類神經網路

Homework #1 MLP with BP

Breast Cancer classification

指導老師：李建誠

學生姓名：林冠廷

學號：1010329

系級：通訊系 4A

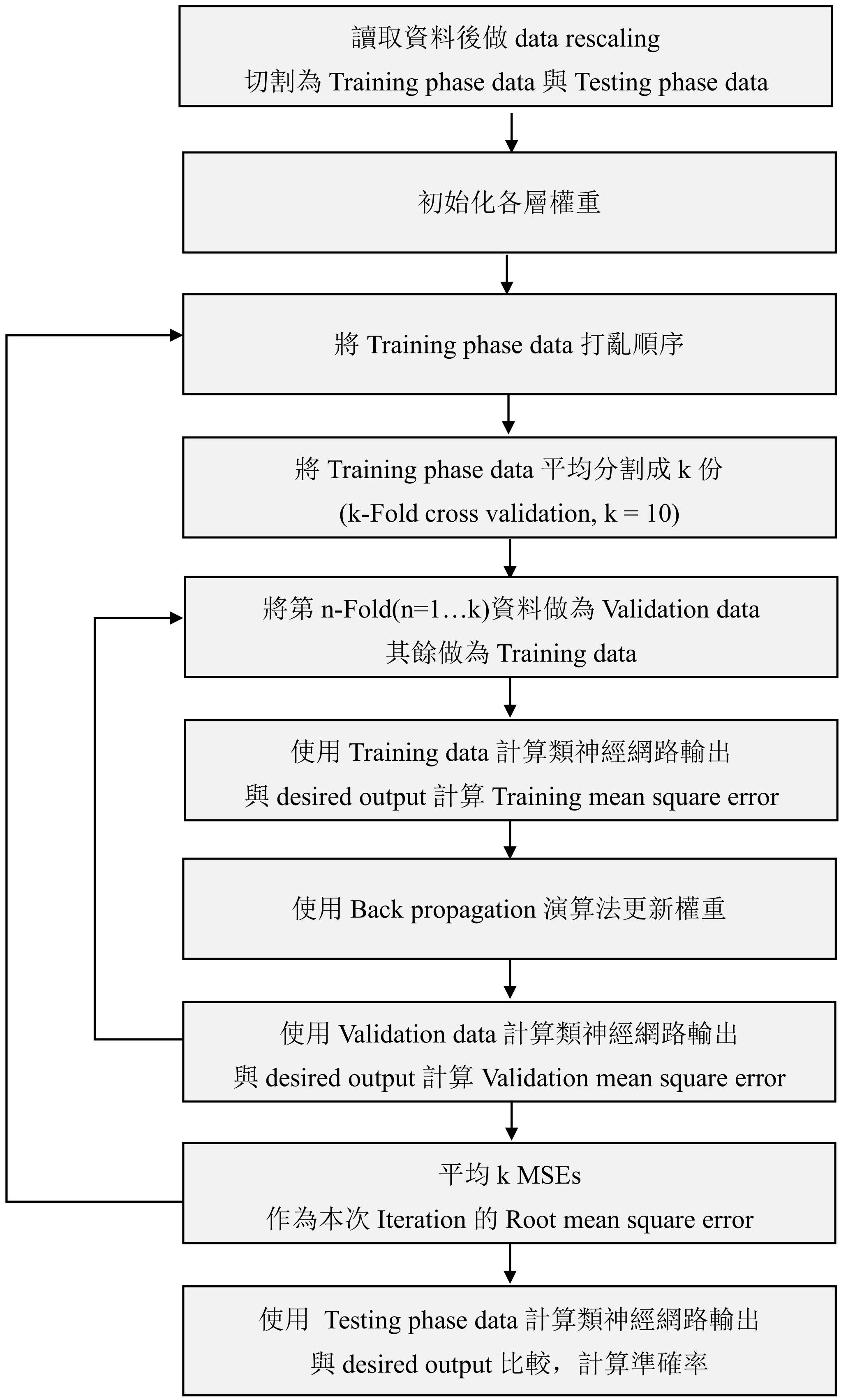
A. 實驗目的

設計 Three-layer Neuron Network 搭配 Back-Propagation 演算法修正權重，對乳癌資料庫做學習，根據Clump Thickness, Uniformity of Cell Size, Uniformity of Cell Shape, Marginal Adhesion, Single Epithelial Cell Size, Bare Nuclei, Bland Chromatin, Normal Nucleoli 與 Mitoses 九個參數，判斷患者罹患良性或惡性腫瘤。

