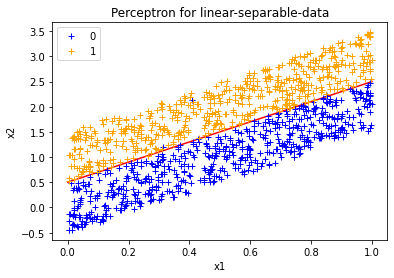
**Assignment 1 Part 2 Report for CSE-512**

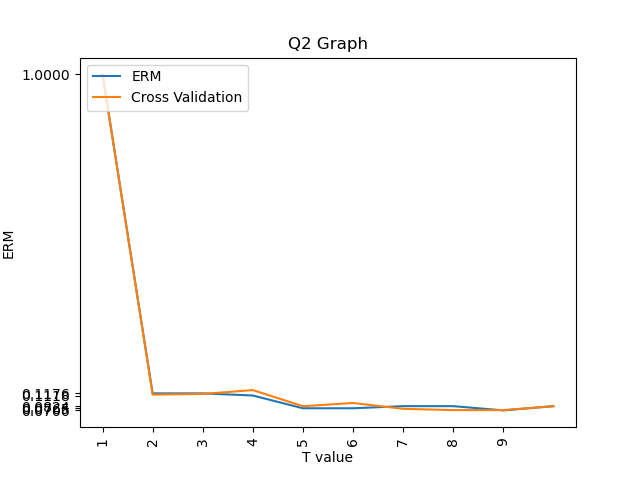
For linearly-separable-dataset

|  |  |  |
| --- | --- | --- |
|  | **Perceptron ERM only** | **Perceptron K-Fold Cross Validation** |
| **Number of Folds** | **Na** | **10** |
| **Number of Epochs** | **493** | **153** |
| **Average ERM value** | **0.0** | **0.0** |
| **Accuracy** | **1.0** | **1.0** |

****

For breast\_cancer\_data

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Perceptron ERM only** | **Perceptron K-Fold Cross Validation** | **AdaBoost ERM only** | **AdaBoost K-Fold Cross Validation** |
| **Number of Folds** | **Na** | **10** | **Na** | **10** |
| **Number of Epochs** | **1000** | **10000** | **5** | **5** |
| **ERM value** | **0.37** | **0.28** | **0.07** | **0.09** |
| **Accuracy**  **(1-ERM)** | **0.63** | **0.72** | **0.93** | **0.91** |



1. **Perceptron:**

* Perceptron is a binary classifier, in our case we run it for linearly separable data and the breast cancer dataset.
* In case of first dataset the number of epochs taken is constant as the data converges after 493 for erm only and 153 for perceptron with k-fold and produces an almost perfect line of separation as can be seen in the first graph on page 1, accuracy of 100% for both ERM only and K-fold cross validations.
* Next we run the perceptron for a data that is not perfectly linearly separable (breast cancer data) here the accuracy of perceptron with ERM only is 63% and for the K-Fold data validation is 72%.

1. **Adaptive Boosting:**

* Accuracy for adaptive boosting in case of ERM only on breast cancer data is 93% while for the k-fold cross validation is 91%
* Ideal number of classifiers found is 10
* Graph above demonstrates the comparison between the normal ERM and the ERM in case of cross validation across 10 folds.
* Increasing number of classifiers leads to increase in accuracy as well as takes more time.

**Readme:**

* To run perceptron for ERM only: python perceptron.py --data FILEPATH --mode erm
* To run perceptron for ERM with 10 folds: python perceptron.py --data FILEPATH --mode kfold
* To run AdaBoost for ERM only: python adaboost.py --dataset FILEPATH –mode erm
* To run perceptron for ERM with 10 folds: python adaboost.py --dataset FILEPATH –mode kfold
* To plot graph for AdaBoost comparison: python adaboost.py --dataset FILEPATH --mode analysis