AUS | الجامعة الأميركية في الشارقة
American University of Sharjah

American University of Sharjah
School of Engineering
Department of Computer Engineering
P. O. Box 26666
Sharjah, UAE

Instructor: Alex Aklson
Office: ESB-2172
Phone: 971-6-515 4893
E-mail: aaklson@aus.edu
Semester: Fall 2024

MLR503 – Data Mining and Knowledge Discovery
Assignment 3
Due: November 17, 2024 by 11:59PM

Question (Support Vector Machines):

a) Generate a scatter plot of the data in "customers_data.csv". Color code the data points to visualize the datapoints belonging to each class.

b) Preprocess the data and build an SVM classifier with very large C value. What type of margin classifier is this?

c) Plot the decision boundary of the resulting SVM model and show the margins and the resulting support vectors. How many support vectors are there?

d) Evaluate the classifier on the "customers_data_test.csv". Report the value(s) of the appropriate evaluation metric(s).

e) Use the "customers_data_full.csv" dataset to build an SVM classifier Set the value of C to be a very large value.

f) Discuss any issues you run into when building an SVM classifier using "customers_data_full.csv" and a large value for C. Explain what could be causing those issues.

g) Build an SVM classifier that is suitable for the data in "customers_data_full.csv". What type of margin classifier is this?

h) Choose the value of C that results in the best performance.

i) Plot the decision boundary of the resulting SVM model and show the margins and the resulting support vectors. How many support vectors are there?