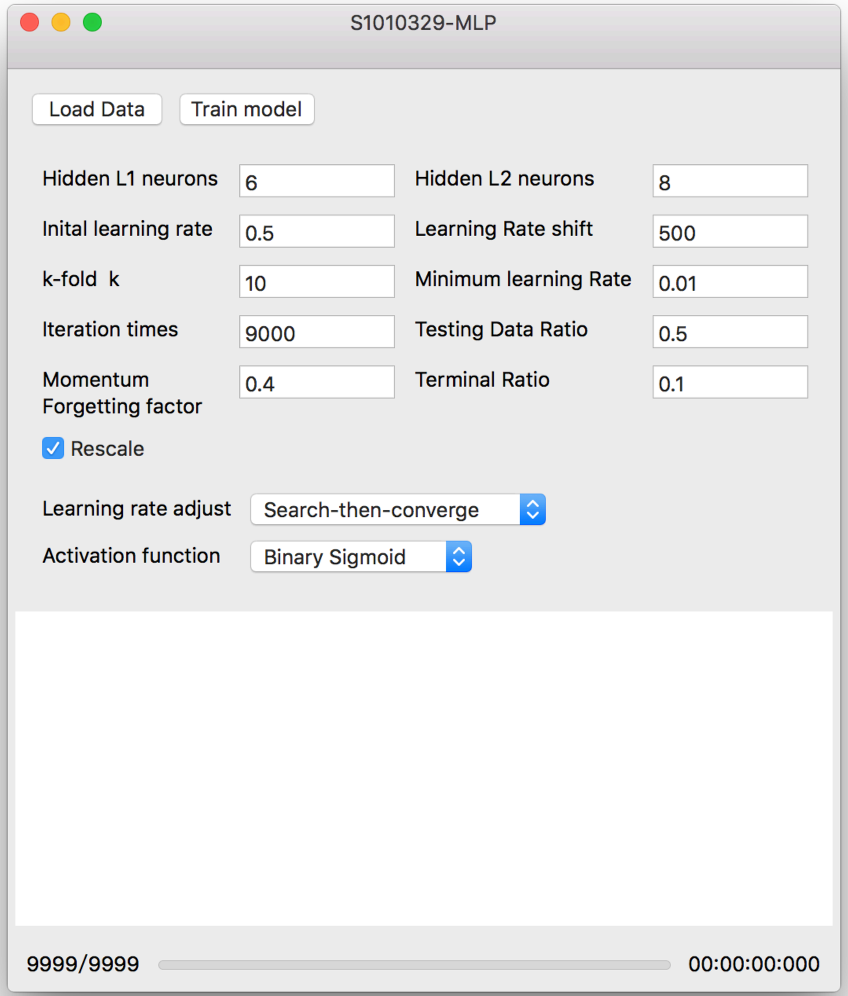
B. 實驗方法

1. 讀入資料
2. 將九項輸入參數Rescale 到 0.01~0.99之間，輸出參數為二維，各表示良性腫瘤與惡性腫瘤
3. 將所有資料分割成Training phase data與Testing phase data
4. 初始化權重(0.1~0.4 uniform)
5. 將Training phase data平均分割成k份
   (k-Fold cross validation, k = 10)
6. 更新權重
7. 計算Training Square Error
8. 根據Error使用Back-Propagation演算法更新權重
9. 計算Validation Square Error
10. 計算RMSE
11. 使用Testing phase data計算類神經網路準確率

C. 實驗流程


讀取資料後做 data rescaling
切割為 Training phase data 與 Testing phase data
初始化各層權重
將 Training phase data 打亂順序
將 Training phase data 平均分割成 k 份
(k-Fold cross validation, k = 10)
將第 n-Fold(n=1…k)資料做為 Validation data
其餘做為 Training data
使用 Training data 計算類神經網路輸出
與 desired output 計算 Training mean square error
使用 Back propagation 演算法更新權重
使用 Validation data 計算類神經網路輸出
與 desired output 計算 Validation mean square error
平均 k MSEs
作為本次 Iteration 的 Root mean square error
使用 Testing phase data 計算類神經網路輸出
與 desired output 比較，計算準確率

