Hand-on Exercise 2

Guidelines

(1) Each group consists of at most two students. Members in the same team will receive the same grade. (Please give your group member names to the TA.)

(2) The deadline for turning in your report is March 12 (Thursday), 1:00 PM.

Assignment

(1) Each team's task is to unveil the embedded information of two given datasets.

(2) **TWO datasets** are posted on the course page.

(3) Each team may use MATLAB or Python as the development platform.

(4) The performance will be judged by the quality of solutions obtained, the associated analysis, and discussions.

(5) For each dataset, perform the following tasks:

(i) LSSVM: Use the LSSVM model (page 26 of Lecture 5) to classify the data contained in "EX2dataset.zip", for $C = 1, 10, 100$.

(a) (16 pts) Find the classification hyperplane, optimal objective value, and computational time for each parameter C.

(b) (8 pts) Visualize the separation hyperplane with all data points for each parameter C.

(c) (4 pts) Design a decision rule for choosing the "best" parameter C. Clearly explain the logic of your decision rule.

(ii) DKSSVM-Gaussian: Use the DKSSVM model (page 39 of Lecture 5) and the Gaussian kernel (page 43 of Lecture 5) with $\sigma = 1, 10, 100$, to classify the data contained in "EX2dataset.zip", for $C = 1, 10, 100$.

(a) (16 pts) Find the optimal objective value and computational time for each combination of (C, σ).

(b) (8 pts) Visualize the separation surface with all data points for each combination of (C, σ).

(c) (4 pts) Design a decision rule for choosing the "best" combination of (C, σ). Clearly explain the logic of your decision rule.

(iii) DKSSVM-Polynomial: Use the DKSSVM model (page 39 of Lecture 5) and the Polynomial kernel (page 41 of Lecture 5) with $r = 1$ and $d = 3$, to classify the data contained in "EX2dataset.zip", for $C = 1, 10, 100$.

(a) (16 pts) Find the optimal objective value and computational time for each parameter C.

(b) (8 pts) Visualize the separation surface with all data points for each parameter C.

(c) (4 pts) Design a decision rule for choosing the "best" parameter C. Clearly explain the logic of your decision rule.

(iv) Analysis and Discussion:

(a) (8 pts) Compare your results of (i), (ii), and (iii), which model and parameters would you like to use for this dataset? Why?

(b) (8 pts) Analyze your findings and discuss what you have learned from doing this hand-on exercise..

Report Format

(1) Title and Team Members

(2) Background and Objectives

(3) LSSVM Work and Results

(4) DKSSVM-Gaussian Work and Results

(5) DKSSVM-Polynomial Work and Results

(6) Graphic Display of the Best Results of Each Case

(7) Analysis and Discussions

(8) Attachment of Computer Program/Codes, Inputs and Outputs