CSEN 240
Homework #5

1. (10 points) Team project
    a. Please form teams with 2-3 students
    b. Create a group in the Student Groups under Camino > People > Groups
    c. Attach a screenshot of the group

[Level 1: high-level understanding]

2. (10 points) Which of the following are true? (Check all apply)
    a. Linear regression predicts a continuous output variable based on one or more input variables.
    b. A goal of linear regression is to find the line of best fit that minimizes the sum of squared errors between the predicted values and the actual values.
    c. Polynomial regression is a machine learning technique used to model non-linear relationships between the input features and the target variable.
    d. Polynomial regression can suffer from overfitting when the degree of the polynomial is too high.
    e. Regularized polynomial regression involves adding a penalty term to the cost function, which reduces the magnitude of the polynomial coefficients and prevents overfitting.
3. (10 points) Which of the following are true? (Check all apply)
    a. Gradient descent is an optimization algorithm that can be used to minimize the cost or loss function of a machine learning model. It is very useful for the optimization problem does not have a close-form solution.
    b. The learning rate determines the size of the step taken during each iteration of gradient descent.
    c. Large learning rate is always better than small learning rate.
    d. Gradient descent can only be used to optimize linear models.
    e. Gradient descent always converges to the global minimum of the loss function.
4. (10 points) Please read about **Lasso Regression** by yourself. Ref: Aurélien Géron, "Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow," O'Reilly Media, Inc., Chapter 4. Which of the following statements are true about Lasso Regression? (Select all that apply)
    a. Lasso Regression uses L1 regularization. That is, the penalty term is the absolute value of the coefficients.
    b. Lasso Regression can be used for reducing overfitting.
    c. An important characteristic of Lasso Regression is that it tends to eliminate the weights of the least important features (i.e., set them to zero).
    d. Lasso Regression automatically performs feature selection and outputs a sparse model with few nonzero feature weights.
    e. The Lasso cost function is not differentiable at coefficients = 0.

[Level 2: manual exercise]

5. (20 points) Linear regression using closed-form solution
    a. Given the following dataset (x, y) ∈ { (1, 2), (3, 3), (4, 4) }
    b. Please use least squared errors to find the best linear regression model, using the closed-form solution
    c. Please use the linear regression model to predict the output y, given input x=2
6. (20 points) Linear regression using gradient descent
    a. Given the following dataset (x, y) ∈ { (1, 2), (3, 3), (4, 4) }
    b. The initial weight of the linear regression model is (0, 1)
    c. Please gradient descent (once) to update the weight, using sum of squared errors as the error function ($E(\boldsymbol{w}) = \sum_{i=0}^{N}(w_0 + w_1 x_i - y_i)^2$), and learning rate = 0.1
    d. Please use the updated linear regression model to predict the output y, given input x=2

[Level 3: computer-based exercise]

7. (20 points) Linear regression using gradient descent (Using HW5-7.ipynb)
    a. Given the house price dataset, you will train a linear regression model using gradient descent (with sum of squared error, $E(\boldsymbol{w}) = \sum_{i=0}^{N}(w_0 + w_1 x_i - y_i)^2$, as the cost function) to predict the house price given the square footage of the house.
    b. Read the dataset. Choose the appropriate fields to create a 1D training set
    c. Normalize the house size using the min-max scaling
    d. Train the model with gradient descent
    e. Plot the line
    f. Print the weight vector
    g. Predict the price of the house if the size is 4000 square feet