No.

Date

Deep Learning - Homework 2
Part 2

Q1 Channel 1:

2	3	4	5	6
14	15	16	1	1
6	7	8	0	0
8	10	12	1	2
22	24	26	3	3

Channel 2:

22	23	24	12	14
14	15	16	0	0
6	7	8	0	0
8	10	12	1	2
22	24	26	5	5

For setting 1, we have a $\frac{(5+0-3)}{1}+1 \Rightarrow 3\times 3$ output

For setting 2, we have $\frac{(5+0-3)}{2}+1 \Rightarrow 2\times 2$ output

Setting 1 : Channel 1:

22	25	28
28	32	36
36	41	46

Channel 2:

14	15	16
6	7	8
8	10	12

Setting 2 : Channel 1:

22	28
36	46

Channel 2:

14	16
8	12

Q2 Given $x = [-0.5,\ 0.3,\ 0.2]^T$

applying softmax to these values we get

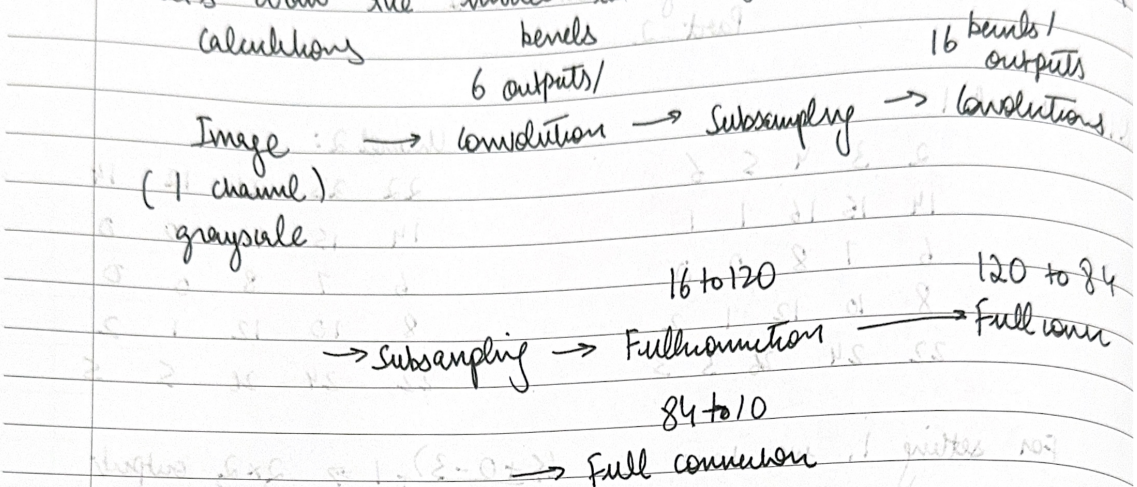
$$f(x)_i = \frac{e^{y_i}}{\sum_{k=1}^{K} e^{y_k}}$$

$\Rightarrow [0.190,\ 0.424,\ 0.3843]$

Now, $L = -\log(0.424)$

$\Rightarrow -0.85$

Q3 Let's draw the model in the format below for easier calculations



Now, we ignore subsampling layers as they have no trainable parameters so we left with

Conv layer 1 parameters ⇒ $(5 \times 5 \times 1 + 1) \times 6 = 156$
Conv layer 2 parameters ⇒ $(5 \times 5 \times 6 + 1) \times 16 = 2416$ (approx.)
(C5) fully-connected layer 1 ⇒ $(5 \times 5 \times 16 + 1) \times 120 = 48120$
fully-connected layer 2 ⇒ $(120 \times 84) + 84 = 10164$
" " layer 3 ⇒ $(84 \times 10) + 10 \Rightarrow 850$

In addition to these, we have 2 for (6×2) trainable parameters in the S2 or first pooling layer, where as the second pooling layer has 32 training parameters. Adding everything up we get:

61,750 training parameters approximately.