Assignment 2

Your Name (your student number)
abc@yorku.ca

February 4, 2021

Note: You have to work individually. You must use the same mathematical notations in textbook or lecture slides to answer these questions. You must use this latex template to write up your solutions. Remember to fill in your information (name, student number, email) at above. No handwriting is accepted. In this assignment, you need to use the *MNIST* data set. Refer to

`https://colab.research.google.com/drive/1FyahMGAE22716sUCrNXPvTrKTwS615Hd`

for how to load it in Python. Direct your queries to Hui Jiang (`hj@eecs.yorku.ca`)

Exercise 1

Dimension Reduction

1. (5 marks) **PCA**: Q4.2 on page 93
2. (5 marks) **LDA**: Q4.4 on page 93
3. (10 marks) **Data visualization**: Lab Project I on page 92, parts a), b) and c)
   Note that you will have to implement PCA and LDA from scratch but you may choose to use a t-SNE implementation from any Python package.

Your answers:

Exercise 2

Linear Models for Regression

1. (10 marks) derive the formula to compute the gradients for the following linear models:
   (a) linear regression
   (b) ridge regression
   (c) LASSO

   follow the style of **Algorithm 2.3** (refer to `https://www.overleaf.com/learn/latex/algorithms`) to derive mini-batch stochastic gradient descent algorithms to optimize these models.
2. (20 marks) implement these three algorithms on a small data set, e.g. the Boston Housing Dataset (`https://scikit-learn.org/stable/modules/generated/sklearn.datasets.load_boston.html`), to predict median value of a home from the 13 attributes. You need to experimentally compare these regression models, discuss your results in terms of how the learning objective function and all learned model weights may differ among three models.

Your answers:

Exercise 3

Support Vector Machine (SVM)

1. (10 marks) Q6.8 on page 130
2. (20 marks) use all training data of two digits 5 and 8 from the MNIST dataset to learn two binary classifiers using linear SVM and nonlinear SVM (with Gaussian RBF kernel), and compare and discuss the performance and efficiency of linear SVM and nonlinear SVM methods for these two digits. Report your best results in the test data of 5 and 8. Don't call any off-the-shelf optimizer. Use the projected gradient descent in Algorithm 6.5 to implement the SVM optimizer yourself.

Your answers:

What to submit?

You must submit:

1. one PDF document (using this latex template) for your solutions to all written questions and all results and discussions for your programming assignments
2. one zip file that includes all of your Python codes and a readme file for TA to run your codes

from eClass before the deadline. No late submission will be accepted.