學號：B05602052 系級： 電機四 姓名：舒泓諭

1.請說明這次使用的 **model 架構，包含各層維度及連接方式**。

Training Epochs: 50

Batch Size: 256

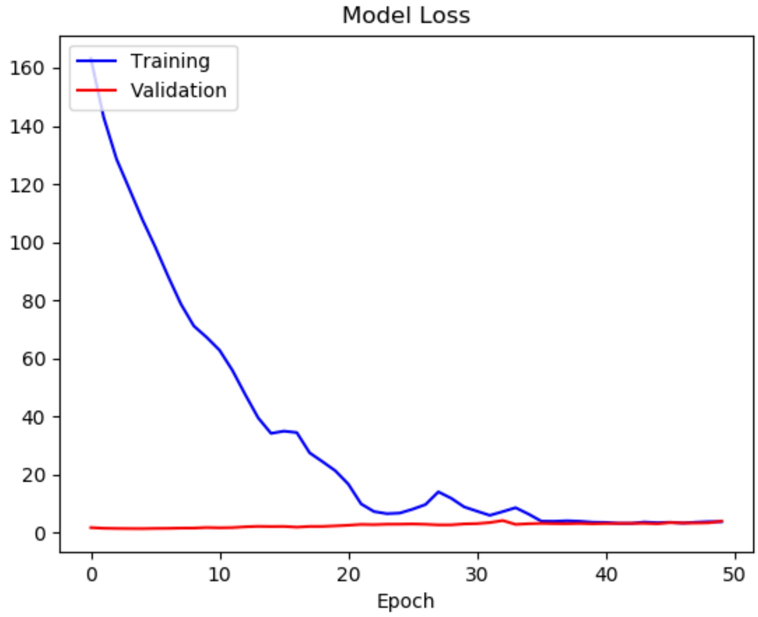
Optimizer: Adam

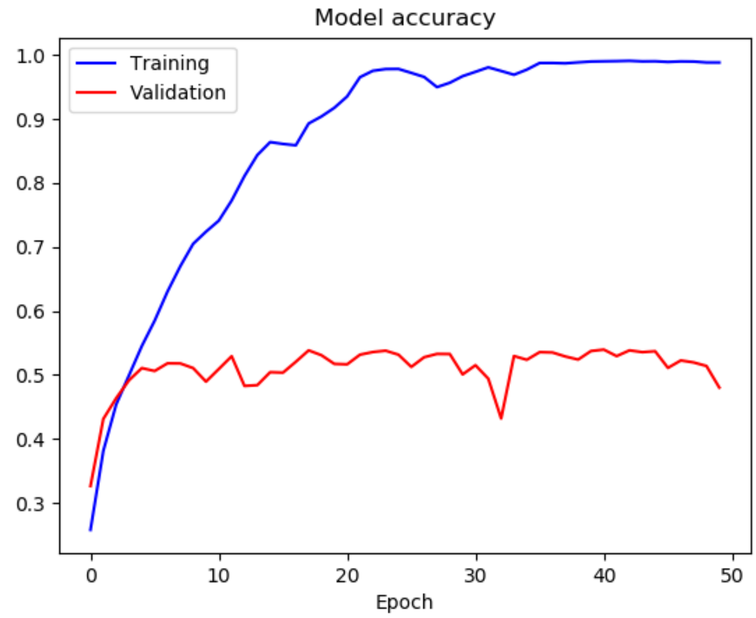
Validation Data: 20% training data

以下用表格的方式來說明 NN 的連結方式：

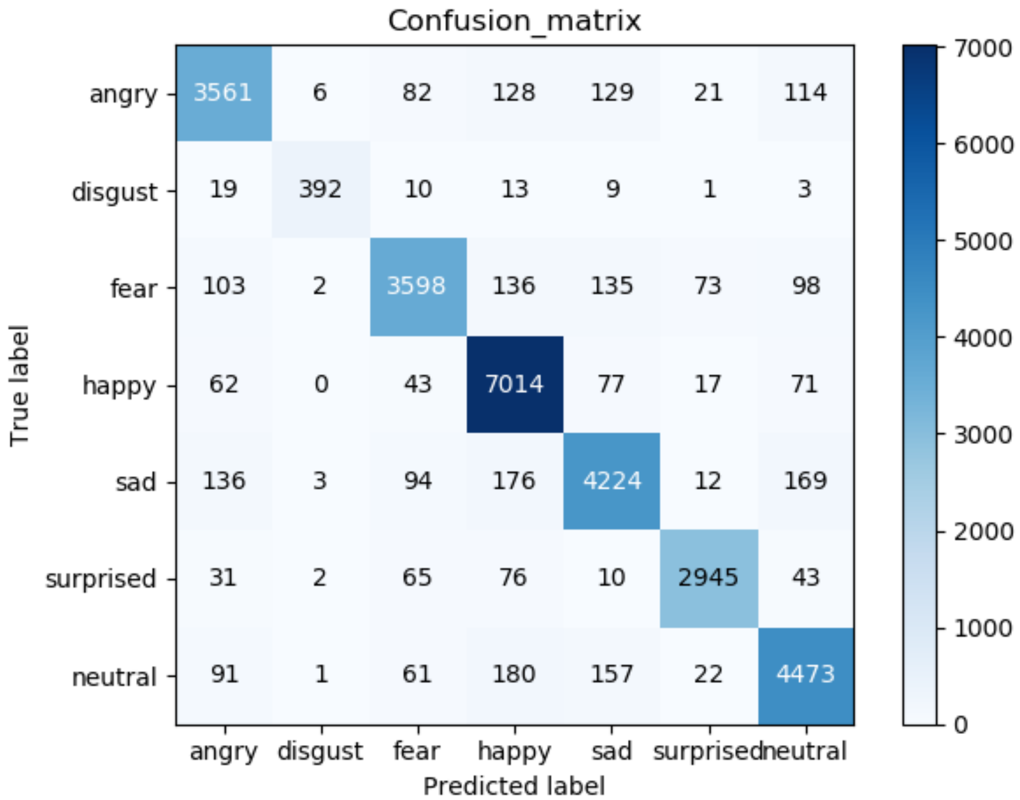
Layer	**Input_chnnel**	**Output_channel**	**Shape**
Conv2D(kernel size =4, stride=2, padding=1)	1	64	48*48
BatchNorm2D	64	64	24*24
LeakyRelu(0.2)	64	64	24*24
Conv2D(kernel size =3, stride=1, padding=1)	64	64	24*24
BatchNorm2D	64	64	24*24
LeakyRelu(0.2)	64	64	24*24
MaxPool2d(2*2)	64	64	12*12
Conv2D(kernel size =3, stride=1, padding=1)	64	128	12*12
BatchNorm2D	128	128	12*12
LeakyRelu(0.2)	128	128	12*12
MaxPool2d(2*2)	128	128	6*6
Conv2D(kernel size =3, stride=1, padding=1)	128	256	6*6
BatchNorm2D	256	256	6*6
LeakyRelu(0.2)	256	256	6*6
MaxPool2d(2*2)	256	256	3*3
Flatten	256	2304	1*1
Fully connected (Dropout(0.5)+ LeakyRelu(0.2))	2304	1024	1*1
Fully connected (Dropout(0.5)+ LeakyRelu(0.2))	1024	512	1*1
Fully connected (Dropout(0.5)+ LeakyRelu(0.2))	512	256	1*1
Fully connected (Dropout(0.5)+ LeakyRelu(0.2))	256	7	1*1

