HW06

To achieve highest accuracy, I used several methods.

Firstly, I changed the parameter M in cross-validation.

```
Confusion matrix is:
[[56. 20.]
 [ 4. 40.]]
Percentage Correct:  80.0
M is  10
```

```
Confusion matrix is:
[[60.  0.]
 [ 0. 60.]]
Percentage Correct:  100.0
M is  15
```

Figure 1

From Figure 1 we find when M is 15, we get a better split of original data and a higher accuracy.

```
nhidden is 1
Stopped 2.390263670131029 2.391143522147944 2.3920887788738234
Count is  38
Confusion matrix is:
[[40.  0.]
 [20. 60.]]
Percentage Correct:  83.33333333333334
nhidden is 2
Stopped 2.071401498854219 2.0723243666987745 2.073303574711268
Count is  40
Confusion matrix is:
[[40.  0.]
 [20. 60.]]
Percentage Correct:  83.33333333333334
nhidden is 3
Stopped 2.4253791244893215 2.4263212799824236 2.4273123503278287
Count is  56
Confusion matrix is:
[[40.  0.]
 [20. 60.]]
Percentage Correct:  83.33333333333334
nhidden is 4
Stopped 0.17013205510445967 0.17104333314244932 0.17200673262777333
Count is  35
Confusion matrix is:
[[60.  0.]
 [ 0. 60.]]
Percentage Correct:  100.0
nhidden is 5
Stopped 0.22074248063760932 0.22170929602727305 0.222692597016871
Count is  65
Confusion matrix is:
[[60.  0.]
 [ 0. 60.]]
Percentage Correct:  100.0
nhidden is 6
Stopped 0.161294266162656 0.1622616714202547 0.1632469919228865
Count is  73
Confusion matrix is:
[[60.  0.]
 [ 0. 60.]]
Percentage Correct:  100.0
```

Figure 2

In Figure 2, I changed the number of nodes on hidden layer. When number of nodes is too small, the network works badly. And when is number is large enough, the network works well. We know network whose number of nodes is too large may be overfitting, but we can’t show that in this lab because we only have training data and it works very well for our training data. Here I choose 6 as my number of nodes.

```
eta is  0.1
Count is  89
Percentage Correct:  100.0
eta is  0.2
Count is  16
Percentage Correct:  100.0
eta is  0.3
Count is  15
Percentage Correct:  100.0
eta is  0.4
Count is  53
Percentage Correct:  100.0
eta is  0.5
Count is  20
Percentage Correct:  100.0
eta is  0.6
Count is  11
Percentage Correct:  100.0
eta is  0.7
Count is  6
Percentage Correct:  100.0
eta is  0.8
Count is  5
Percentage Correct:  100.0
eta is  0.9
Count is  12
Percentage Correct:  100.0
```

Figure 3

```
eta is  2
Stopped 1.821938377712026 1.8218704459902701 1.822117770861831
Count is  79
Confusion matrix is:
[[40.  0.]
 [20. 60.]]
Percentage Correct:  83.33333333333334
```

Figure 4

```
eta is  1e-05
Stopped 4.709788024673868 4.710782735554716 4.711782143360013
Count is  611
Confusion matrix is:
[[50. 40.]
 [10. 20.]]
Percentage Correct:  58.333333333333336
```

Figure 5

Figure 3&4&5 show results changing learning rate eta. When eta is too small or too large, we got a low accuracy. Because we are able to be trapped at local optima with a small eta and it’s also hard to get global optimal when eta is too large. Meanwhile, eta influences the number of iterations. In Figure 3, although we got same correct percentage, the number of iteration changing from large to small then to large again. Here the number of iteration equals to count*100 (default niteration).

```
iteration is  1
Stopped 2.275575363693911 2.2765691731665574 2.277568733641949
Count is  378
Confusion matrix is:
[[40. 20.]
 [20. 40.]]
Percentage Correct:  66.66666666666666
```

Figure 6

```
iteration is  5
Stopped 5.06480895843032 1.9416517052000781 0.9976571738550354
Count is  151
Confusion matrix is:
[[25.  0.]
 [35. 60.]]
Percentage Correct:  70.83333333333334
```

Figure 7

```
iteration is  50
Stopped 0.17244382942915393 0.17153612834478266 0.16828708306500328
Count is  54
Confusion matrix is:
[[60.  0.]
 [ 0. 60.]]
Percentage Correct:  100.0
```

Figure 8

From Figure 6&7&8, when iteration is too small, we get a relative bad result. We get a higher accuracy as number of iterations increases. But even using earlystopping function, the total iteration which equals to count * iteration increases and we need a longer time to get result.

All in all, I changed M, eta, number of hidden layer and number of iteration to get a better accuracy.