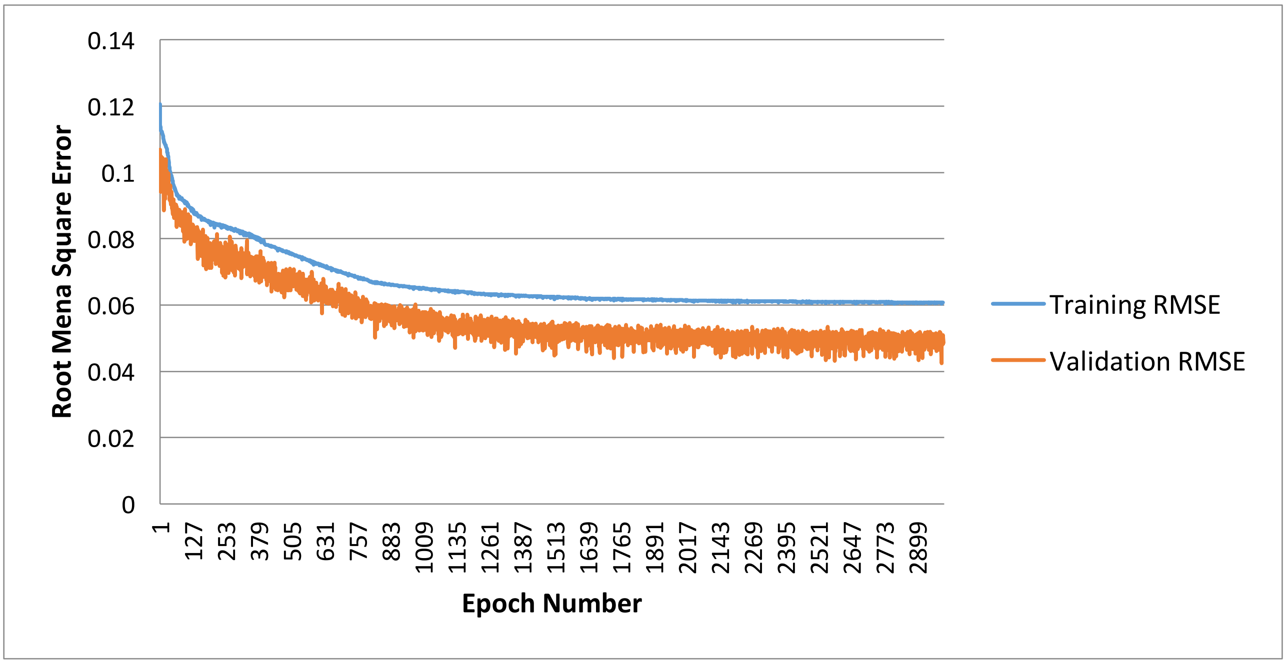
D. 程式操作介面


S1010329-MLP
Load Data
Train model
Hidden L1 neurons
6
Hidden L2 neurons
8
Inital learning rate
0.5
Learning Rate shift
500
k-fold k
10
Minimum learning Rate
0.01
Iteration times
9000
Testing Data Ratio
0.5
Momentum Forgetting factor
0.4
Terminal Ratio
0.1
Rescale
Learning rate adjust
Search-then-converge
Activation function
Binary Sigmoid
9999/9999
00:00:00:000

E. 實驗參數

1. Neuron Network
    a. Three-Layer : 兩層隱藏層與一層輸出層
        1. Number of neurons：
            a. Hidden layer 1 : 6
            b. Hidden layer 2 : 8
            c. Output layer : 2
2. Activation Function
    a. Binary Sigmoid, slope = 0.5
3. Learning rate adjusting
    a. Search then converge , slope = 0.5
4. cross validation
    a. K-fold, k = 10
5. Momentum
    a. Forgetting factor = 0.4

