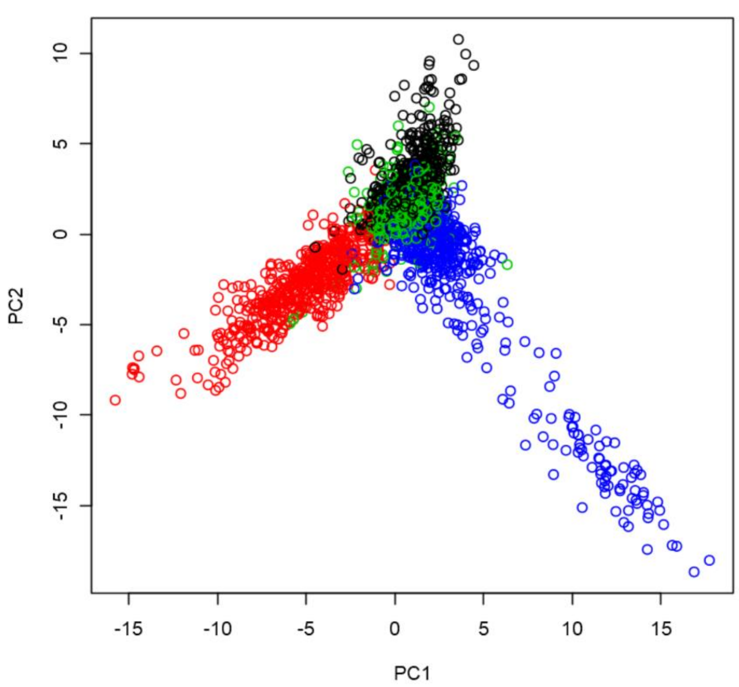
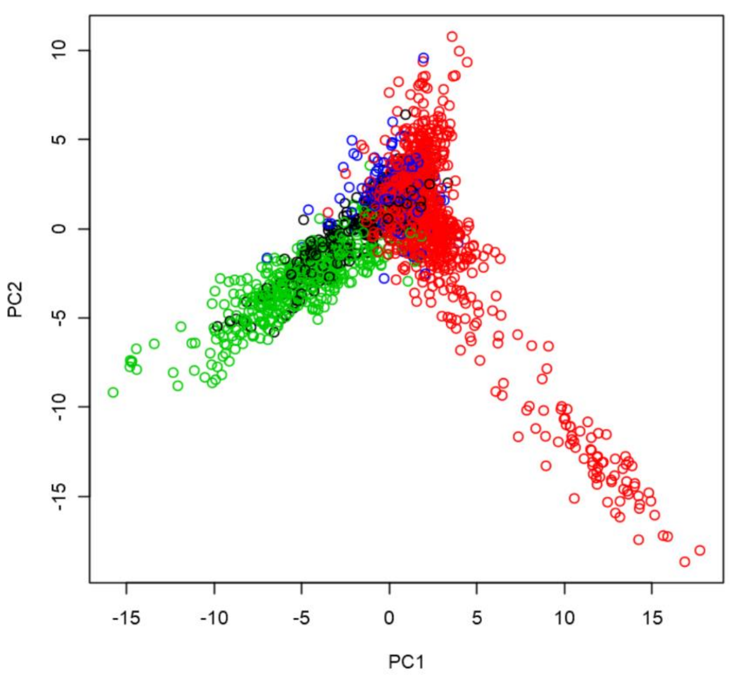
FIT5201 Assessment 2 report
Changze Chen 27717704
Question1