2.請附上 model 的 training/validation history (loss and accuracy)。





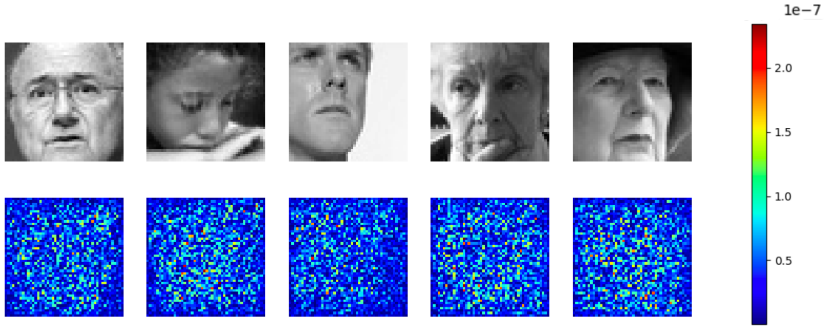
3.畫出 confusion matrix 分析哪些類別的圖片容易使 model 搞混，並簡單說明。

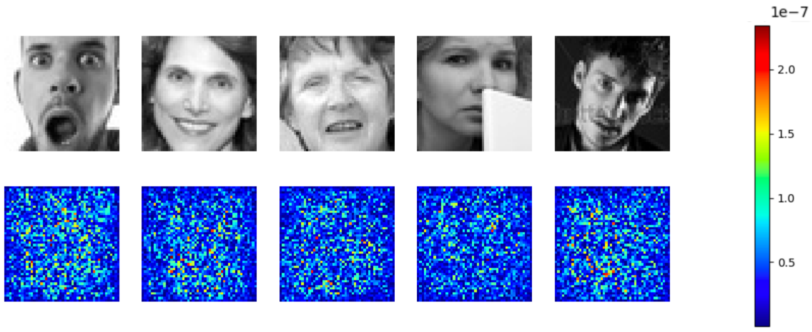


可以看出最容易辨別出來的類別是 happy 的類別。此外，比較容易混肴的類別分析，用下面表格來說明**(辨別錯次數>150)**，從以上表格可以發現，對於 CNN 這個架構而言，Neutral 和 Sad 這兩個 label 是比較難區分出來的，也是筆記容易混淆的類別。

True label V.S. Predicted label	**Error number**
Happy V.S. Sad	176
Sad V.S. Neutral	169
Neutral V.S. Happy	180
Neutral V.S. Sad	157