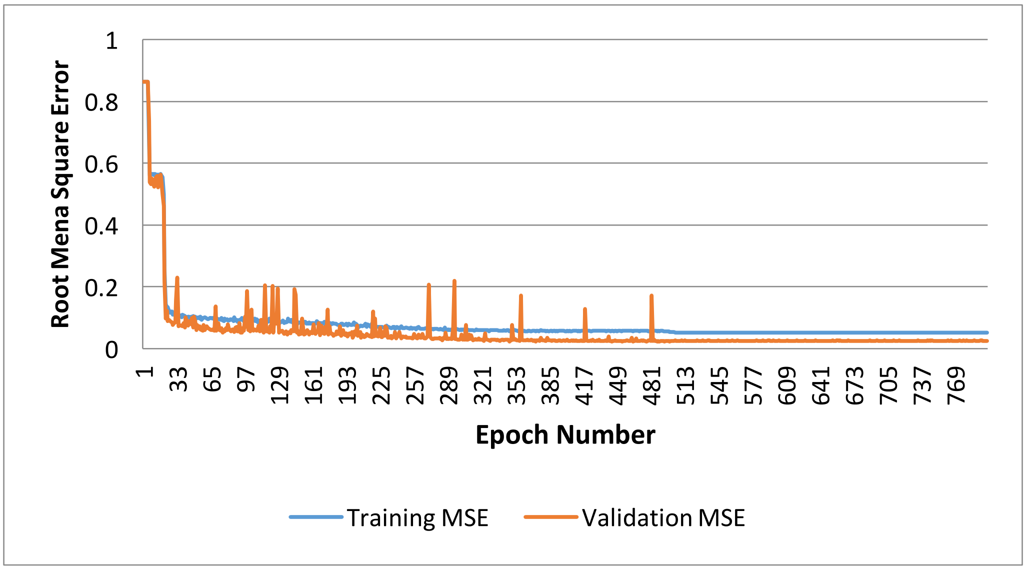
F. 實驗結果



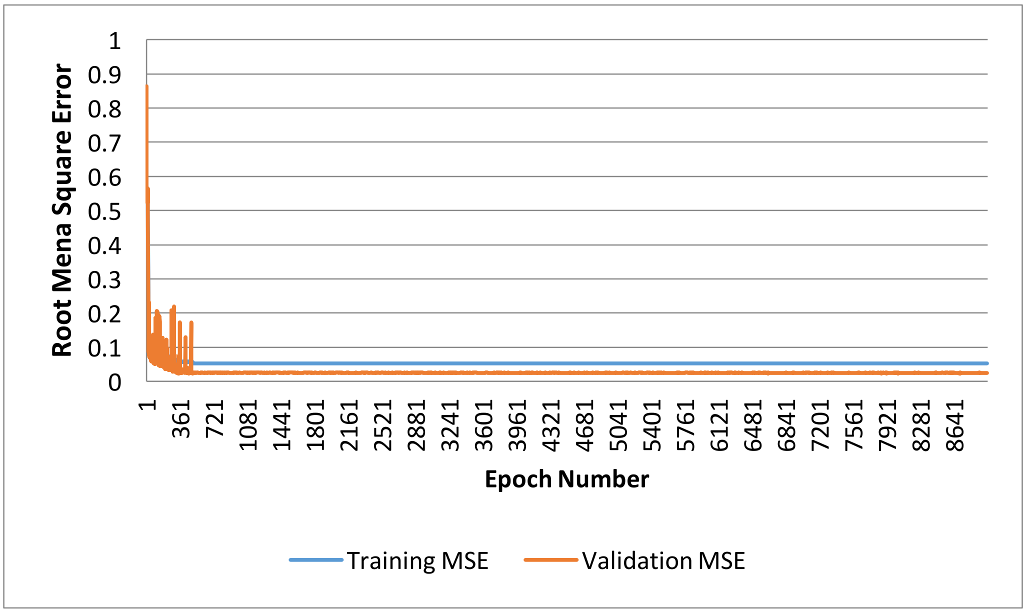
圖一

Epoch	Training MSE	Validation MSE
1	0.120542	0.101057
300	0.082235	0.075398
600	0.072357	0.065119
900	0.066045	0.059805
1200	0.063655	0.052102
1500	0.062317	0.046954
1800	0.061705	0.049519
2100	0.061318	0.050235
2400	0.060878	0.047251
2700	0.060878	0.048456
3000	0.060719	0.048781

