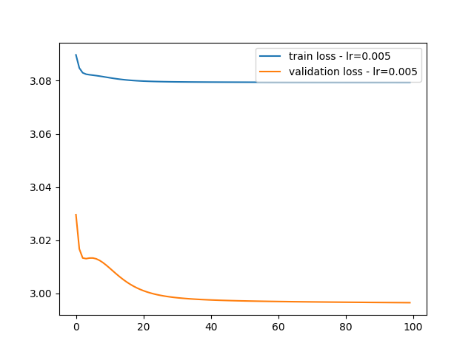
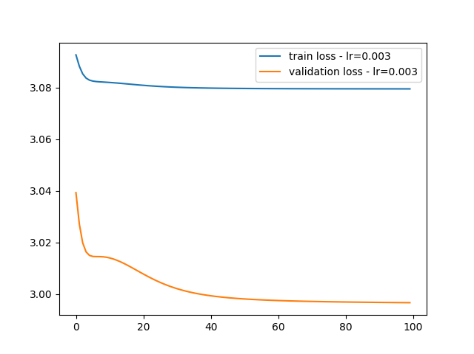
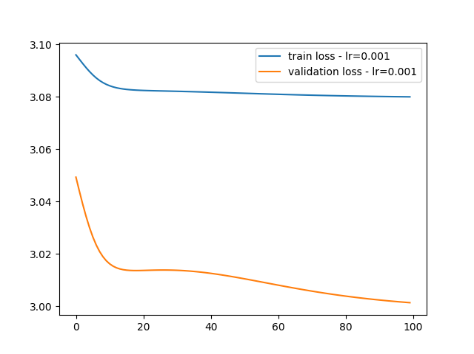
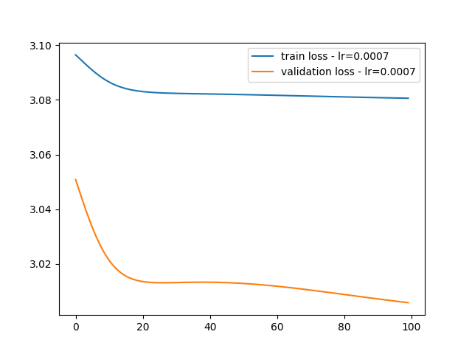
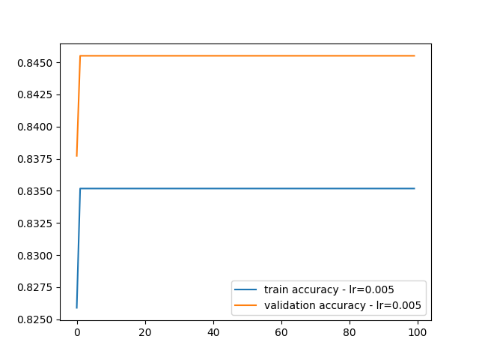
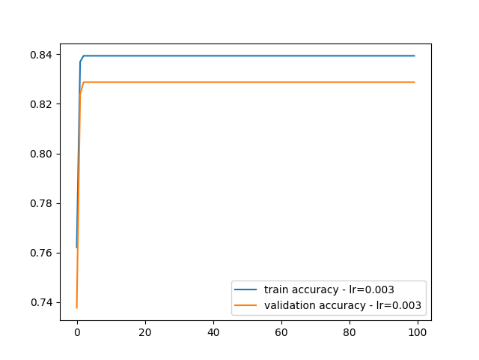
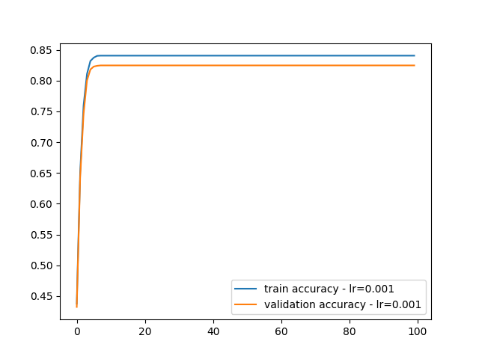
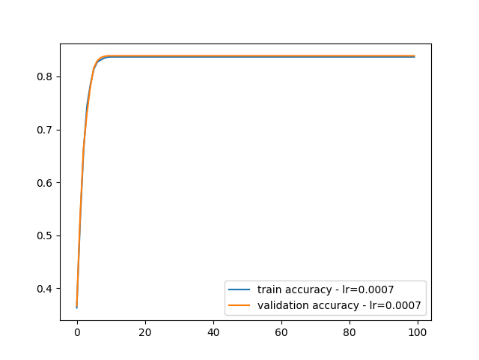
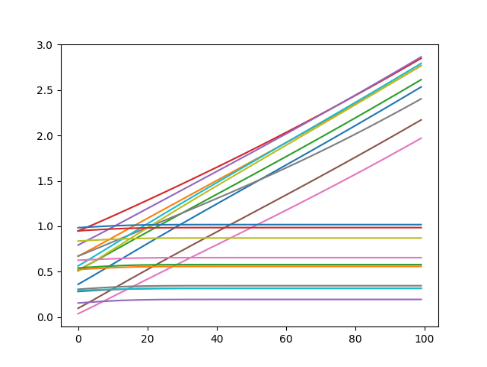
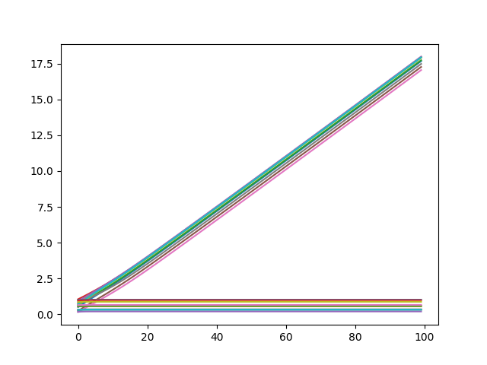
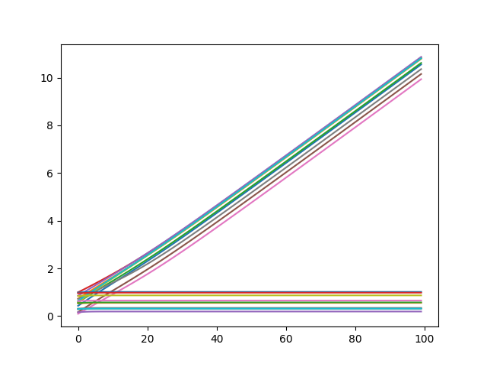
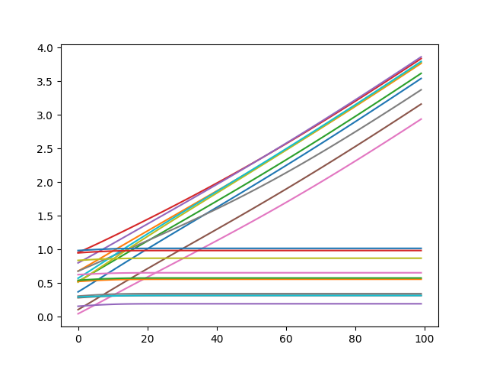
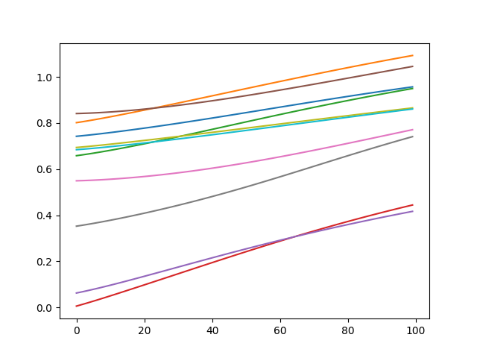
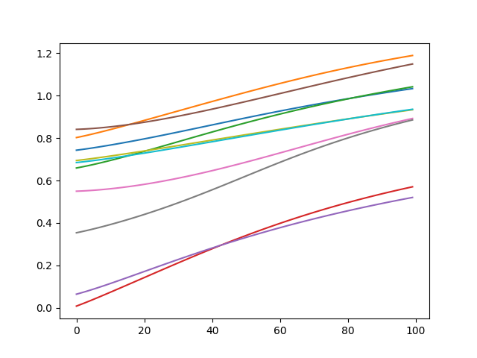
