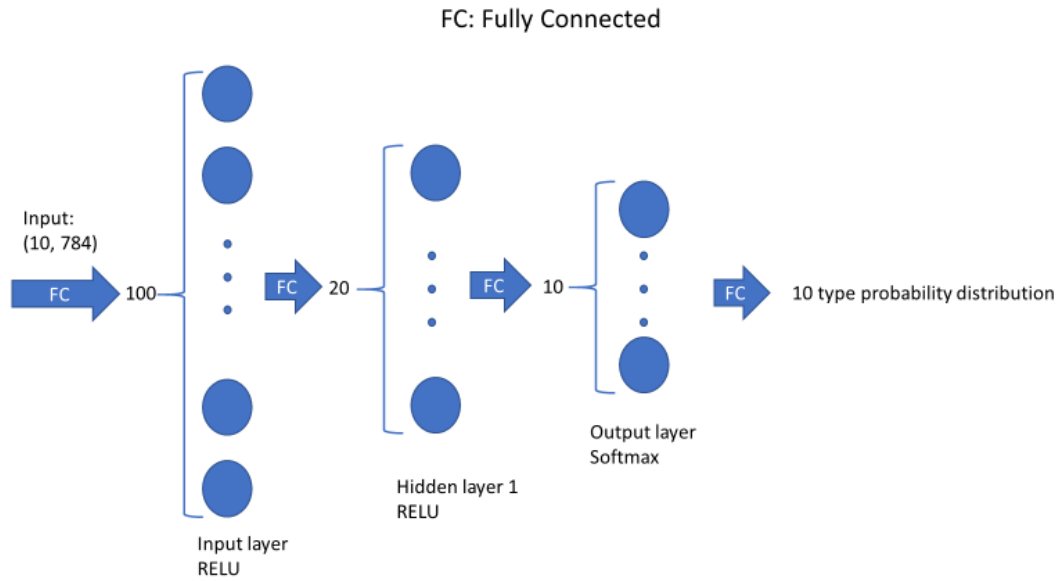


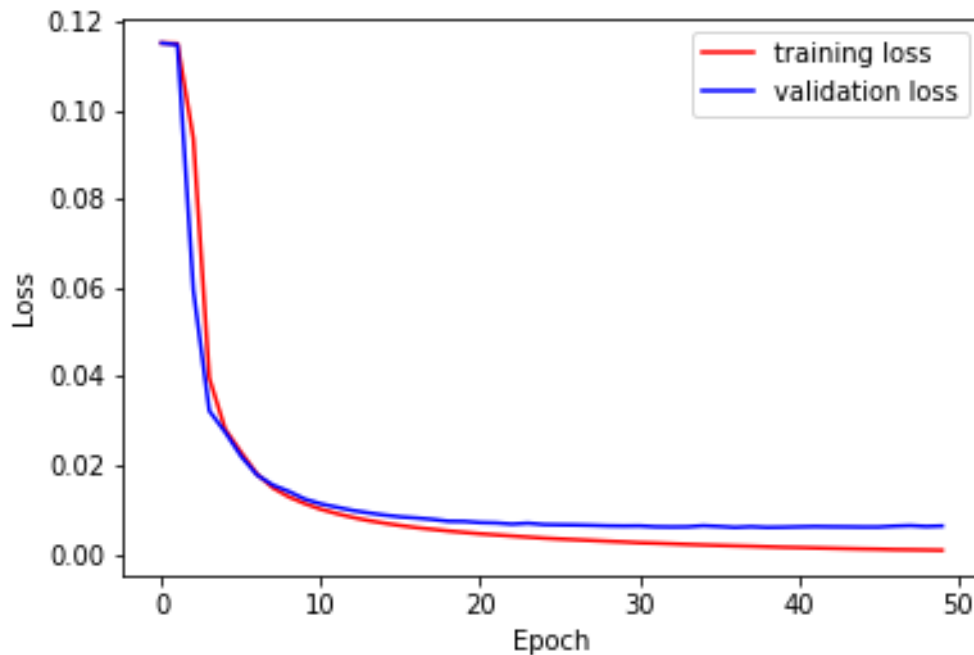
1. Show your model architecture and testing accuracy.



Select epoch 45 based on validation set
Epoch 45 on testing set: 97.32 %

2. How do you implement feed forward and backward propagation? A brief explanation is fine.
 - Feed forward: 是指透過 input matrix 與 weight matrix 相乘再加上 bias，並通過 activation function 得到 layer output，一組 weight matrix & bias & activation function 可視為一層，可以疊加多層，最後根據 softmax 計算出各個類別的機率分布。
 - backward propagation: 先用 ground truth one hot vector 與 feed forward 產生的各類別的機率分布兩者做 cross entropy 產生 loss，並用該 loss 對神經網路中的每一個 weight & bias 做偏微分，該偏微分結果就是對某個參數的更新方向，通常需要對此更新方向乘上 learning rate 再對參數做更新。
3. Plot training loss and validation loss. (loss vs. epochs figure)

下圖的 loss 值是每個 epoch 看過所有筆資料的平均



4. If we use a very deep NN with a large number of neurons, will the accuracy increase? Why or why not?

越深越寬的神經網路意味著參數量更多，model size 越大，通常會讓 model 在 training data 上學到更多更細微的東西，所以在 training dataset 上面的 accuracy 表現會更好，但是這麼做可能會造成 overfitting，也就是在 testing dataset 上的 accuracy 表現會變差，因為 model 過於緊密或精確地匹配 training dataset，以致於無法良好地調適其他資料或預測未來的觀察結果。

5. Why do we need to validate our model?

因為在眾多次更新參數中，我們需要挑選哪一次更新後的參數模型是我們想要拿來做測試的，所以需要 validation set 來做挑選，我的作法是根據每一個 epoch 的 model 來做 validation，看哪個 epoch 在 validation set 上的 accuracy 表現最好，最後就挑選該 epoch 的 model 來做 testing。