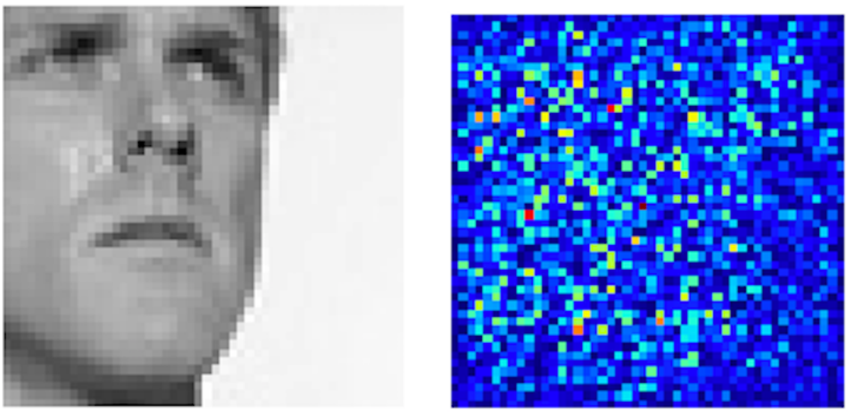
4.畫出 CNN model 的 saliency map，並簡單討論其現象。





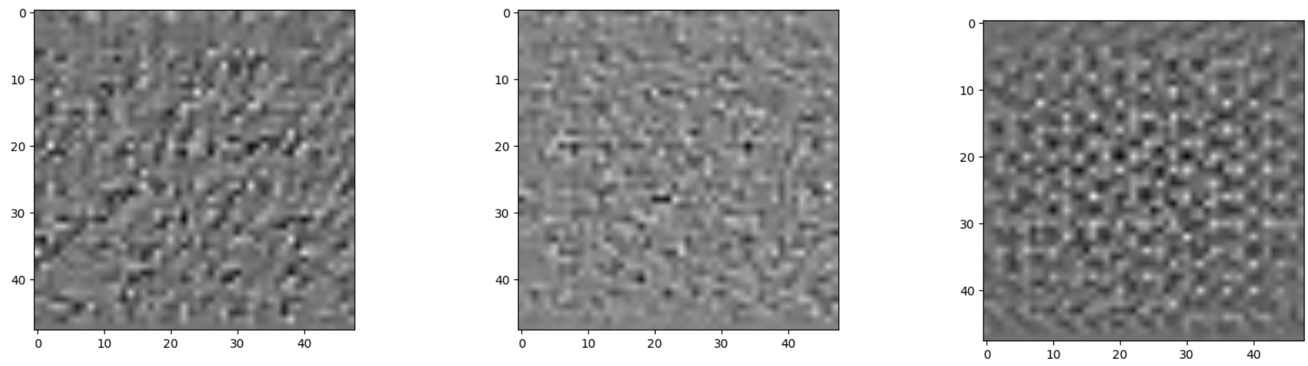
可以發現其實 CNN 所辨別的主要特徵，還是在**人的五官**以及**臉的輪廓，**也就是 saliency map 中藍色以外的部分。

而當 training data 的照片中如果，有**人臉之外的物件時**，我們可以發現沒有人臉的地方會有一大片的藍色，例如下圖，從此更可以看出其實 CNN 所辨別的主要特徵還是在於人臉的輪廓。

