Pattern Recognition, Homework 1

Deadline: March 31, 23:55

Part. 1, Coding (70%):

In this coding assignment, you are required to implement linear regression by using only NumPy, then train your model using **gradient descent** on the provided dataset, and evaluate the performance on testing data. Find the sample code and data on the GitHub page https://github.com/NCTU-VRDL/CS_ILE5065/tree/main/HW1

Please note that only NumPy can be used to implement your model, you will get 0 points by calling sklearn.linear_model.LinearRegression. Moreover, please train your regression model using gradient descent, not the closed-form solution.

1. (15%) Implement the linear regression model and train it by using **gradient descent** with **mean absolute error** and **mean square error** as the objective function, respectively
2. (15%) Plot the learning curve of the training with both losses in the same figure, you should find that loss decreases and converges after a few iterations (x-axis=iteration, y-axis=loss, Matplotlib or other plot tools is available to use)
3. (15%) What're the mean square error and mean absolute error between your predictions and the ground truths on the testing data (prediction=model(x_test), ground truth=y_test)
4. (10%) What're the weights (β_1) and intercepts (β_0) of your linear model trained from both losses?
5. (10%) What's the difference between gradient descent, mini-batch gradient descent, and stochastic gradient descent?
6. (5%) All your codes should follow the PEP8 coding style and with clear comments

Part. 2, Questions (30%):

1. (10%) Suppose that we have three colored boxes R (red), B (blue), and G (green). Box R contains 3 apples, 4 oranges, and 3 guavas, box B contains 2 apples, 0 orange, and 2 guavas, and box G contains 12 apples, 4 oranges, and 4 guavas. If a box is chosen at random with probabilities p(R)=0.2, p(B)=0.4, p(G)=0.4, and a piece of fruit is removed from the box (with equal probability of selecting any of the items in the box), then what is the probability of selecting guava? If we observe that the selected fruit is in fact an apple, what is the probability that it came from the blue box?

2. (10%) Using the definition $var[f] = E[(f(x) - E[f(x)])^2]$ show that $var[f(x)]$ satisfies $var[f] = E[f(x)^2] - E[f(x)]^2$.

3. (10%) Consider two variables x and y with joint distribution $p(x, y)$. Prove the following result

$$E[x] = E_y[E_x[x|y]]$$

Here $E_x[x|y]$ denotes the expectation of x under the conditional $p(x|y)$, with a similar notation for the conditional variance.

Hint: Please check the definitions of the expectation operator, the sum rule, and the product rule.