

1. A very brief description of your model and its implementation.

The model we have implemented is the one detailed in Bishop's "Pattern Recognition and Machine Learning". The dataset iterates through the dataset for a maximum of 10^6 iterations (In the case that the dataset is not linearly separable, otherwise it will converge under the perceptron convergence theorem) and for each point incorrectly classified, the parameter vector (named "weights") according to the equation as shown below.

$$\mathbf{w}^{(\tau+1)} = \mathbf{w}^{(\tau)} - \eta \nabla E_P(\mathbf{w}) = \mathbf{w}^{(\tau)} + \eta \phi_n t_n$$

Where "eta" is the learning rate and t_n is the correct classification. If all the points are correctly classified, then the algorithm has converged to an exact solution. The prediction function is defined as.

```
def pred(row, weights):  
    activation = np.dot(row, weights)  
    return 1 if activation >= 0.0 else 0
```

Where it returns the prediction value, 1 or 0. The dataset is given a 70:30 split as mentioned in the assignment handout.

2. Accuracy of your model on both the datasets.

The training and testing accuracy of the datasets given "dataset_LP_1.txt" and "dataset_LP_2.csv" respectively are shown below.

```
Number of iterations : 1000000  
The training accuracy for dataset 1 is 98.64583333333333  
The test accuracy for dataset 1 is 99.27184466019418  
Number of iterations : 10  
The training accuracy for dataset 2 is 100.0  
The test accuracy for dataset 2 is 100.0
```

3. Dataset which was more linearly separable.

As you can see above, as "dataset_LP_2.csv" took significantly less iterations and achieved a testing and training accuracy of 100, the algorithm clearly converged on an exact solution, as compared to "dataset_LP_1.txt" data which implies that the former is more linearly separable according to the perceptron convergence theorem.

4. Major limitations of the Perceptron classifier.

The main limitation of this algorithm is that it can only be used confidently with datasets that are linearly separable, it should be noted however, that with multiple perceptrons, datasets and problems of increasing complexity can be solved.