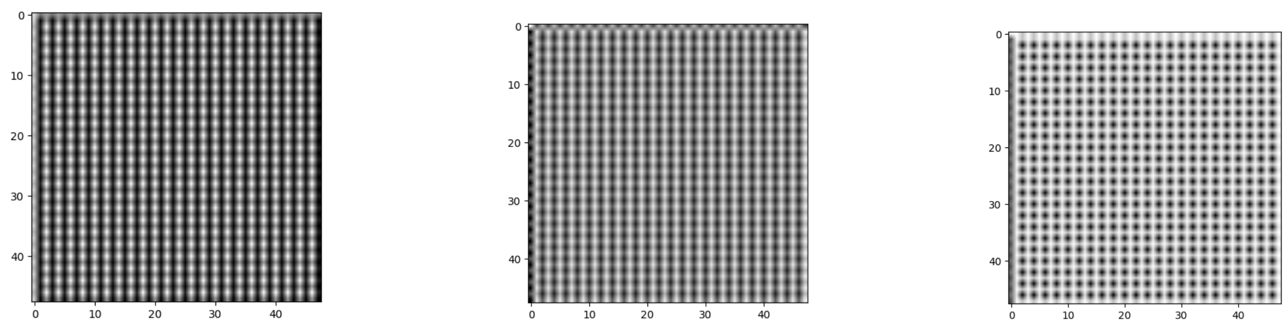


5.畫出最後一層的 filters 最容易被哪些 feature activate。

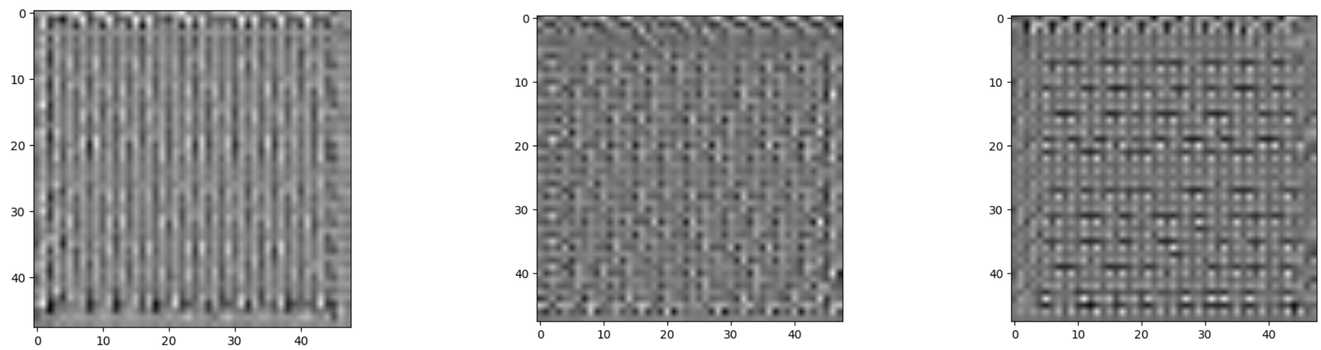
conv2d_3 作為最後一層的 filter，可以辨認出相較其他層比較複雜的 feature。



conv2d_3 (最後一層 filter)



