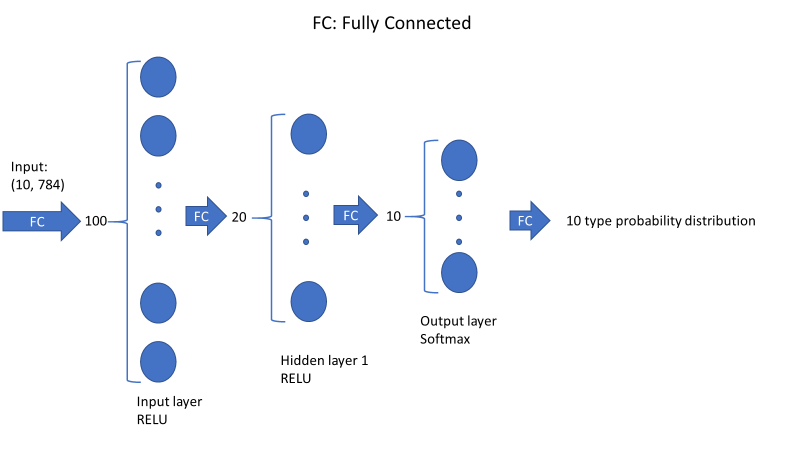
1. Show your model architecture and testing accuracy.



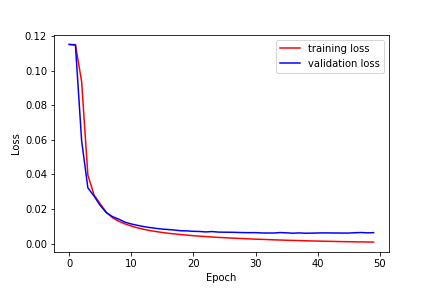


1. How do you implement feed forward and backward propagation? A brief explanation is fine.

* 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。

1. 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



1. 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。

1. 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，所以需要用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。