

Report for
CSE 474/574 Introduction to Machine Learning
Programming Assignment 1
Handwritten Digits Classification

Course Number: CSE 474

Student Name:

Steven Mundt

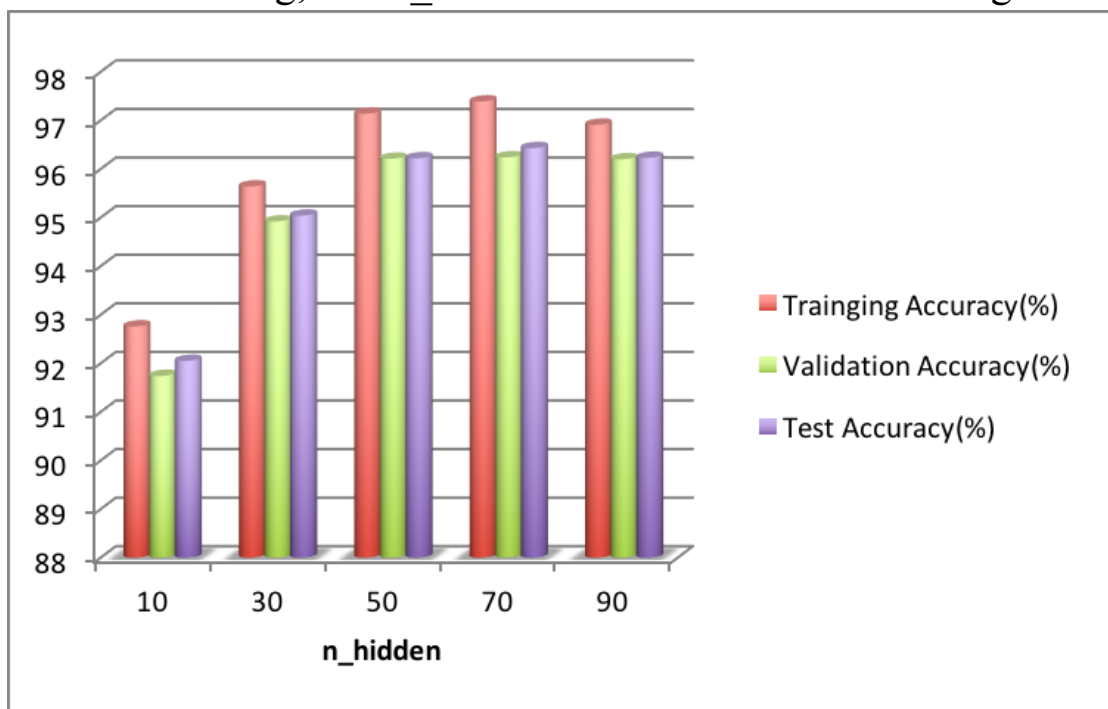
Qinxin Tian

Fengyu Wu

1. How to choose hyper-parameter: *number of hidden units*

In this program, 5 different n_hidden which are 10, 30, 50, 70, 90 were tested when keeping *iterations* and *lambdaval* unchanged (*lambdaval*=0.6). We set *iterations*=100 because it was found we were underfitting the model when *iterations*=50.

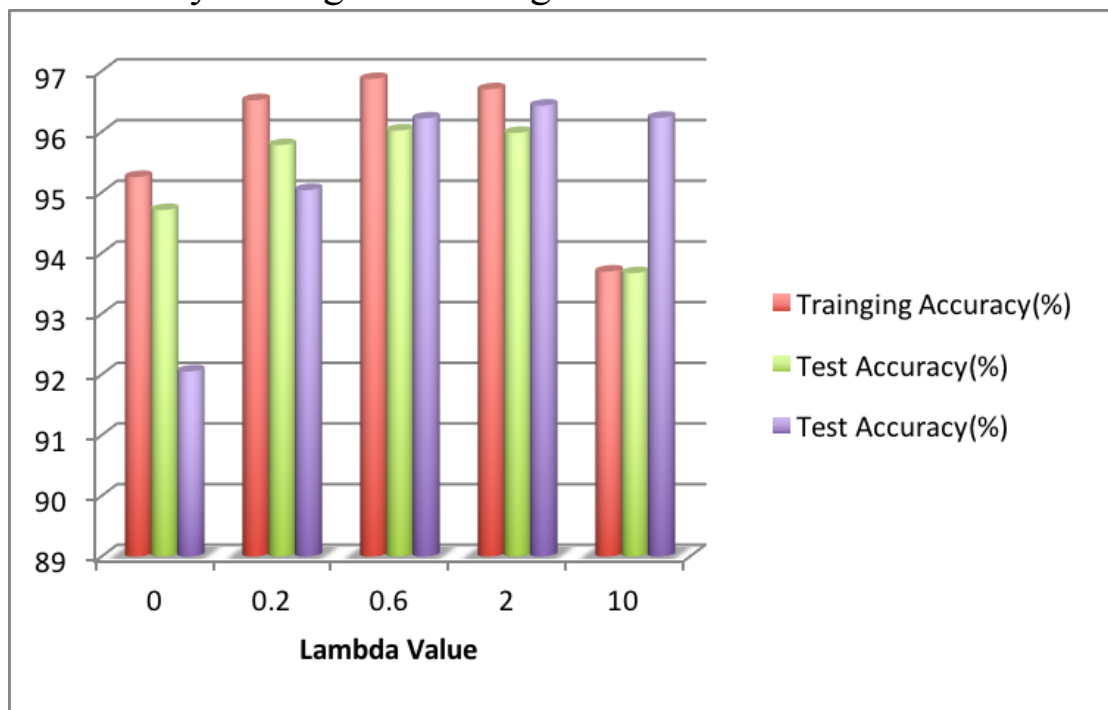
It is found that $n_hidden=50$ is the best, n_hidden less than 50 is underfitting, and n_hidden more than 50 is overfitting.