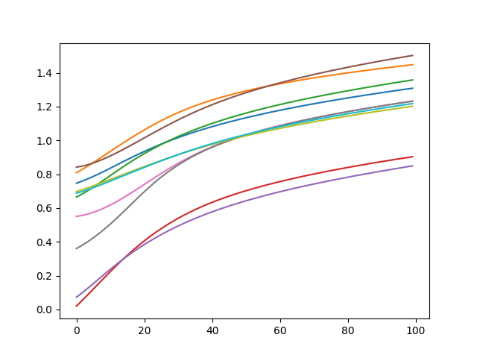
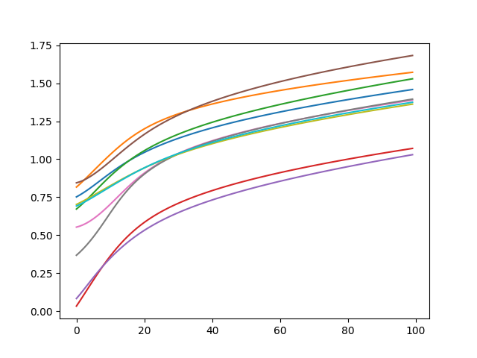
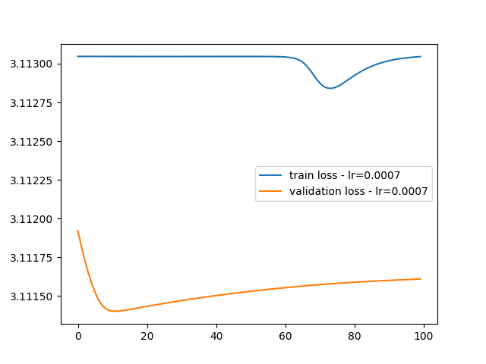
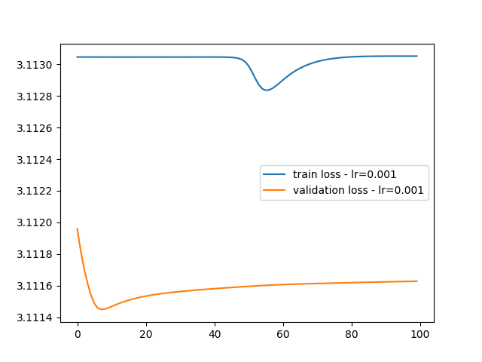
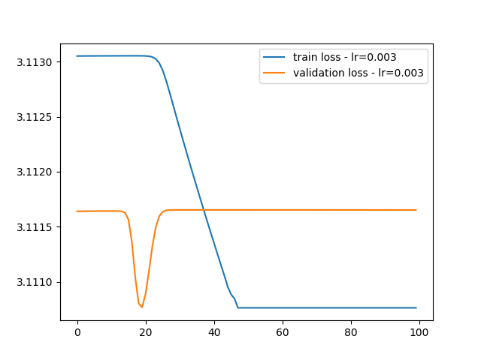
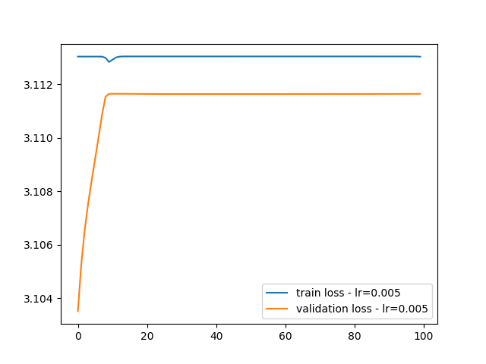
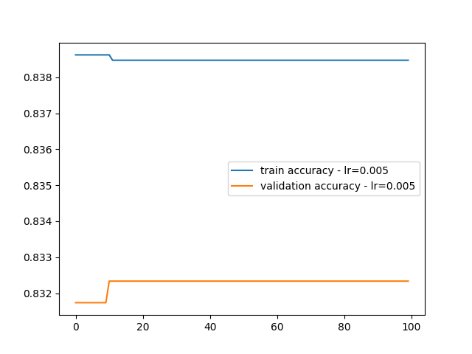
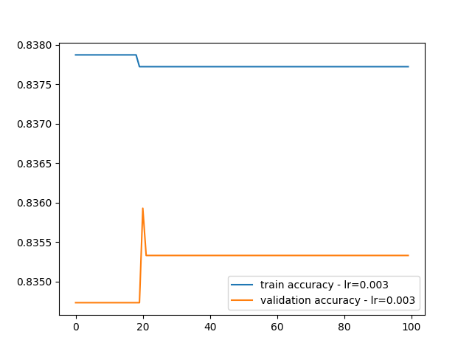
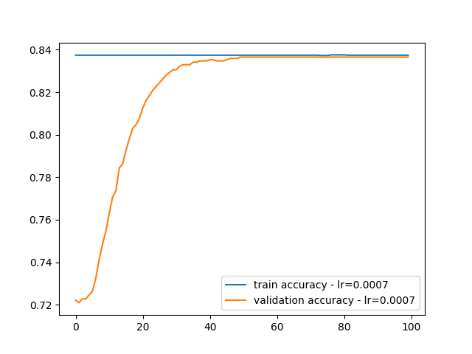
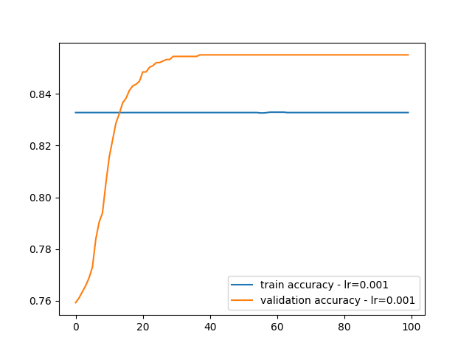
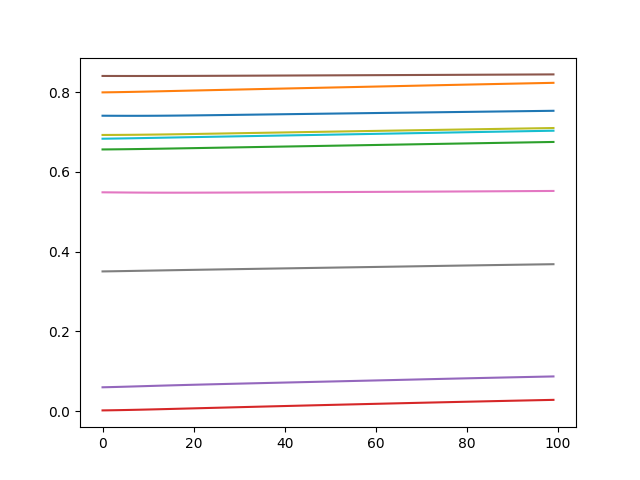
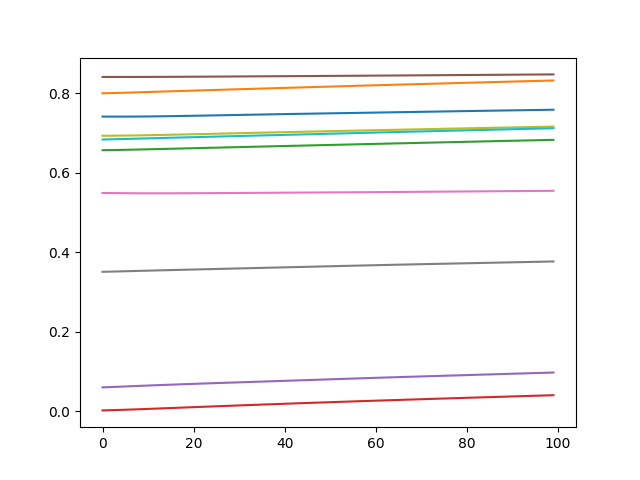
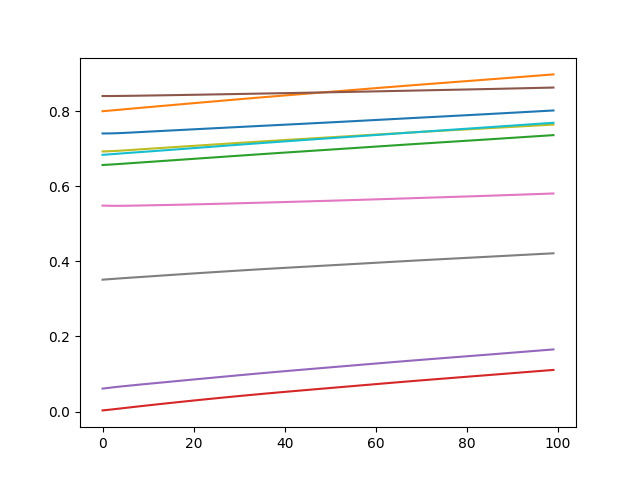
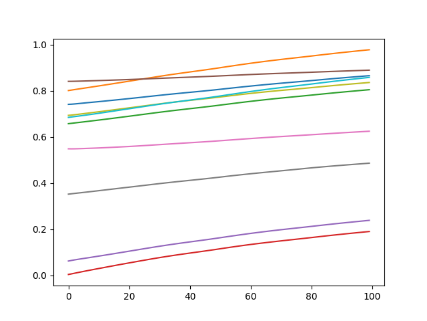