表一



(a) Epoch 1~800



(b) Epoch 1~9000

圖二 、第一層隱藏層與第二層隱藏層兩層皆使用99個神經元

Activation function使用Binary Sigmoid(slope = 0.5)

學習率調整使用Search then converge

Momentum Forgetting factor = 0.4

Epoch	**Training RMSE**	**Validation RMSE**
1	0.863128	0.863128
1000	0.051896	0.026077
2000	0.051872	0.024871
3000	0.051834	0.025112
4000	0.051733	0.023816
5000	0.051728	0.024109
6000	0.051726	0.024511
7000	0.051680	0.025230
8000	0.051649	0.024131
9000	0.051616	0.023899

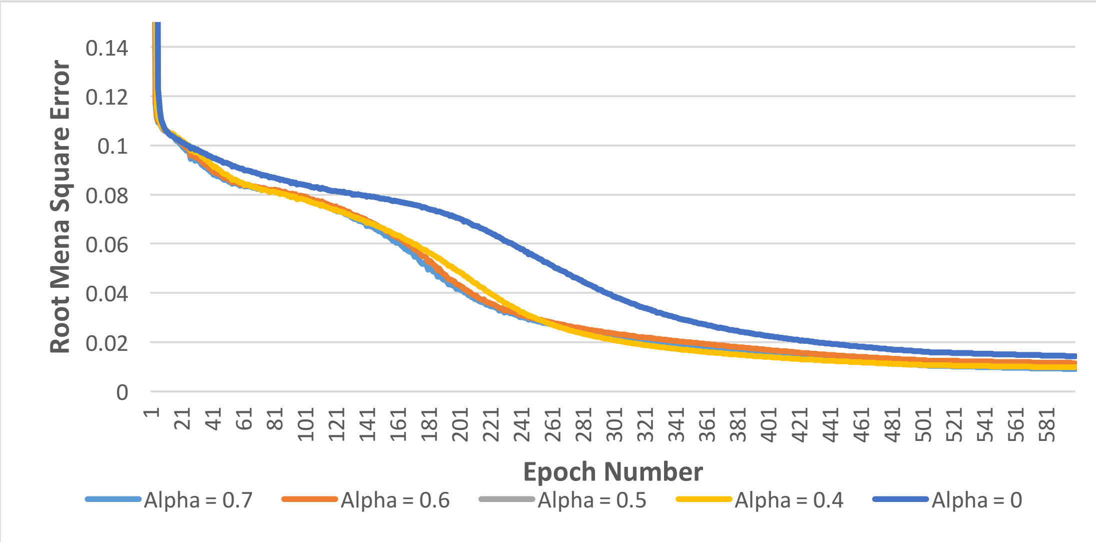
表二、MSE Numerical data Result

圖一與表一為使用兩層隱藏層，第一層隱藏層與第二層隱藏層分別使用6個與8個神經元、Activation function使用Binary Sigmoid(slope = 0.5)，學習率調整使用Search then converge之實驗結果，在Testing Data中的準確率為97.654%。