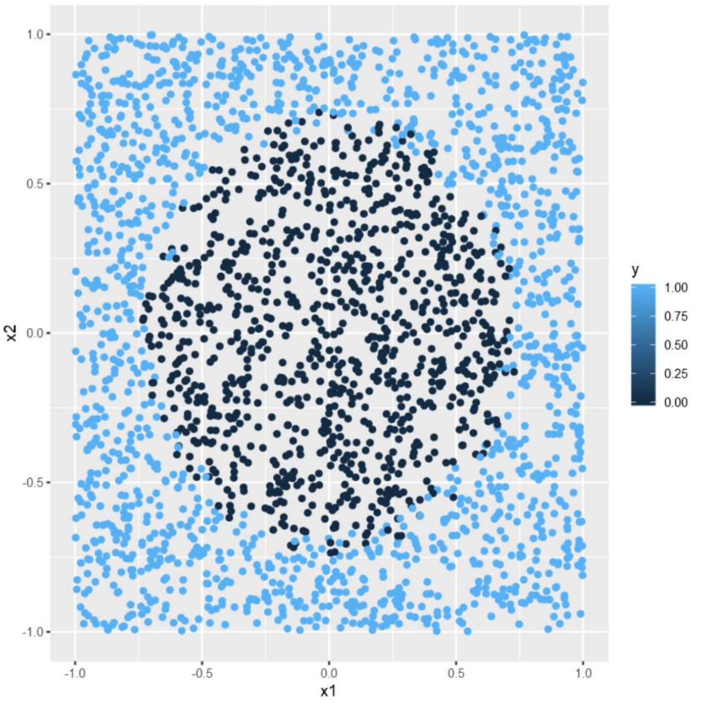
Estimated Clusters (Soft EM)



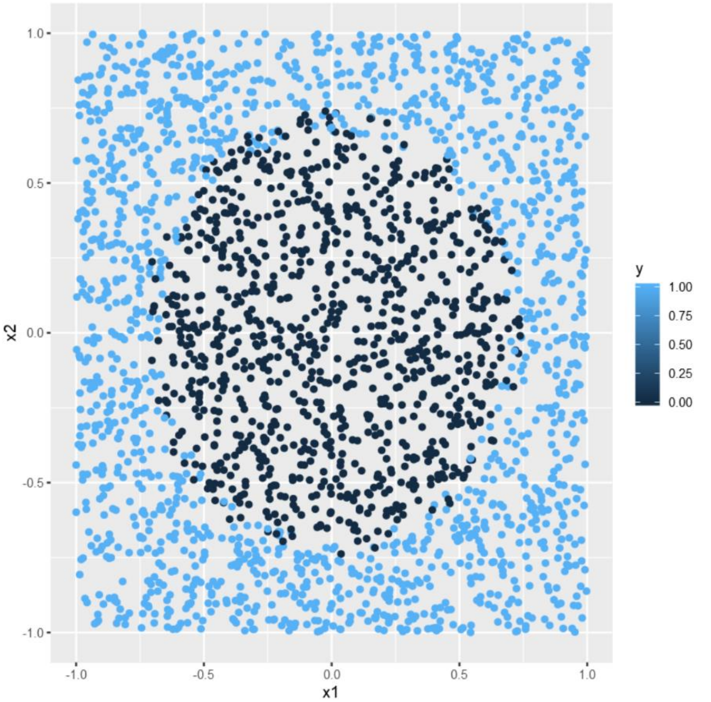
Estimated Clusters (Hard EM)



