**ISE 789/OR 791 Exercise 2**

**Issued: November 14, 2024**
**Due: December 3, 12:00 PM NOON**

**Rules and Reports: Same as in Exercise 1**

**Total points = 110**

Given 4 data sets, ``data1.csv", ``data2.csv", ``data3.csv" and ``data4.csv" (provided on the course webpage), use various SVM/SVR models we learned in the class to investigate the following binary classification and linear regression problems.

**Problem 1 (25 points)**
Consider the data set data1.csv, which is linearly separable.

a) (5 points) Use the LSVM model (page 9 of Lecture 9) to classify the data contained in data1.csv. Find the classification hyperplane H1 and its supporting vectors. Also, record the computational time T1.
b) (5 points) Use the DLSVM model (page 15 of Lecture 9) to classify the data contained in data1.csv. Find the classification hyperplane H2 and its supporting vectors. Also, record the computational time T2.
c) (5 points) Visualize H1 and H2 with all data points and highlight the corresponding supporting vectors. Are H1 and H2 the same? Why? Are those supporting vectors the same? Why? Are T1 and T2 the same? Why? If the results are not the same, why are they different? Which model do you prefer to use in this case? Why?
d) (5 points) Use the LSSVM model (page 22 of Lecture 9) with parameter $C = 100$ to classify the data contained in data1.csv. Find the classification hyperplane H3 and its supporting vectors. Also, record the computational time T3.
e) (5 points) Visualize H1 and H3 with all data points and highlight the corresponding supporting vectors. Are H1 and H3 the same? Why? Are the corresponding supporting vectors the same? Why? Arre T1 and T3 the same? Why? If the results are not the same, why are they different? Which model do you prefer to use in this case? Why?

**Problem 2 (30 points)**
Consider the data set data2.csv, which is not linearly separable.

a) (5 points) Use the LSVM model (page 9 of Lecture 9) to classify the data contained in data2.csv. Find the classification hyperplane H4 and its supporting vectors. Also, record the computational time T4. If you cannot find H4, please explain the situation.
b) (10 points) Use the LSSVM model (page 22 of Lecture 9) to classify the data contained in data2.csv, for $C$ = 0, 1, 10, 100, 10000. Find the classification hyperplane

H5, optimal objective value, and computational time in each case. Among those obtained H5's, which one is the best for your choice? Why?

c) (10 points) Use the DLSSVM model (page 27 of Lecture 9) to classify the data contained in data2.csv, for $C$ = 0, 1, 10, 100, 10000. Find the classification hyperplane H6, optimal objective value, and computational time in each case. Among those obtained H6's, which one is the best for your choice? Why?

d) (5 points) Visualize the best H5 and H6 with all data points. Are H5 and H6 the same? Why? If not so, why are they different? Which model do you prefer to use in this case? Why?

**Problem 3 (35 points)**

Consider the data set data3.csv, which is totally not linearly separable.

a) (10 points) Use the LSSVM model (page 22 of Lecture 9) to classify the data contained in data3.csv, for $C$ = 1, 10, 100. Find the classification hyperplane H7, optimal objective value, and computational time in each case. Among those obtained H7's, which one is the best? Why? Visualize the best H7 with all data points.

b) (10 points) Use the DKSSVM model (page 36 of Lecture 9) and the Gaussian kernel (page 40 of Lecture 9) with $\sigma = 1, 10, 100$, to classify the data contained in data3.csv, for $C$ = 1, 10, 100. Find the optimal objective value and computational time in each case. Which one is the best? Why? Visualize the best separation surface S1 with all data points.

c) (10 points) Use the DKSSVM model (page 36 of Lecture 9) and the Polynomial kernel (page 38 of Lecture 9) with $r = 1\ and\ d = 3$, to classify the data contained in data3.csv, for $C$ = 1, 10, 100. Find the optimal objective value and computational time in each case. Which one is the best? Why? Visualize the best separation surface S2 with all data points.

d) (5 points) Compare the results of (a), (b) and (c), which model and parameters you would like to use for this case? Why?

**Problem 4 (20 points)**

Consider the data set data4.csv.

a) (5 point) Use the least-squares-estimation linear regression method (page 14 of Lecture 4) to find the hidden linear relationship R1 between the input and output variables. Record the computational time and compute the MSE (mean squared errors).

b) (5 points) Use the LSSVR model (page 52 of Lecture 9) with parameters $C = 1\ and\ \varepsilon = 3$ to find the hidden linear relationship R2 with optimal $\boldsymbol{w}, b\ and\ \boldsymbol{\zeta}$. Record the computational time and compute the corresponding MSE.

c) (5 points) Visualize R1 and R2 with all data points and the tubes with radius $\varepsilon = 3$.

d) (5 points) Compare the MSE value, the number of data points outside the tube, and the computational time of R1 and R2. What are the advantages and disadvantages of the regular linear regression model and support vector regression model?