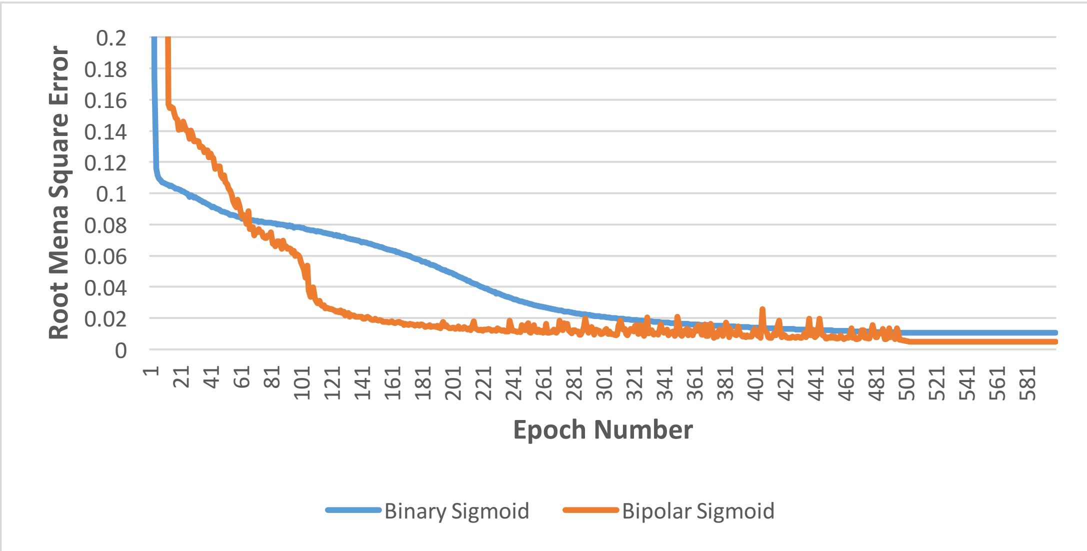
圖二與表二為使用兩層隱藏層，第一層隱藏層與第二層隱藏層兩層皆使用99個神經元、Activation function使用Binary Sigmoid(slope = 0.5)，學習率調整使用Binary Sigmoid(初始學習率0.5，最終學習率0.01，epoch number 500以後逼近0.01)之實驗結果，在Testing Data中的準確率為96.187%。

以上兩個類神經網路為本次實作準確率較佳之結果，可以觀察到對於類神經網路來說，有好的收斂策略是非常重要的，不僅不能收斂過快造成over fitting且還要兼顧訓練的速度。

以下所有比較皆基於此參數設定做更動:Hidden Layer 1 Neurons: 6, Hidden Layer 2 Neurons: 8, Iteration times: 9000, 10-Fold, Learning rate adjust: Search then converge, Activation function: Binary Sigmoid

