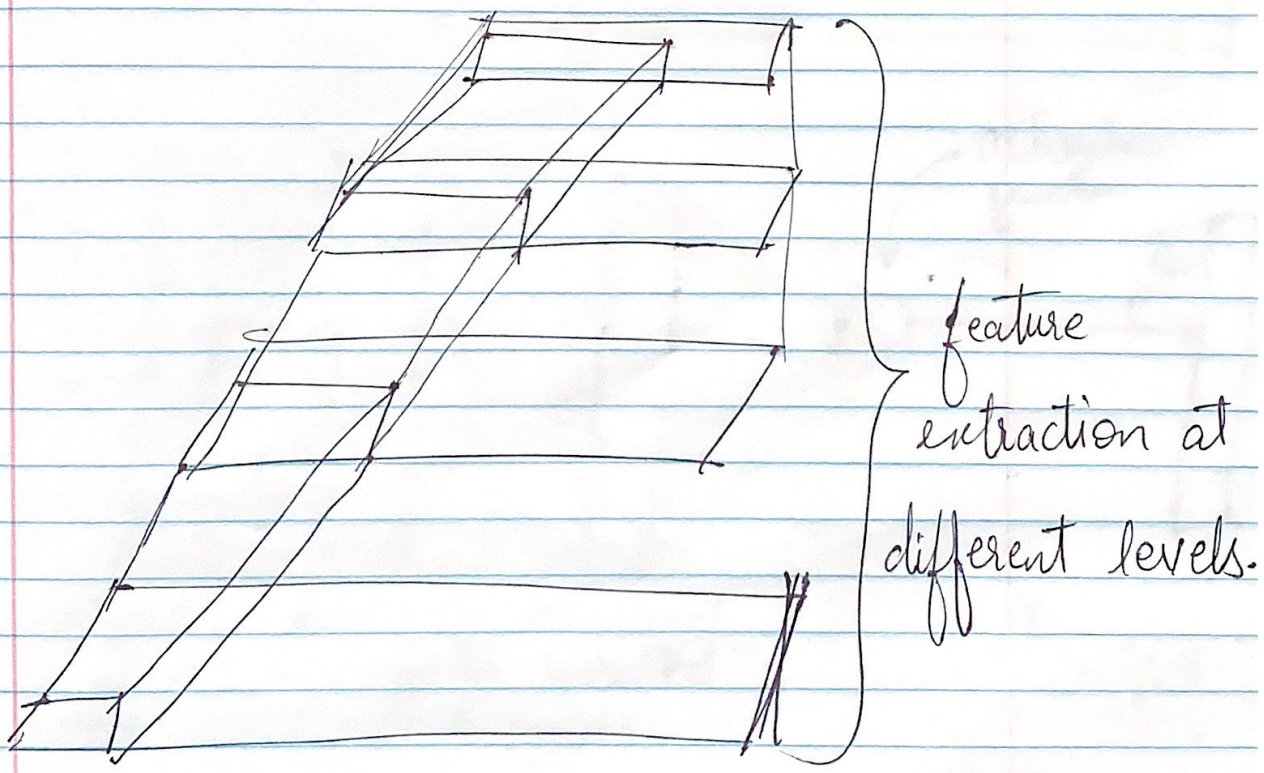
$$\partial J(\theta) / \partial \theta_{Z_2}^{(3)} = g'(z_2^{(3)}) \quad [\text{... derivative}]$$

So,

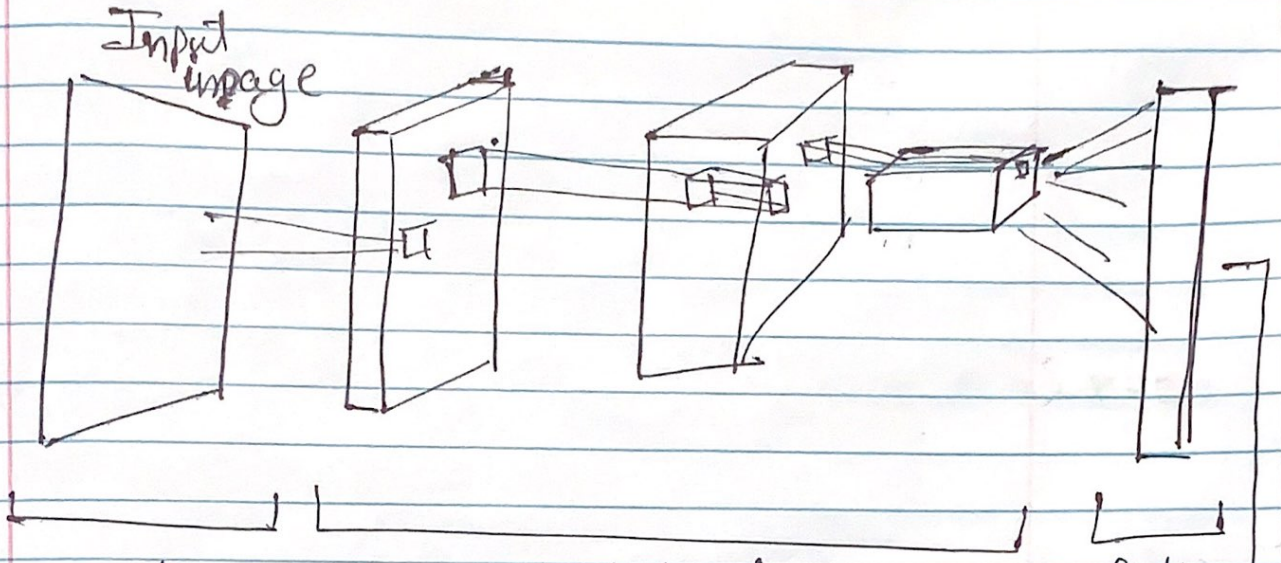
$$\partial J(\theta) / \partial \theta Z_1^{(2)} = g'(z_1^{(2)}) \cdot \left(g'(z_1^{(3)}) + g'(z_2^{(3)})\right)$$