Question 2
I.
Training data

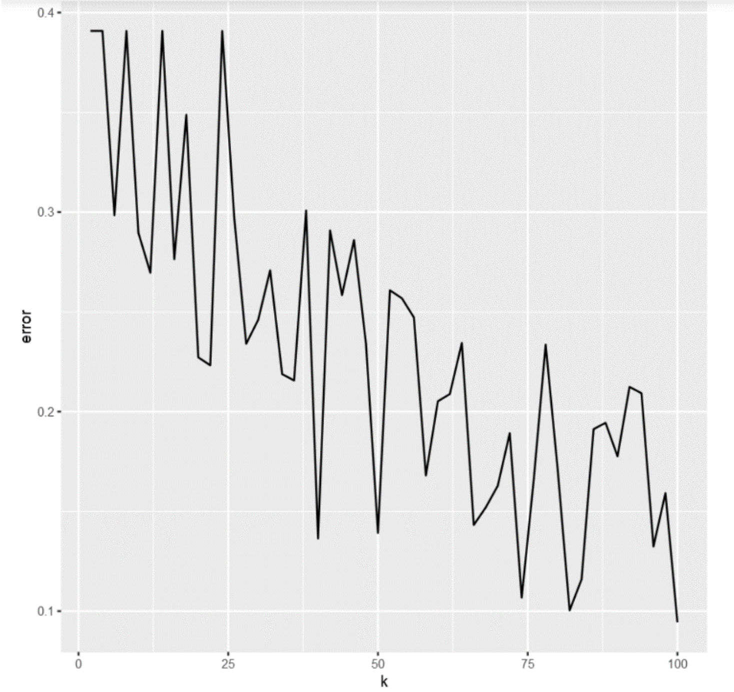


Testing data

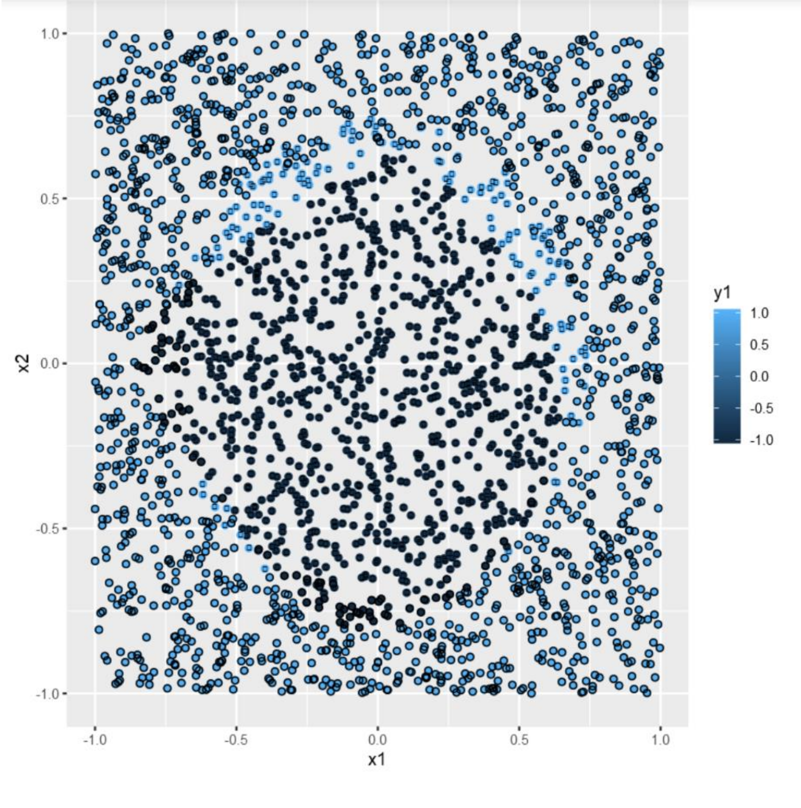


III.

test error vs k



The shape of data points represents to the true label (with the black circle and without the black circle). The colour of data points represent to predicted label



IV.

k	type	error
1	Perceptron	0.5140
2	Neural Network	0.3908
4	Neural Network	0.3908
6	Neural Network	0.2984
8	Neural Network	0.3908
10	Neural Network	0.2896
12	Neural Network	0.2696
14	Neural Network	0.3908
16	Neural Network	0.2764
18	Neural Network	0.3488
20	Neural Network	0.2272
22	Neural Network	0.2232
24	Neural Network	0.3908
26	Neural Network	0.2972
28	Neural Network	0.2340
30	Neural Network	0.2460

k	type	error
32	Neural Network	0.2708
34	Neural Network	0.2188
36	Neural Network	0.2156
38	Neural Network	0.3008
40	Neural Network	0.1364
42	Neural Network	0.2908
44	Neural Network	0.2584
46	Neural Network	0.2860
48	Neural Network	0.2340
50	Neural Network	0.1392
52	Neural Network	0.2608
54	Neural Network	0.2568
56	Neural Network	0.2472
58	Neural Network	0.1680
60	Neural Network	0.2052
62	Neural Network	0.2088