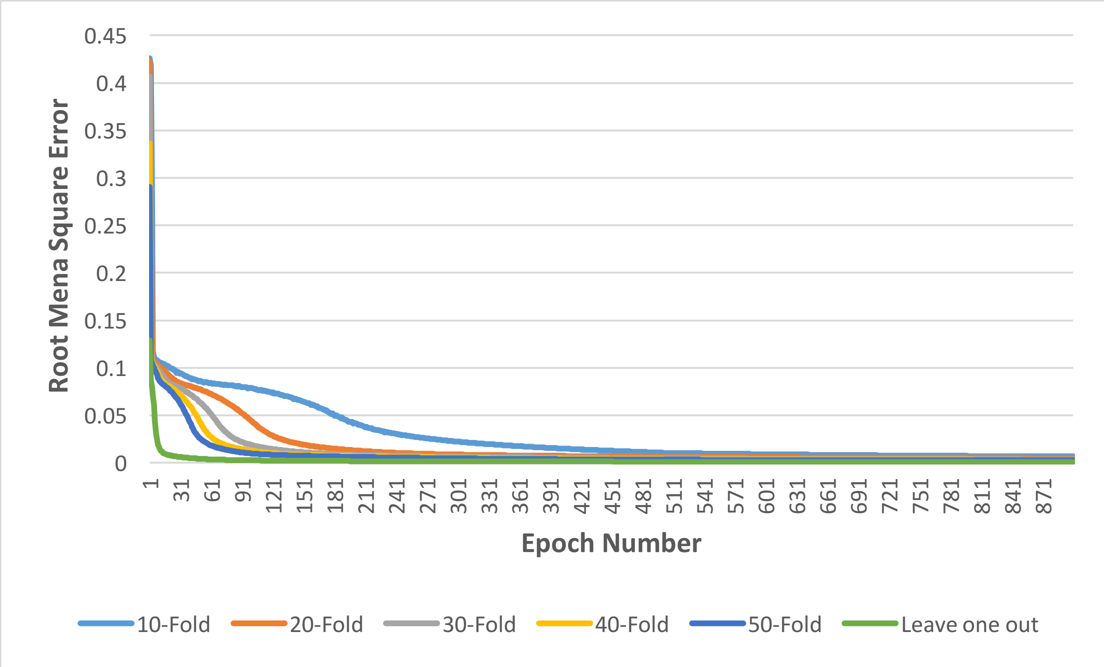


圖三、比較Momentum Forgetting factor之RMSE曲線差異

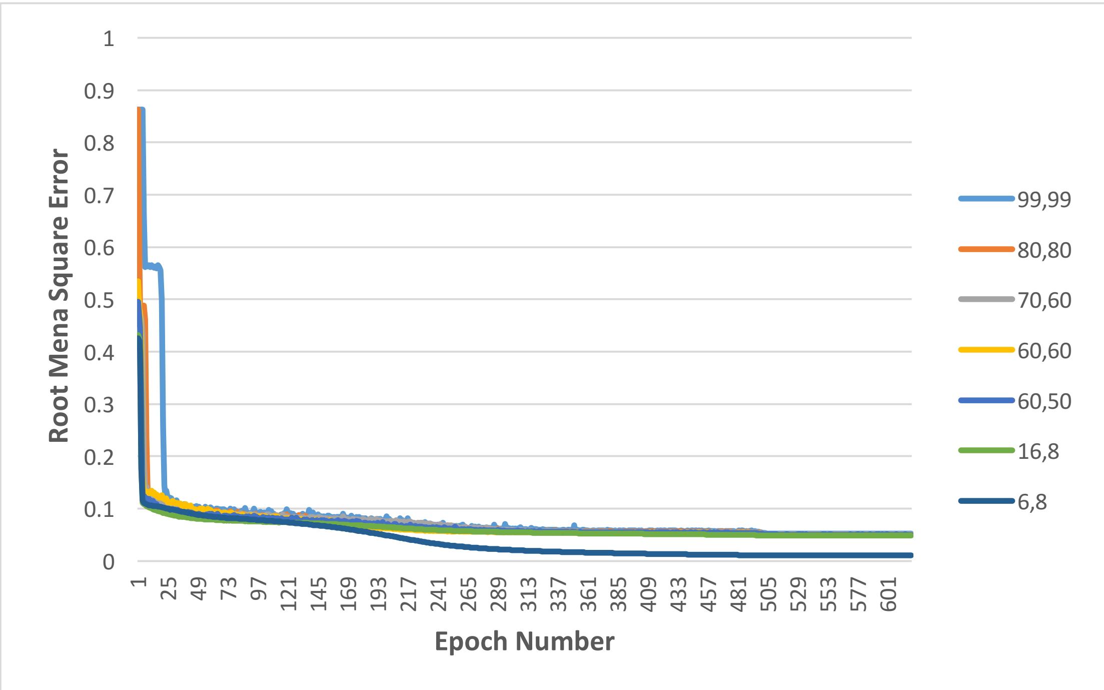


圖四、比較使用Binary Sigmoid與Bipolar Sigmoid作為Activation function之RMSE曲線差異

圖三為比較Momentum之Forgetting factor之差異。其中Alpha=0，即為Standard Learning與Alpha=0.4,0.5,0.6,0.7有明顯差異，Alpha較大會影響一開始的學習速率，RMSE在前期下降的會比較快速。圖四為比較使用Binary Sigmoid與Bipolar Sigmoid作為Activation function之RMSE曲線差異，可以看到Bipolar Sigmoid相較之下較快將RMSE縮小。



圖五、比較使用10, 20, 30, 40, 50 Fold 與 Leave one out Cross Validation 之RMSE曲線差異



圖六、比較神經元數目與RMSE曲線之差異

圖五為比較10, 20, 30, 40, 50 Fold 與 Leave one out Cross Validation 之差異，可以明顯觀察到k值越大在同樣Epoch Number下RMSE較小，但同時的也需要花費更長的時間。

G. 結論

在本作業中學習到了MLP(Multilayer perceptron) 的架構與精神，並且搭配使用Back-Propagation演算法更新權重實作了一類神經網路，根據Clump Thickness, Uniformity of Cell Size, Uniformity of Cell Shape, Marginal Adhesion, Single Epithelial Cell Size, Bare Nuclei, Bland Chromatin, Normal Nucleoli 與 Mitoses 九個參數，判斷患者罹患良性或惡性腫瘤，在測試各種參數過後，發現在第一層隱藏層與第二層隱藏層分別使用6個與8個神經元、Activation function使用Binary Sigmoid(slope = 0.5)，學習率調整使用Search then converge，將資料平均分為兩部分:訓練資料與測試資料，使用訓練資料訓練類神經網路後再使用測試資料測試類神經網路，得到97.654%的準確率。