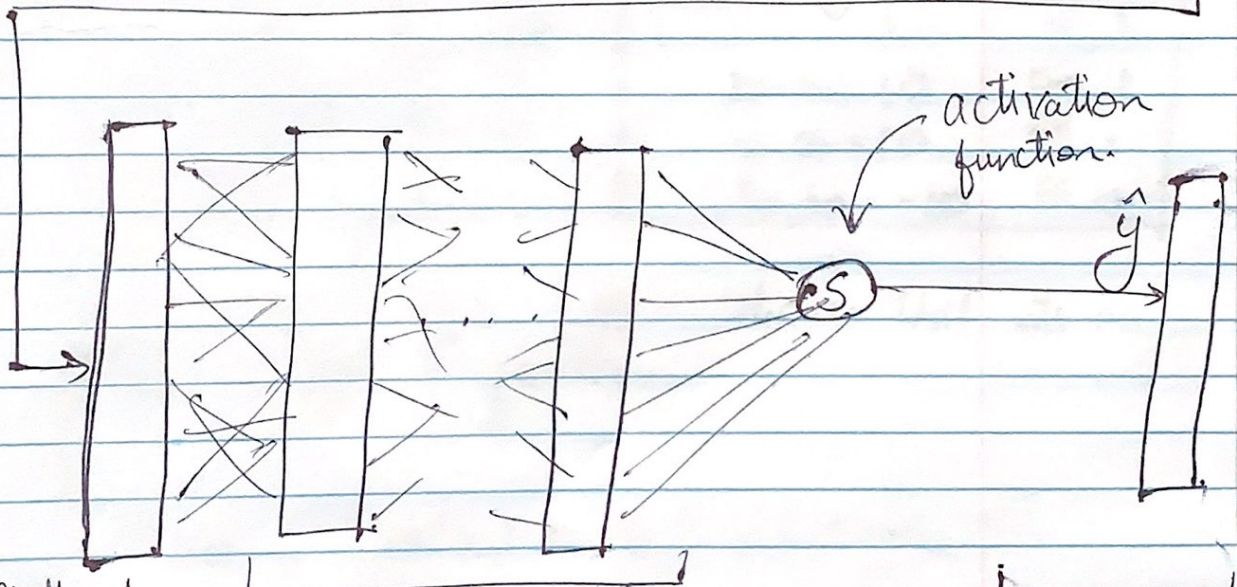
$$\partial J(\theta) / \partial \theta_{11}^{(1)} = g'(z_1^{(2)}) \cdot g'(z_1^{(3)}) + g'(z_1^{(2)}) \cdot g'(z_2^{(3)})$$

Q11. Convolutions make use of ~~fute~~ filters of certain kernal sizes which is a useful feature in neural networks for images as we are ~~Ey~~ trying to extract feature which are localized over a neighborhood. The filters of certain kernal sizes convolve i.e. slide in windows over the image extracting / looking for features. This window sliding of filter is defined using stride & decay. We usually maxpool images and repeat this process as we intend to extract features at different levels i.e. like pyramid.



12.





Classification

13. AND

$h\theta(x)$ is given as,

$$h\theta(x) = +1 \times 30 + x_1 \times -20 + x_2 \times -20$$

bias is constant $= 30$

Thus for x_1 & x_2, we can get following outcomes,

x_1	x_2	$\hat{y} = h\theta(x)$	
0	0	30	$\cong 1$
0	1	$30-20=10$	$\cong 1$
1	0	$30-20=10$	$\cong 1$
1	1	$30-40=-10$	$\cong 0$

As our logic table shows that its an AND logical function.

Thus, $h\theta(x) = (x_1)$ AND (x_2).