conv2d_2

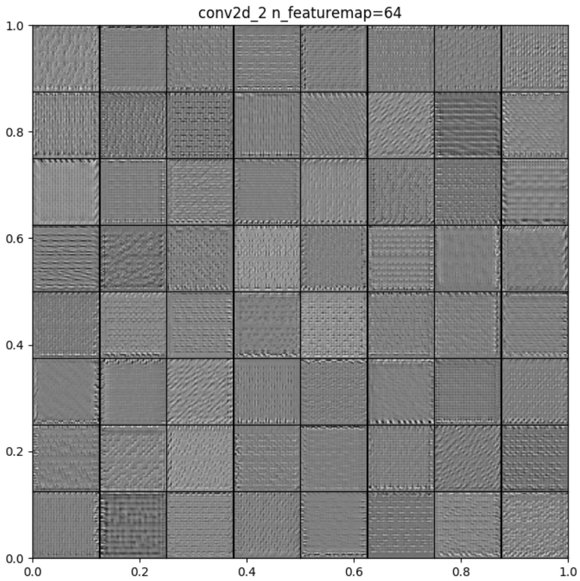


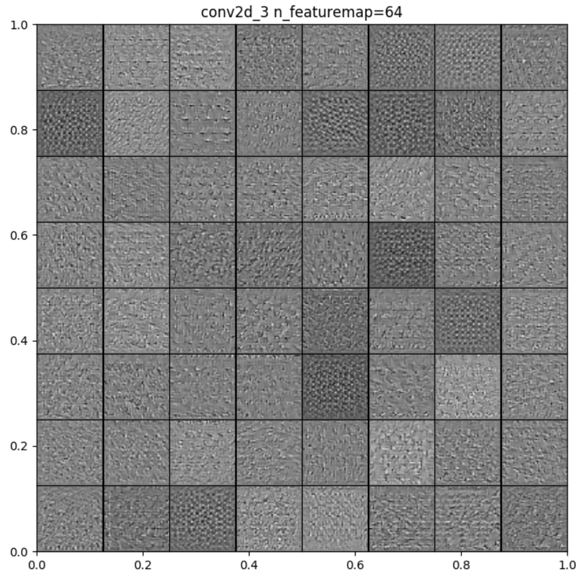
conv2d_1

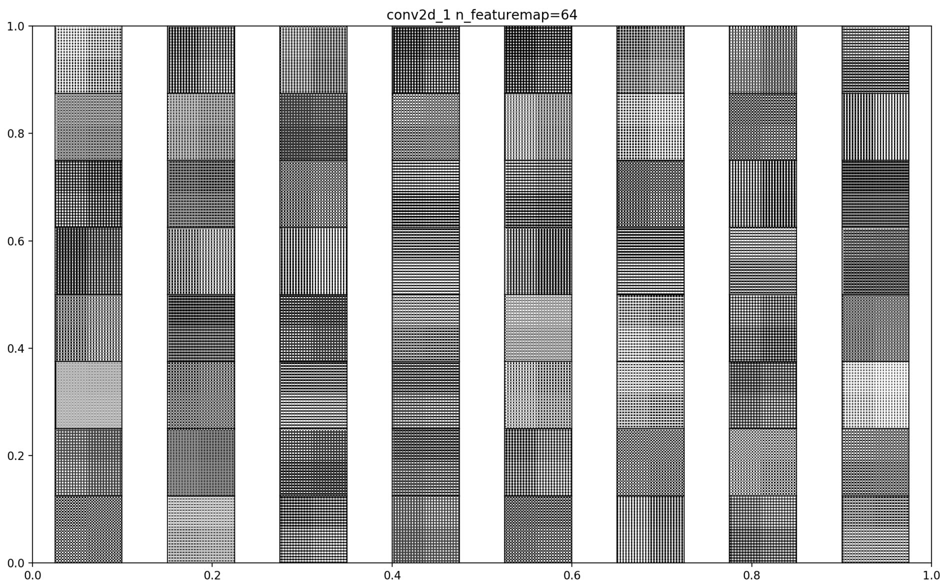
以下將分別比較 **conv2d_1,conv2d_2,conv2d_3 這三層 layer(三者順序由淺到深)**，而 **conv2d_3** 則是我 CNN 在 Fully connectd 之前的最後一層 layer。

conv2d_1 作為第一層的 layer 只能辨別出，區域的顏色強度，而到 **conv2d_2** 時則可以辨識出稍微複雜的 feature，相對於前面幾層的 layer，**conv2d_3** 作為最後一層的 filter 可以辨認出比相較起來較複雜的 feature。

此外，從以上三張圖都可以看出來，有些 filter 明明長的一樣，只是方向不太一樣而已。這也蠻符合當初我們對 CNN 運算的直觀，因爲如果 CNN 要能清楚的辨別出圖片，那麼出現在不同的角度的相同 feature，那麼就一定需要不同角度的 filter。



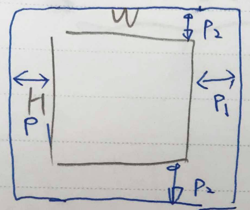




6.手寫做業

1.

1.

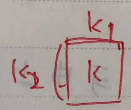


$$W' = W + 2P_1$$

$$H' = H + 2P_2$$

$$B_{out} = B$$

$$\therefore H_{out} = \left\lfloor \frac{H + 2P_2 - k_2}{S_2} + 1 \right\rfloor$$



$$W_{out} = \left\lfloor \frac{W + 2P_1 - k_1}{S_1} + 1 \right\rfloor$$

output_channel = input_channel

$$\therefore (B', W', H', channels')$$

$$= \left(B, \left\lfloor 1 + \frac{W + 2P_1 - k_1}{S_1} \right\rfloor, \left\lfloor 1 + \frac{H + 2P_2 - k_2}{S_2} \right\rfloor, output\_channels\right) \#$$

2.

3.

$$\hat{y}_t = \text{softmax}(z_t) = \frac{e^{z_t}}{\sum_{i=1} e^{z_i}}$$

$$\frac{\partial L_t}{\partial z_t} = \frac{\partial L_t}{\partial \hat{y}_t} \frac{\partial \hat{y}_t}{\partial z_t}$$

consider ①. $\frac{\partial \hat{y}_i}{\partial z_j} = \frac{\partial \frac{e^{z_t}}{\sum e^{z_i}}}{\partial z_t} = \frac{e^{z_t}\sum e^{z_i} - e^{z_t} \cdot e^{z_t}}{(\sum e^{z_i})^2}$

as $\boxed{i = j = t}$

$$= \frac{e^{z_t}}{\sum e^{z_i}} \frac{\sum e^{z_i} - e^{z_t}}{\sum e^{z_i}}$$

$$= \hat{y}_t (1 - \hat{y}_t)$$

consider ②. $\frac{\partial \hat{y}_i}{\partial z_j} = \frac{\partial \frac{e^{z_i}}{\sum e^{z_i}}}{\partial z_j} = \frac{0 - e^{z_i} e^{z_j}}{(\sum e^{z_i})^2} = -\frac{e^{z_i}}{\sum e^{z_i}} \frac{e^{z_j}}{\sum e^{z_i}}$

as $i \neq j$.

