CS 789
INTELLIGENT SYSTEMS AND MACHINE LEARNING

SPRING 2018
Assignment #5

Due Date/Time: **04/02/2018 @ 11:59PM**
Total Points: **100**

This is an individual assignment; no teamwork is allowed.

For this assignment, you will implement logistic regression to model the provided dataset `fisherIrisVersicolorVirginicaData`, which is extracted from `fisherIrisData`.

- The dataset has 100 instances.
- It has three columns: column 1 is **sepal length** column 2 is **petal length**, and column 3 is **species**.
- Species 1 is ***Iris Versicolor*** and species 2 is ***Iris Virginica*****.**
- Logistic regression will be used to predict **species** based on **sepal length** and **petal length**.

Important: You are not allowed to use any built-in logistic regression functions (i.e., you need to implement your own algorithms.)

1) **(5 points)** Plot `fisherIrisVersicolorVirginicaData` dataset by setting `xlabel` and `ylabel` to `SepalLength` and `PetalLength` respectively. Submit your code and figure.

2) **(45 points) Offline learning:** Implement gradient descent to fit the logistic regression parameters to `fisherIrisVersicolorVirginicaData` dataset.
   - Use the first 80 instances in the dataset to train the model.
   - Keep track of the number of iterations it takes for gradient descent to converge.
   - Use `tic` and `toc` functions to measure and display run-time.
   - Use the last 20 instances in the dataset to test the model.
   - Submit your code.

3) **(45 points) Online learning:** Implement online gradient descent to fit the logistic regression parameters to `fisherIrisVersicolorVirginicaData` dataset.
   Important: Please note that this is not the perceptron algorithm.
   - Use the first 80 instances in the dataset to train the model.
   - Keep track of the number of iterations it takes for online gradient descent to converge.
   - Use `tic` and `toc` functions to measure and display run-time.
   - Use the last 20 instances in the dataset to test the model.
   - Submit your code.

4) **(5 points)**
   - Compare the learning rates used, accuracy rates, running times, and number of iterations of offline learning and online learning.

Subbmission Instructions:

Compress all the files and name the submission file **<YourLastName>_Assignment5**. For example, if your last name is Smith and you are submitting a .zip file, the file should be named Smith_Assignment5.zip.

IMPORTANT: WebCampus Maintenance Downtimes

The regular WebCampus maintenance period is every Saturday from 12:01am - 6am. This period is reserved for necessary updates and maintenance to the system. Users should not plan on accessing WebCampus during this time. **http://wchelp.unlv.edu/students/home**