**COMP 642 Assignment --- MODULE 3**

1. Perceptron
   1. Suppose you have a dataset with 4 data points shown below with each point having a label attached to it. Let the starting weight vector . Apply the perceptron learning algorithm for this dataset until convergence. For each step, write down the classification result and the resulting update of the weight vector. Assume learning rate is 0.1.

Dataset:

Hint:

Recall the process of perceptron algorithm:

* You are given
* You predict
* You are given
* Update the weight on the basis of of and
  1. If we change the dataset to contain the 4 data points shown below, can a single perceptron solve the problem? Why or why not? (Hint—work the problem using the beginning weight vector and learning rate from part a for 10 iterations and observe what happens). What kind of problems do you think can be solved by perceptron and what kind of problems can’t be solved by perceptron?

Dataset:

* 1. Run the hw\_3.ipynb and fill out the code cells related to perceptron. . Include the answers in your homework document submission as well as in the Jupyter notebook.

1. Adaline
   1. Gradient descent is an optimization algorithm used to minimize a function by iteratively moving in the direction of steepest descent as defined by the negative of the gradient.

The rule for Adaline to update the weights is in fact the stochastic gradient descent of linear regression. In this exercise you will derive gradient descent for Linear Regression cost function.

The notations for this problem:

Given the cost function for Linear Regression:

Given the update rule for Linear Regression:

The stochastic gradient descent computes the gradient of a single randomly chosen data point which means every individual observation will help to make an update on the parameter. Can you complete the intermediate steps for deriving the stochastic gradient descent for Linear Regression?

* 1. Run the hw\_3.ipynb and fill out the code cells related to Adaline. Include the answers in your homework document submission as well as in the Jupyter notebook.

1. Logistic Regression.
   1. Run the hw\_3.ipynb and fill out the code cells related to Logistic Regression. Include the answers in your homework document submission as well as in the Jupyter notebook.
   2. In hw\_3.ipynb provided, we only use “sepal length” and “petal length” as features for parametric models. Try out other feature combinations to train the model. Which combination do you think is the best one?
2. Compare your results of the Perceptron, Adaline, and Logistic Regression.