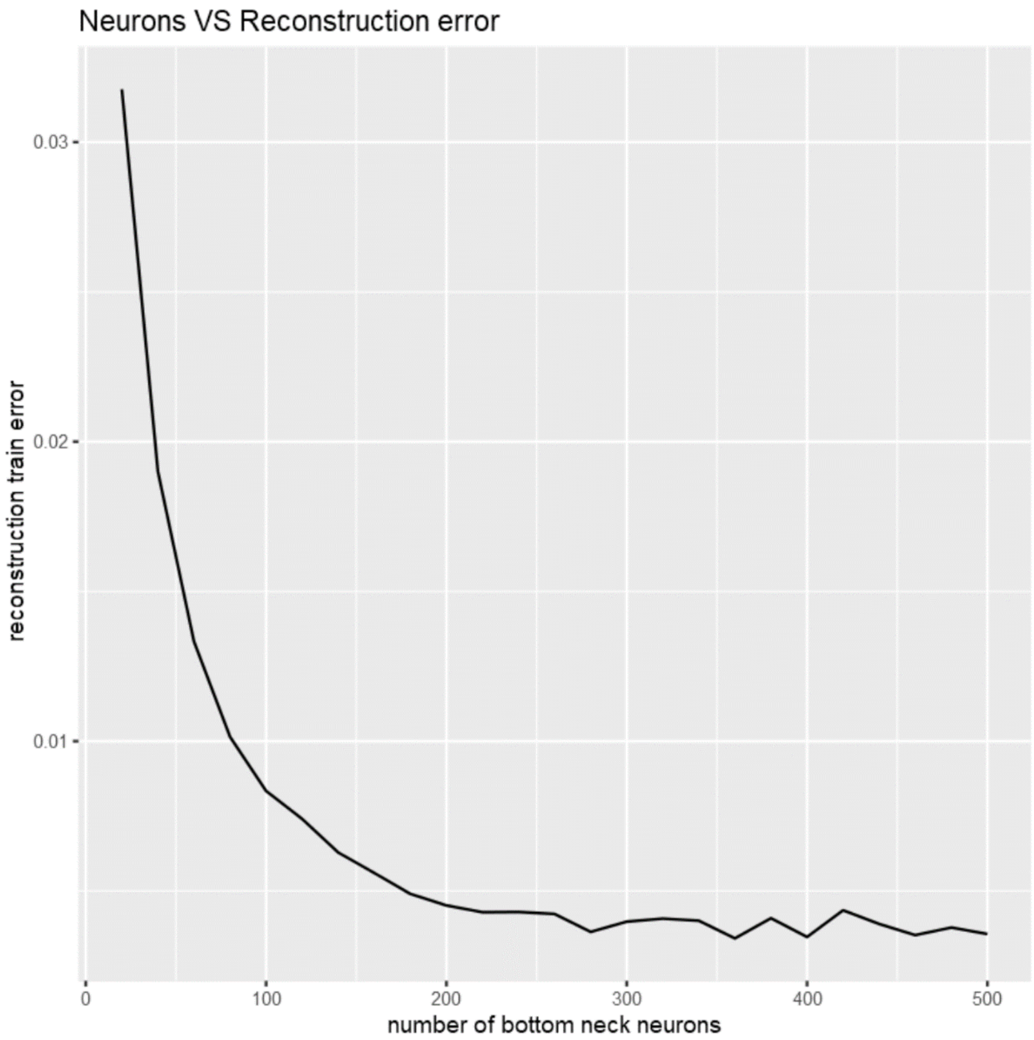
k	type	error
64	Neural Network	0.2344
66	Neural Network	0.1432
68	Neural Network	0.1520
70	Neural Network	0.1628
72	Neural Network	0.1892
74	Neural Network	0.1068
76	Neural Network	0.1656
78	Neural Network	0.2336
80	Neural Network	0.1712
82	Neural Network	0.1004
84	Neural Network	0.1160
86	Neural Network	0.1912
88	Neural Network	0.1944
90	Neural Network	0.1776
92	Neural Network	0.2124
94	Neural Network	0.2092

k	type	error
96	Neural Network	0.1324
98	Neural Network	0.1592
100	**Neural Network**	**0.0944**

V.

