Homework 2 **CSCE 421**
Due: 5.00 pm on October 3, 2019

Instructions for homework submission
a) For the **math problems**, please typewrite your answers in Latex, or handwrite your solution *very clearly*. Non-visible solutions will not be graded: we wouldn't like our TA to have to guess what you are writing :)
b) For the **experimental problems**, please write a brief report. At the end of the report, please include your code. Print the report, including the code.
c) **Staple all your answers and hand them out in paper in class or in the instructor's office.**
d) Please start early :)
e) The maximum grade for this homework, excluding bonus questions, is **10 points** (out of 100 total for the class).

Question 1 (5 points)
Linear Perceptron Algorithm: The goal of this problem is to run a linear perceptron algorithm. Assume that you have three training samples in the 2D space:

1. Sample $\mathbf{x_1}$ with coordinates $(1, 3)$ belonging to Class 1 ($y_1 = 1$)
2. Sample $\mathbf{x_2}$ with coordinates $(3, 2)$ belonging to Class 2 ($y_2 = -1$)
3. Sample $\mathbf{x_3}$ with coordinates $(4, 1)$ belonging to Class 2 ($y_2 = -1$)

The linear perceptron is initialized with a line with corresponding weight $\mathbf{w(0)} = [2, -1, 1]^T$, or else the line $2 - x + y = 0$.
In contrast to the example that we have done in class, in this problem **we will include the intercept** w_0 **or** x_0**.**

(0.5 points) (i) Plot $\mathbf{x_1}$, $\mathbf{x_2}$, and $\mathbf{x_3}$ in the given 2D space. Plot the line corresponding to weight $\mathbf{w(0)}$, as well as the direction of the weight $\mathbf{w(0)}$ on the line.

(0.5 points) (ii) Using the rule $sign(\mathbf{w(t)}^T\mathbf{x_n})$, please indicate the class in which samples $\mathbf{x_1}$, $\mathbf{x_2}$, and $\mathbf{x_3}$ are classified using the weight $\mathbf{w(0)}$. Which samples are not correctly classified based on this rule?
Note: You have to compute the inner product $\mathbf{w(0)}^T\mathbf{x_n}$, $n = 1, 2, 3$, and see if it is greater or less than 0.

(1.5 points) (iii) Using the weight update rule from the linear perceptron algorithm, please find the value of the new weight $\mathbf{w(1)}$ based on the misclassified sample from question **(ii)**. Find and plot the new line corresponding to weight $\mathbf{w(1)}$ in the 2D space, as well as the direction of the weight $\mathbf{w(0)}$ on the line. Indicate which samples are correctly classified and which samples are not correctly classified.
Note: The update rule is $\mathbf{w(t+1)} = \mathbf{w(t)} + y_s\mathbf{x_s}$, where $\mathbf{x_s}$ and $y_s \in \{-1, 1\}$ is the feature and class label of misclassified sample s.
Hint: The line corresponding to a vector $\mathbf{w} = [w_0, w_1, w_2]$ can be written as $w_0 + w_1 x + w_2 y = 0$. Make sure that you get the direction of the vector $\mathbf{w}$ correctly based on the sign of w_1 and w_2.

(2.5 points) (iv) Using the rule $sign(\mathbf{w(t)}^T\mathbf{x_n})$, run the linear perceptron algorithm until it converges, find and plot the weights $\mathbf{w(2)}, \mathbf{w(3)}, \ldots$ and the corresponding lines in each iteration.

For each iteration, please indicate which samples are classified correctly and which samples are not classified correctly.
Hint: In order to make the linear perceptron algorithm converge as fast as possible, you can always update the weight based on sample $\mathbf{x_1}$. Why?

Question 2 (5 points)
Predicting sound pressure in NASA wind tunnel: We would like to predict the sound pressure in an anechoic wind tunnel. This can help NASA design ground-based wind tunnels in order to assess the effect of noise during spaceflight. We use data from the Airfoil Self-Noise Data Set in the following UCI Machine Learning Repository: `https://archive.ics.uci.edu/ml/datasets/Airfoil+Self-Noise#`.
Inside "Homework 2" folder on Piazza you can find two files including the train and test data (named "hw1_ question1_train.csv" and "hw1_ question1_test.csv") for our experiments. The rows of these files refer to the data samples, while the columns denote the features (columns 1-5) and the sound pressure output (column 6), as described bellow:

1. Frequency, in Hertz
2. Angle of attack, in degrees
3. Chord length, in meters
4. Free-stream velocity, in meters per second
5. Suction side displacement thickness, in meters
6. Scaled sound pressure level, in decibels (this is the **outcome**)

(i) (1.5 point) Data exploration: Using the training data, plot the histogram of each feature and the output (i.e., 6 total histograms). How are the features and the output distributed?

(ii) (2 points) Using the train data, **implement** a linear regression model using the ordinary least squares (OLS) solution. How many parameters does this model have?
Hint: You will build the data matrix $\mathbf{X} \in \mathcal{R}^{N_{train} \times 6}$, whose rows correspond to the training samples $\mathbf{x_1}, \ldots, \mathbf{x_{N_{train}}} \in \mathcal{R}^{5 \times 1}$ and columns to the features (including the constant 1 for the intercept): $\mathbf{X} = \begin{bmatrix} 1, \; \mathbf{x_1}^T \\ \vdots \\ 1, \; \mathbf{x_N}^T \end{bmatrix} \in \mathcal{R}^{N_{train} \times 6}$. Then use the ordinary least squares solution that we learned in class: $\mathbf{w}^* = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$.
Note: You can use libraries for matrix operations.

(iii) (1.5 points) Test your model on the test data and compute the residual sum of squares error (RSS) between the actual and predicted outcome variable.

(iv) (Bonus, 2 points) Experiment with different feature combinations and report your findings. What do you observe?