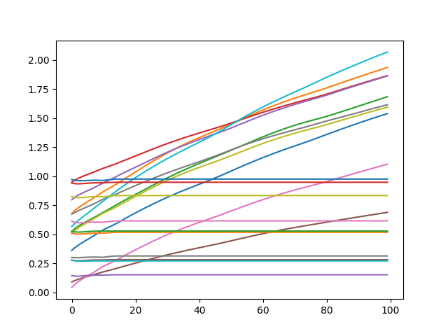
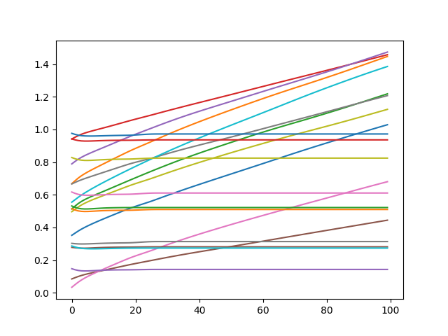
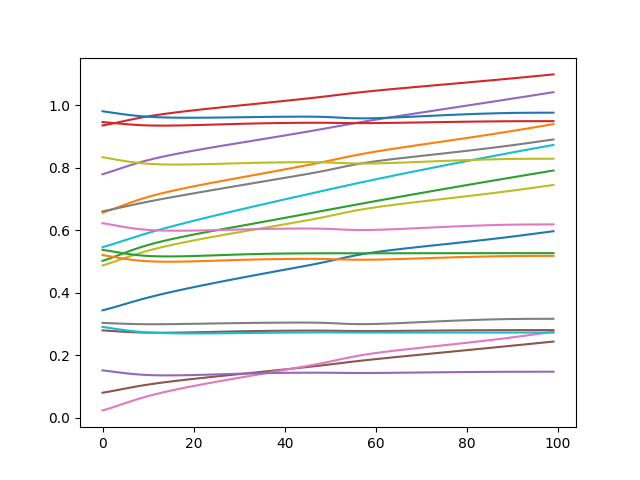
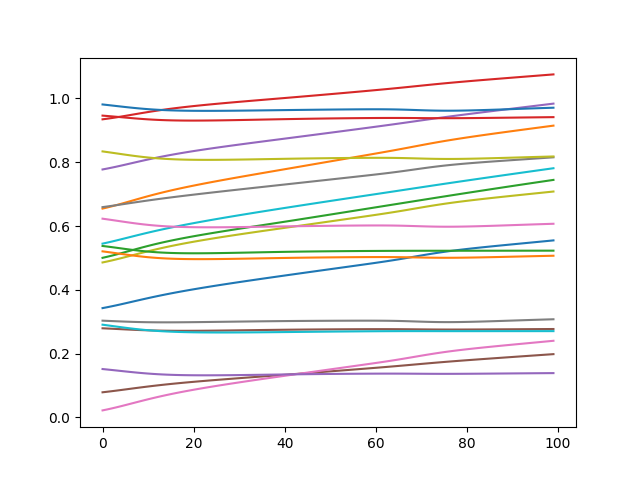
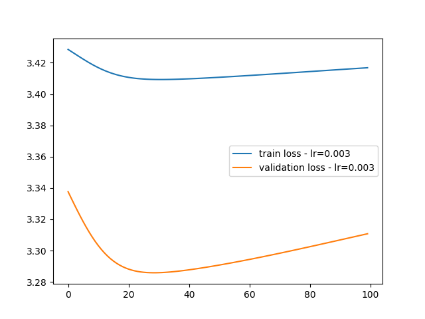
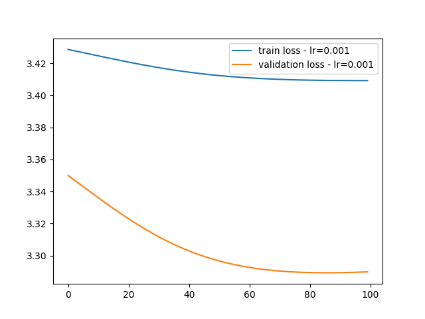
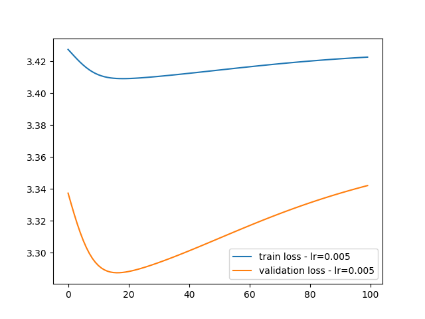
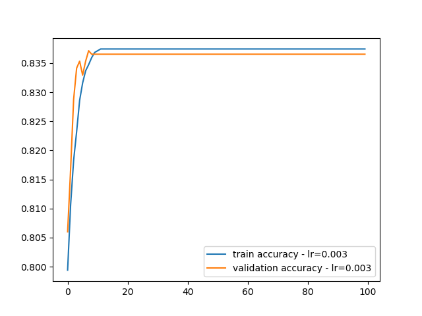
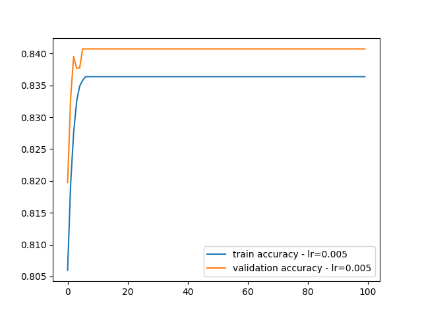
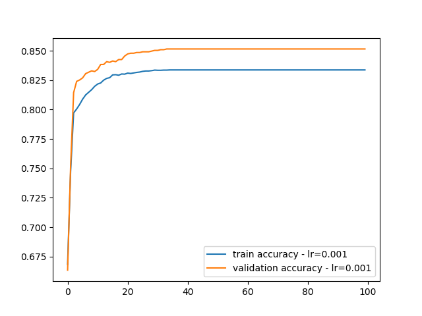
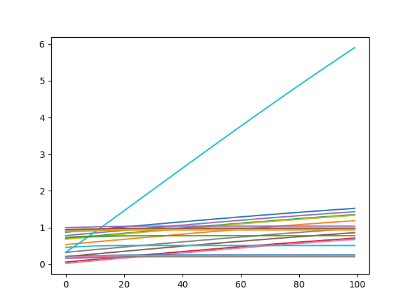
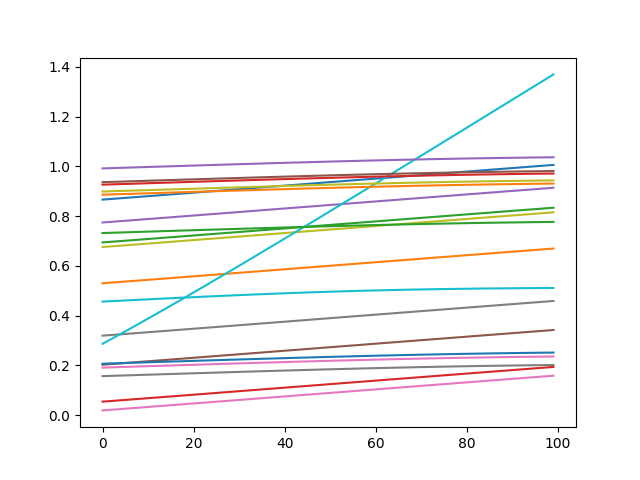
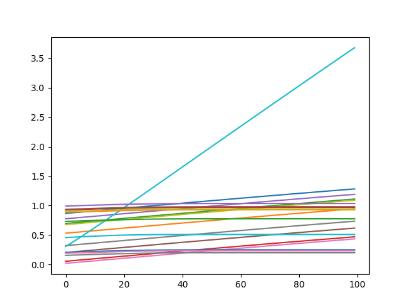
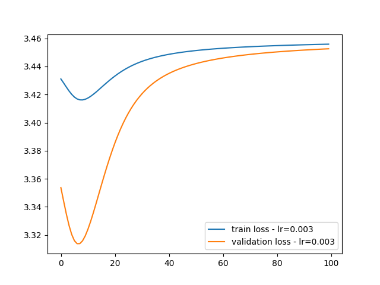
