* Compare and analyze the average accuracy, precision, recall and F1-score of semi-supervised learning with the ones obtained during the supervised learning.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Recall | Precision | Accuracy |
| Supervised | 33.27% | 94.12% | 94.42% |
| Semi-Supervised | 30.57% | 86.00% | 90.97% |

In three cases Recall,Precision and Accuracy of Supervised learning > Semi supervised learning.

In supervised learning all the train data have correct labels.

But in Semi supervised learning the if any one of the unlabeled data gets a wrong label on step 4 ( where using the learned tree we give the unlabeled training data label, then error is introduced in the train data.

* Compare and analyze the accuracies obtained using cross validation for k=5,10,20, leave-one-out cross validation and without using any cross validation.

|  |  |
| --- | --- |
| K=? | Accuracy % |
| 5 | 93.5020138592 |
| 10 | 98.5680985681 |
| 20 | 98.5857198357 |
| 1 (Leave out one Cross Validation) | 99.8505231689 |

For Leave out one cross Validation ( K = 1 ) gives highest accuracy.

Because we are giving All data except one which is the test data.

But it take excessive computation of order Ω(n^2) where n is the data set numbers.

As the value of K increases from K = 5 to 20 the accuracy also increases.