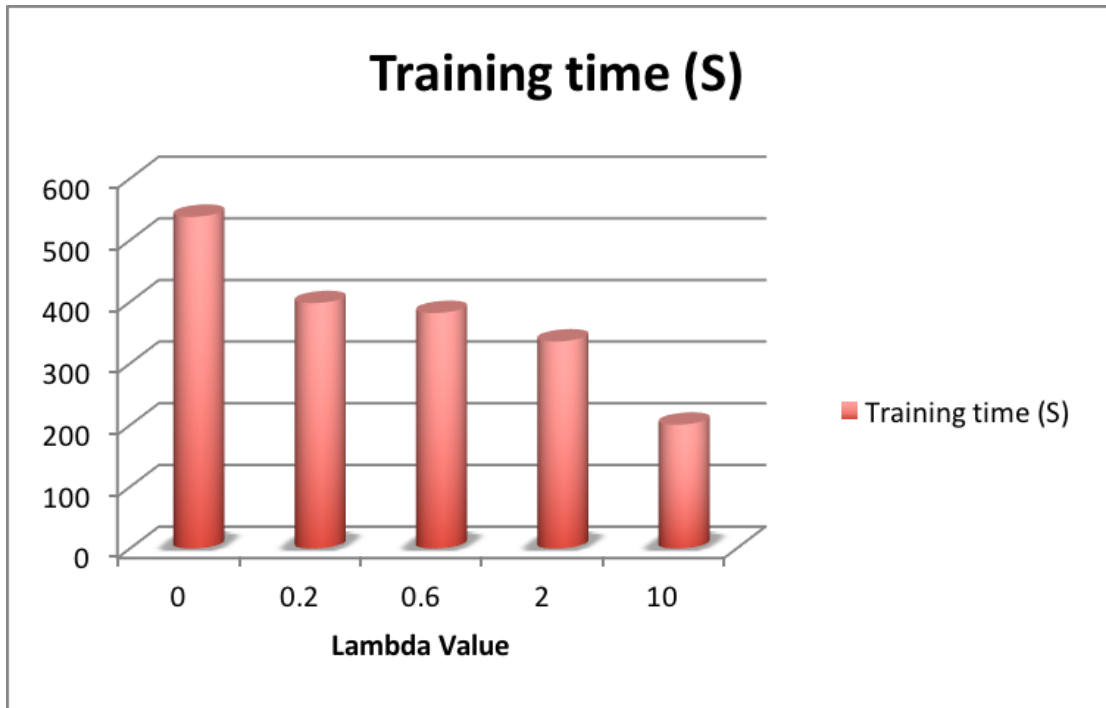


2. How to choose hyper-parameter: *regularization term λ*

We tested 5 different λ which are 0.0, 0.6, 0.2, 2.0, 10 when keeping n_{hidden} ($n_{\text{hidden}}=50$) unchanged.

It is found that $\lambda=0.6$ is the best given it has the best accuracy and a good training time.





3. Classification accuracy on handwritten digits test data:

When choosing the following parameters ($n_{\text{hidden}}=50$, $\text{lambdaval}=0.6$, $\text{iterations}=100$), we got the following results:

Training accuracy: 97.39%

Validation accuracy: 96.24%

Test accuracy: 96.43

We think this is the best combination, because test accuracy is the most important for us.

```
preprocess done.
```

```
Training...
```

```
Training done.
```

```
Final loss: 0.03461201844005686
```

```
Training set Accuracy:97.394%
```

```
Validation set Accuracy:96.24%
```

```
Test set Accuracy:96.43%
```

```
In [64]:
```