

Lab 01

8.

	Correctly classified instances	Incorrectly classified instances
Training Set	143	12
Cross Validation(10 folds)	130	25

9.

A confusion matrix is a summary of prediction results on a classification problem.

The number of correct and incorrect predictions are summarized with count values and broken down by each class.

The confusion matrix shows the ways in which your classification model is confused when it makes predictions.

Simply the confusion matrix shows that how many times the correct classification occurs and how many times an incorrect classification occurs. The number of times the prediction differs from the actual expected output.

Example:

Training set

a b <-- classified as

22 10 | a = DIE

2 121 | b = LIVE

In here, the number of times

DIE classified as DIE = 22

DIE classified as LIVE = 2

LIVE classified as DIE = 10

LIVE classified as LIVE = 121

Cross validation (10 folds)

a b <-- classified as

14 18 | a = DIE

7 116 | b = LIVE

DIE classified as DIE = 14

DIE classified as LIVE = 7

LIVE classified as DIE = 18

LIVE classified as LIVE = 116

10.

	Correctly classified instances	Incorrectly classified instances
Training Set	149	6
Cross Validation(10 folds)	126	29

Above shown are the results obtain after the given parameters are set. It has affected both the situations in a different manner.

When comparing the results obtain from the training set in this situation, with the default, we can see that changing those parameters has increased the correctly classified instances and decreased the incorrectly classified instances.

But with the cross validation (10 folds), changed parameters has decreased the correctly classified instances, and increased the incorrectly classified instances.

Let's compare the confusion matrix in both situations.

Training set (default)

a b <-- classified as

22 10 | a = DIE

2 121 | b = LIVE

Training set (changed parameters)

a b <-- classified as

28 4 | a = DIE

2 121 | b = LIVE

As you can see it has increased the correctly classified instances of DIE classified as DIE and also it has decreased the incorrectly classified instances of LIVE classified as DIE. And others has not changed.

Cross validation (default)

```
a      b <-- classified as  
14     18 | a = DIE  
7      116 | b = LIVE
```

Cross validation (changed parameters)

```
a      b <-- classified as  
18     14 | a = DIE  
15     108 | b = LIVE
```

Although it has increased the DIE classified as DIE, LIVE classified as LIVE has decreased.