準確率	測試資料比例	神經元個數(1)	神經元個數(2)	迴圈數	k-fold	Initial Learning rate	min Learning Rate	MomentumAlpha	Leatning rate	Activation function	花費時間
0.976540	0.5	6	8	3.0k	10	0.5	0.001	0.4	Search then converge	Binary Sigmoid	11
0.961877	0.5	99	99	9.0k	10	0.5	0.01	0.4	Binary Sigmoid	Binary Sigmoid	57
0.958944	0.5	80	80	9.0k	10	0.5	0.01	0.4	Binary Sigmoid	Binary Sigmoid	51
0.958944	0.5	70	60	9.0k	10	0.5	0.01	0.4	Binary Sigmoid	Binary Sigmoid	40
0.958944	0.5	60	50	9.0k	10	0.5	0.01	0.4	Binary Sigmoid	Binary Sigmoid	36
0.956012	0.5	60	60	9.0k	10	0.5	0.01	0.4	Binary Sigmoid	Binary Sigmoid	37
0.954212	0.8	6	8	9.0k	10	0.5	0.01	0	Search then converge	Binary Sigmoid	10
0.954212	0.8	6	8	9.0k	10	0.5	0.01	0.4	Search then converge	Binary Sigmoid	10
0.953079	0.5	16	8	9.0k	10	0.5	0.01	0.4	Binary Sigmoid	Binary Sigmoid	19
0.941423	0.7	6	8	9.0k	10	0.5	0.01	0.4	Search then converge	Binary Sigmoid	16
0.939739	0.9	6	8	9.0k	10	0.5	0.01	0.4	Search then converge	Binary Sigmoid	5
0.93643	0.6	6	8	9.0k	10	0.5	0.01	0.4	Search then converge	Binary Sigmoid	21
0.932551	0.5	6	8	9.0k	10	0.5	0.01	0.4	Binary Sigmoid	Binary Sigmoid	25
0.932551	0.5	6	8	9.0k	10	0.5	0.01	0	Search then converge	Binary Sigmoid	24
0.932551	0.5	6	8	9.0k	10	0.5	0.01	0	Search then converge	Binary Sigmoid	19

0.929619	0.5	8	6	9.0k	10	0.5	0.01	0.4	Binary Sigmoid	Binary Sigmoid	26
0.929619	0.5	6	8	9.0k	10	0.5	0.01	0.5	Search then converge	Binary Sigmoid	19
0.926686	0.5	6	8	9.0k	10	0.5	0.01	0.4	Binary Sigmoid	bipolar sigmoid	29
0.923754	0.5	6	8	9.0k	10	0.5	0.01	0.4	Search then converge	Binary Sigmoid	24
0.923754	0.5	6	8	9.0k	10	0.5	0.01	0.7	Search then converge	Binary Sigmoid	21
0.923754	0.5	6	8	9.0k	50	0.5	0.01	0.4	Search then converge	Binary Sigmoid	54
0.923754	0.5	6	8	9.0k	341	0.5	0.01	0.4	Search then converge	Binary Sigmoid	339
0.920821	0.5	6	8	9.0k	10	0.5	0.01	0.6	Search then converge	Binary Sigmoid	21
0.920821	0.5	6	8	9.0k	20	0.5	0.01	0.4	Search then converge	Binary Sigmoid	25
0.920821	0.5	6	8	9.0k	30	0.5	0.01	0.4	Search then converge	Binary Sigmoid	35
0.920821	0.5	6	8	9.0k	40	0.5	0.01	0.4	Search then converge	Binary Sigmoid	45
0.871795	0.4	6	8	9.0k	10	0.5	0.01	0.4	Search then converge	Binary Sigmoid	26
0.852941	0.3	6	8	9.0k	10	0.5	0.01	0.4	Search then converge	Binary Sigmoid	29
0.843137	0.3	6	8	9.0k	10	0.5	0.01	0.5	Search then converge	Binary Sigmoid	30
0.536657	0.5	600	800	1.5k	10	0.5	0.01	0.4	Search then converge	Binary Sigmoid	686