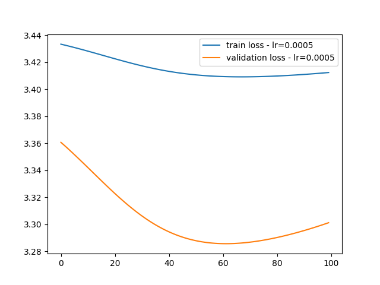
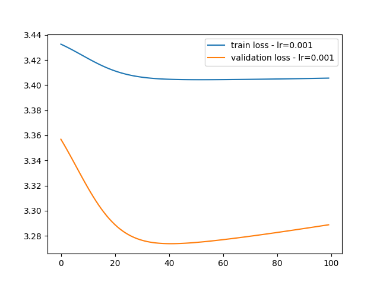
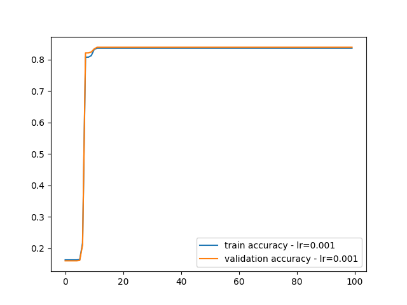
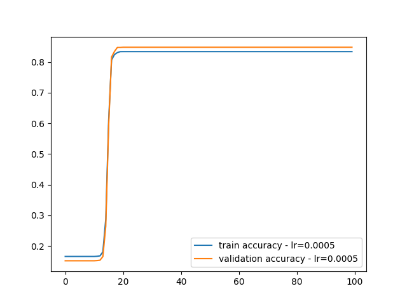
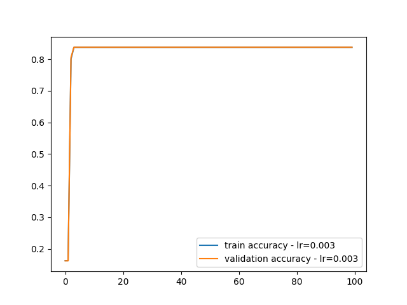
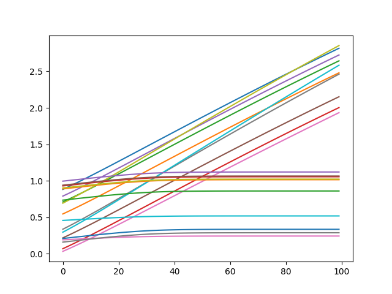
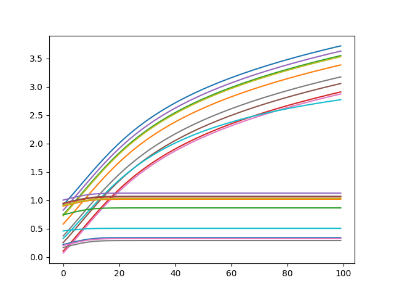
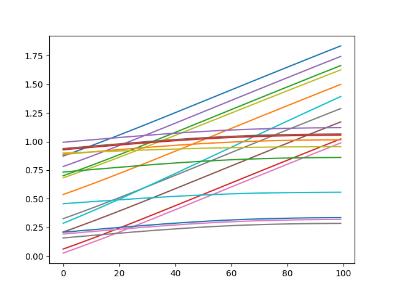
Introduction to Neural Networks Homework1

N26100618 李姵萱

1. Data preprocessing
   1. 類別型資料 : 做one-hot encoding。
   2. 數值型資料 : 做標準化，使平均值=0，標準差=1。
   3. Y : 將yes、no轉換為1與0後，做one-hot encoder。
   4. Imbalance data : 答案為0的資料筆數過多，隨機刪減至與答案為1的資料筆數相同，各4640筆。
