ALL EXPERIMENTS RUN ON MAC

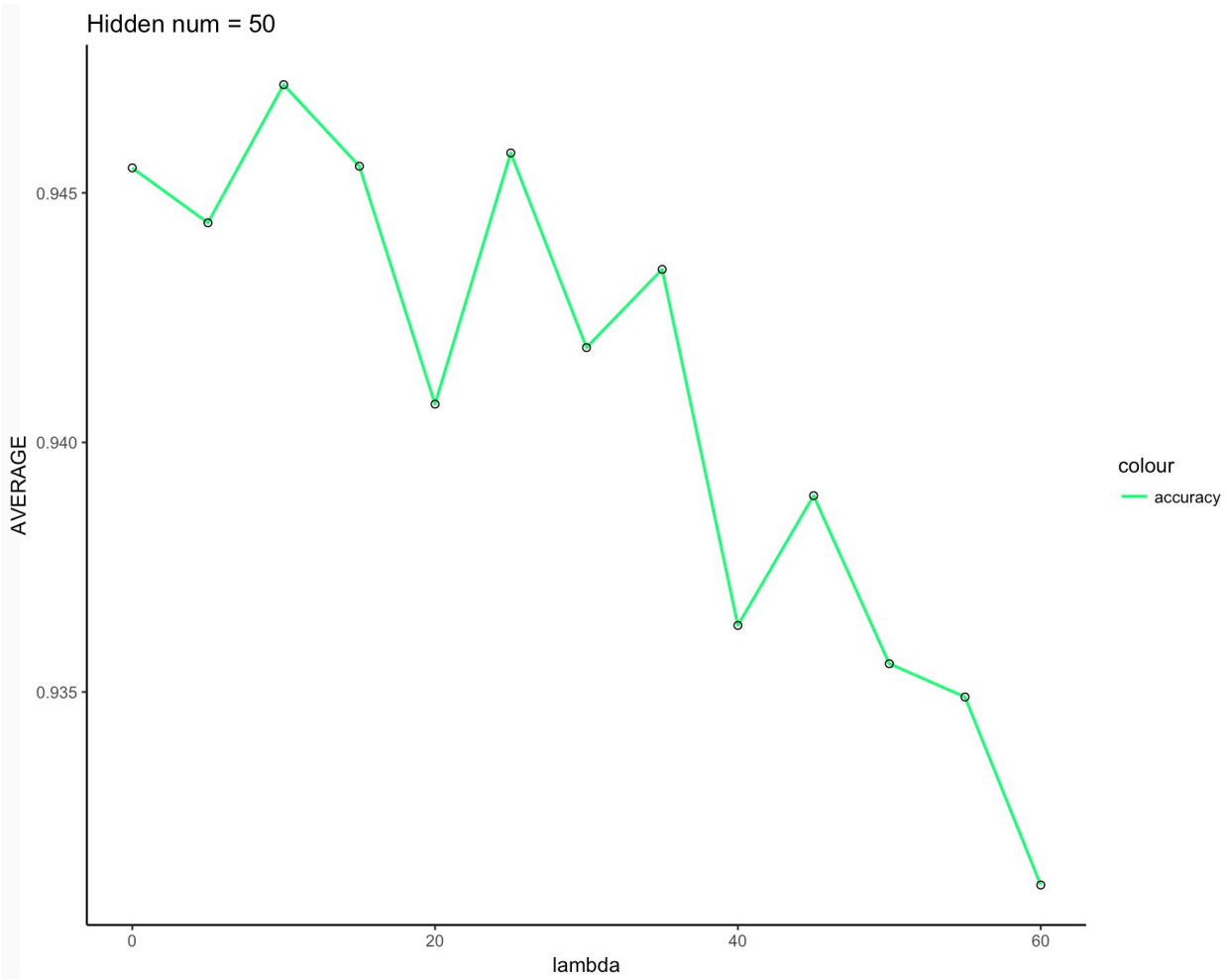
Explanation with supporting _gures of how to choose the hyper-parameter for Neural Net-work:

One hidden layer neural network

First, here is the **validation set accuracy** of result correspond to different lambda value with same number of hidden units. Hidden unit = 50
In this test, for each lambda value, I run 3 times and took the average , here is the experiment data.

lambda	1st	2nd	3rd	average accuracy
0	0.946	0.9418	0.9487	0.9455
5	0.949	0.9429	0.9423	0.944733333
10	0.9435	0.95	0.948	0.947166667
15	0.949	0.947	0.9406	0.945533333
20	0.9454	0.9369	0.94	0.940766667
25	0.9435	0.9439	0.95	0.9458
30	0.943	0.939	0.9437	0.9419
35	0.9419	0.9441	0.9444	0.943466667
40	0.9382	0.9336	0.9372	0.936333333
45	0.9379	0.9365	0.9424	0.938933333
50	0.9345	0.9386	0.9336	0.935566667
55	0.9369	0.9351	0.9327	0.9349
60	0.9314	0.9237	0.9293	0.928133333