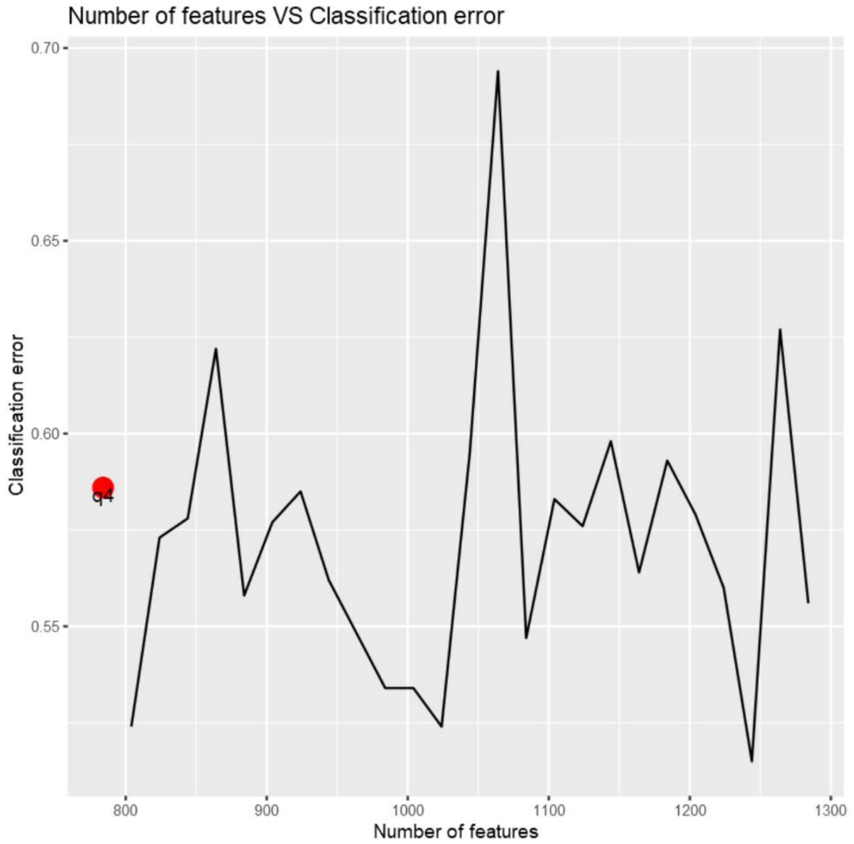
The decision boundary of binary perceptron classifier is always a straight line. However, a complicated shape decision boundary can exist in the neural network classifier. In my opinion, since NN uses a non-linear activation function in hidden layers, it can comprise more complex functions to the classifier input. So, the decision boundary can be any shape.

Question 3
III.



From the plot, we can see, in the beginning, when the number of neurons in the middle layer increases, the reconstruction error decrease. However, as the number of neurons reaches 280, the reconstruction error rises slightly and starts fluctuating, which means the model may tend to be overfitting or we can explain that is because the model may be not stable. In my personal opinion, the slight increment and fluctuation of reconstruction error reflect the instability of our model.
When an autoencoder learns to project the data onto its middle layer neurons, generally, if we have an autoencoder with n inputs and one hidden layer of dimension n (or more), it could potentially just end up learning the identity function. That is called overfitting, and we should avoid it. But the reconstruction error should be close to 0 instead of fluctuating even our model is overfitting.

VI.



VIII.

It seems that if the classification error is low, the reconstruction error must be small. But, if the reconstruction error is small, the classification error may not always be small.
In my personal opinion, if the reconstruction error is low, it means that the new features in the middle layer units figure high dimensional training features well. But it does not say that those features are suitable for classification. On the other hand, in a classification training process, when the classification error is small, it means that not only the model performs well, but also the features which we chose are very representative. Thus, the low classification error must have a small reconstruction error, but a low reconstruction error may not cause a low classification error.