   5. Split Data : 將資料分為train 72%、validation 18%、test 1，並對training data做shuffle。
2. Neural Network ( 1 hidden layer )
   1. 設定hidden layer 有10個neuron，output size為2 ( 二分類，因此output layer的activation function固定為softmax )，loss function為cross entropy。
   2. 變化learn rate、hidden layer的activation function，訓練100epochs，觀察loss、accuracy、weight change。
      1. Activation function = sigmoid，learning rate = 0.0007、0.001、0.003、0.005。
         1. Loss變化
         2. Accuracy變化
         3. Weight變化
            * Output layer weights
            * Hidden layer bias
      2. Activation function = relu，learning rate = 0.0007、0.001、0.003、0.005
         1. Loss變化
         2. Accuracy變化
         3. Weight變化
            * Output layer weights
            * Hidden layer bias
3. Radial Basis Function Network
   1. 設定hidden layer 有10個neuron，output size為2 ( 二分類，因此output layer的activation function固定為softmax )，loss function為cross entropy。
   2. 變化learn rate、radial basis function，訓練100epochs，觀察loss、accuracy、weight change。
      1. Radial basis function ，learning rate = 0.001、0.003、0.005。
         1. Loss變化
         2. Accuracy變化
         3. Weight變化
            * Output layer weights