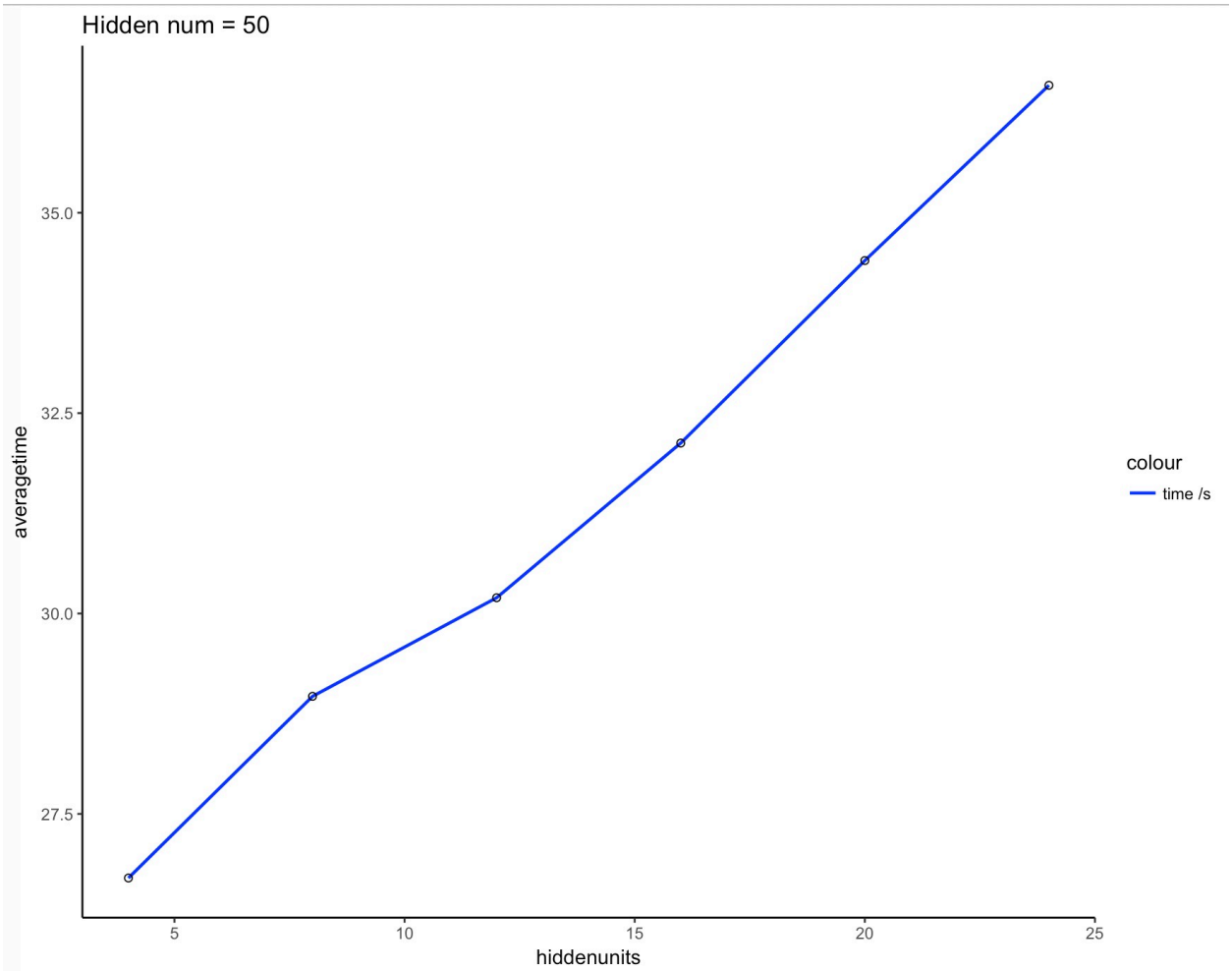
Here is chart correspond to the lambda value and average accuracy



From above chart, we could found that after we set a suitable lambda, we could get a higher accuracy on validation set. However how could we find a suitable lambda value? Usually, we should set lambda as 0 at first, and then according to the validation accuracy to increase lambda value. After you tried to set series value of lambda, you could choose a optimal lambda for the dataset.

After I confirm the optimal value of lambda, I **fixed lambda value**, **change** the number of **hidden units** from 4 to 36 to observe time cost, for each hidden unit, I **run 3 times** and took the **average** time, the experiment data and chart as follows.

lambda	hiddenunits	time 1st	time 2nd	time 3rd	average
10	4	26.43	27.91	25.76	26.7
10	8	29.24	28.81	28.85	28.9666667
10	12	30.23	31.04	29.32	30.1966667
10	16	32.69	31.439	32.25	32.1263333
10	20	34.22	33.99	35	34.4033333
10	24	36.26	35.67	37.84	36.59



From the time experiment, we could found that with the increase number of hidden units, the time cost also increase.

Accuracy of classification method on the handwritten digits test data:

Here is the **accuracy of handwritten digit**.

	train	validation	test
accuracy	95.35%	94.78%	95.08%

Accuracy of classi_cation method on the CelebA data set:

Then, I run the **facennScript**, lambda value is 10, hidden units is 256, I **run 3 times** and **take** the **average accuracy based on test data**. Here is the experiment data.

	1st	2nd	3rd	average
test accuracy	86.60%	85.69%	85.39%	85.89%

Comparison of your neural network with a deep neural network (using TensorFlow) in terms of accuracy and training time:

Here is the comparison between the results of **deep neural network with 3,5,7 hidden layers** and **neural network with one hidden layer** on the CelebA data set. All accuracy results are test set accuracy.

	deepnn 3 hidden layers	deepnn 5 hidden layers	deepnn 7 hidden layers	1 hidden layer nn
Test set Accuacy	79.41%	76.91%	75.51%	86.15%
Time cost(s)	259.4335401	307.9706678	368.2495339	105.903285

From the data, we could find that deepnn takes more time than 1 hidden layer neural network. However, when deepnn has 3 hidden layers the accuracy is lower than 1 hidden layer neural network, maybe it because overfitting happened on deepnn. With the increase hidden layer numbers of deepnn, the accuracy decreased.

Report the results from convolutional neural network in terms of accuracy and training time.

Here is the **time cost** and **test set accuracy** run by cnnScript, I run 3 times and then take the average.

	1st	2nd	3rd	average
time(s)	598	597	592	595.666667
accuracy	98.60%	98.60%	98.80%	98.67%