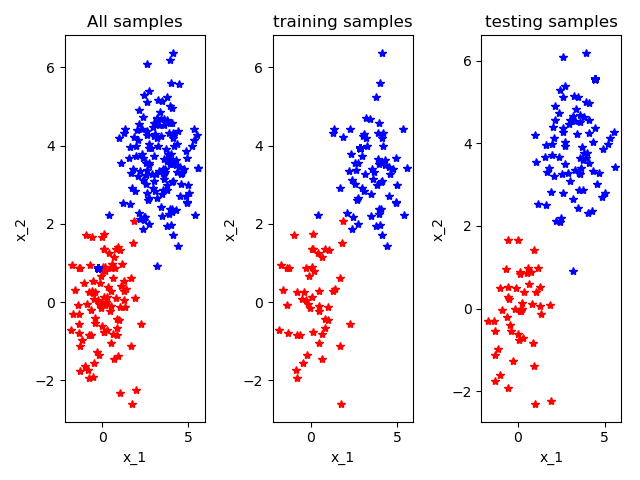
Matthew Hess

Red-Id: 818299658

Assignment 4

**Problem I.**

Here is a visual graph of my data being split:



Initial cost [0.68980433]

Final theta values:

[[0.04546512]

[0.04134964]]

cost [0.66180904]

shape of updated theta (2, 1)

trainX.shape (120, 2)

trainY.shape (120, 1)

testX.shape (130, 2)

testY.shape (130, 1)

yHat.shape (130, 1)

SK learn accuracy: 0.8692307692307693

SK learn true-positive precision: 0.8658536585365854

SK learn true-negative precision: 0.875

SK learn true-positive recall: 0.922077922077922

SK learn true-negative recall: 0.7924528301886793

SK learn confusion matrix:

[[71 6]

[11 42]]

sklearn average error: 0.13076923076923078 (0.33714780001247885)

score Sklearn: 0.8692307692307693

GD learn accuracy: 0.8

GD learn true-positive precision: 0.7475728155339806

GD learn true-negative precision: 1.0

GD learn true-positive recall: 1.0

GD learn true-negative recall: 0.5094339622641509

GD learn confusion matrix:

[[77 0]

[26 27]]

Own GD implementation average error: 0.2 (0.4)

score GD: 0.8

**Problem II.**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Y^ = Dog | Y^ = Cat | Y^ = Monkey |
| Y = Dog | 3 | 3 | 2 |
| Y = Cat | 3 | 1 | 1 |
| Y = Monkey | 2 | 2 | 3 |

True cat count: 5

True dog count: 8

True monkey count: 7

**Accuracy:**

(correctly predicted (a.k.a. diagonals) / all boxes or number of samples)

(3 + 1 + 3) / (3 + 3 + 2 + 3 + 1 + 1 + 2 + 2 + 3) == (7 / 20)

**Precision:**

For each class (# of correctly predicted / total # of predictions of that specific class (a.k.a. vertical count))

Cat: (1 / 6)

Dog: (3 / 8)

Monkey: (3 / 6)

**Recall Rates:**

For each class (# of correctly predicted / total # of ground truths of that class (a.k.a. horizontal count))

Cat: (1 / 5)

Dog: (3 / 8)

Monkey: (3 / 7)