      2. Radial basis function ，learning rate = 0.0005、0.001、0.003。
         1. Loss變化
         2. Accuracy變化
         3. Weight變化
            * Output layer weights
4. 比較與討論
   1. Data preprocessing
      1. Training data有shuffle對於validation data的accuracy提升有幫助。
      2. 資料標準化後訓練時會較順利，即較容易找到訓練得起來的hyperparameters。
      3. 做undersampling較不會只學到單一種資料的資訊，對於準確度提升亦有幫助。
   2. Neural network
      1. 比較相同activation function，不同learning rate
         1. learning rate越大，loss、accuracy、weight收斂越快。
      2. 比較relu與sigmoid兩個activation function
         1. Sigmoid收斂的曲線較平滑，也較不易有overfitting的情形。
         2. Relu只要learning rate稍微過大，很容易就有嚴重的overfitting，甚至比訓練前的結果還糟。
   3. Radial Basis Function Network
      1. 比較相同radial basis function，不同learning rate
         1. learning rate越大，loss、accuracy、weight收斂越快。
         2. learning rate稍大loss就容易overfitting。
      2. 比較2個radial basis function
         1. 需要較大的learning rate。
         2. accuracy的部分在一開始會有停滯期，不像一下子就有明顯成長。
   4. NN與RBFN比較
      1. NN可以訓練的參數較多，搭配足夠多的數據，較容易找到hyperparameter並訓練起來。
      2. RBFN較不容易訓練成功，需多試幾組hyperparameters。且訓練起來後loss非常容易一下子就overfitting。
      3. 使用NN，並用sigmoid當作hidden layer的activation function，最適合本次的資料集。