$$= -\hat{y}_i \cdot \hat{y}_j.$$

$$\frac{\partial L}{\partial z_t} = -\sum_{i=1} \frac{\partial y_i \log \hat{y}_i}{\partial z_t} = -\frac{\partial y_t \log \hat{y}_t}{\partial z_t} - \sum_{i \neq t} \frac{\partial y_i \log \hat{y}_i}{\partial z_t}$$

$$= -y_t \frac{\partial \log \hat{y}_t}{\partial z_t} - \sum_{i \neq t} y_i \frac{\log \hat{y}_i}{\partial z_t}$$

$$= -y_t \frac{1}{\cancel{\hat{y}_t}} \cancel{(\hat{y}_t)}(1 - \hat{y}_t) - \sum_{i \neq t} \frac{y_i}{\hat{y}_i} - \hat{y}_i \hat{y}_t.$$

$$\boxed{\sum_{i=1} y_i = 1}$$

total probability.

$$= -y_t + \underline{y_t \hat{y}_t} + \hat{y}_t \underline{\sum_{i \neq t} y_i}.$$

$$= -y_t + \hat{y}_t \underbrace{\sum_{i=1} y_i}_{=1}$$

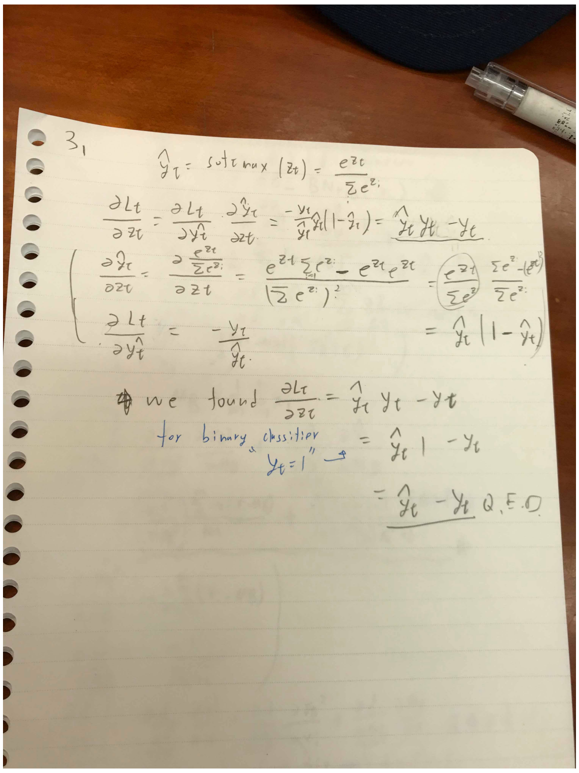
$$= \underline{\hat{y}_t - y_t} \quad \text{Q.E.D}$$

$$\boxed{\frac{\partial \bar{l}}{\partial \gamma}} = \sum_{i=1}^{m} \frac{\partial l}{\partial y_i} \frac{\partial y_i}{\partial \gamma} = \underline{\sum_{i=1}^{m} \frac{\partial l}{\partial y_i} \hat{x}_i} \quad \#$$

$$\boxed{\frac{\partial l}{\partial \beta}} = \sum_{i=1}^{m} \frac{\partial l}{\partial y_i} \frac{\partial y_i}{\partial \beta} = \underline{\sum_{i=1}^{m} \frac{\partial l}{\partial y_i}} \quad \#$$

1.

3.for binary classifier



3,

$$\hat{y}_t = \text{softmax}(z_t) = \frac{e^{z_t}}{\sum e^{z_i}}$$

$$\frac{\partial L_t}{\partial z_t} = \frac{\partial L_t}{\partial \hat{y}_t} \frac{\partial \hat{y}_t}{\partial z_t} = \frac{-y_t}{\hat{y}_t} \hat{y}_t (1-\hat{y}_t) = \underline{\hat{y}_t y_t - y_t}$$

$$\frac{\partial \hat{y}_t}{\partial z_t} = \frac{\partial \frac{e^{z_t}}{\sum e^{z_i}}}{\partial z_t} = \frac{e^{z_t} \sum_{i=1} e^{z_i} - e^{z_t} e^{z_t}}{(\sum e^{z_i})^2} = \frac{e^{z_t}}{\sum e^{z_i}} \frac{\sum e^{z_i} - (e^{z_t})}{\sum e^{z_i}}$$

$$= \hat{y}_t (1 - \hat{y}_t)$$

$$\frac{\partial L_t}{\partial \hat{y}_t} = \frac{-y_t}{\hat{y}_t}$$

we found $\frac{\partial L_t}{\partial z_t} = \hat{y}_t y_t - y_t$

for binary classifier "$y_t = 1$" → $= \hat{y}_t 1 - y_t$

$$= \underline{\hat{y}_t - y_t} \quad \text{Q.E.D}$$

For general case:

3.

$$\hat{y}_t = \text{softmax}(z_t) = \frac{e^{z_t}}{\sum_{i=1} e^{z_i}}$$

$$\frac{\partial L_t}{\partial z_t} = \frac{\partial L_t}{\partial \hat{y}_t}\frac{\partial \hat{y}_t}{\partial z_t}$$

consider ① $\frac{\partial \hat{y}_i}{\partial z_j}$ as $\boxed{i=j=t}$

$$= \frac{\partial \frac{e^{z_t}}{\sum e^{z_i}}}{\partial z_t} = \frac{e^{z_t}\sum e^{z_i} - e^{z_t}\cdot e^{z_t}}{(\sum e^{z_i})^2}$$

$$= \frac{e^{z_t}}{\sum e^{z_i}} \cdot \frac{\sum e^{z_i} - e^{z_t}}{\sum e^{z_i}}$$

$$= \hat{y}_t(1-\hat{y}_t)$$

consider ② $\frac{\partial \hat{y}_i}{\partial z_j}$ as $i \neq j$

$$= \frac{\partial \frac{e^{z_i}}{\sum e^{z_i}}}{\partial z_j} = \frac{0 - e^{z_i}e^{z_j}}{(\sum e^{z_i})^2} = -\frac{e^{z_i}}{\sum e^{z_i}} \cdot \frac{e^{z_j}}{\sum e^{z_i}}$$

$$= -\hat{y}_i \cdot \hat{y}_j$$

$$\frac{\partial L}{\partial z_t} = -\sum_{i=1} \frac{\partial y_i \log \hat{y}_i}{\partial z_t} = -\frac{\partial y_t \log \hat{y}_t}{\partial z_t} - \sum_{i \neq t} \frac{\partial y_i \log \hat{y}_i}{\partial z_t}$$

$$= -y_t \frac{\partial \log \hat{y}_t}{\partial z_t} - \sum_{i \neq t} y_i \frac{\log \hat{y}_i}{\partial z_t}$$

$$= -y_t \frac{1}{\cancel{\hat{y}_t}} \cancel{\hat{y}_t}(1-\hat{y}_t) - \sum_{i \neq t} \frac{y_i}{\hat{y}_i} \cdot -\hat{y}_i \hat{y}_t$$

$$= -y_t + y_t \hat{y}_t + \hat{y}_t \sum_{i \neq t} y_i$$

$\boxed{\sum_{i=1} y_i = 1}$

Total probability.

$$= -y_t + \hat{y}_t \underbrace{\sum_{i=1} y_i}_{=1}$$

$$= \hat{y}_t - y_t \quad \text{Q.E.D}$$