

מבוא ללמידת מכונה

תרגיל 4

תאריך הגשה: באתר הקורס

שאלה 1:

בשאלה זו תעבדו על הדאטהסט שבקובץ `simple_classification.csv` (נמצא באתר).

א. השתמשו **בתכנית הריבועית של SVM** מהשיעור והריצו על הדאטהסט.

- הדפיסו את ווקטור המשקלות (w) המתקבל.
- ציירו (plot) את הדאטה עם המפריד הלינארי.
- ציירו גם את קווי השוליים.

ב. השתמשו **בתכנית הדואלית של SVM** מהשיעור והריצו על הדאטהסט.

- הדפיסו את ווקטור המשקלות (w) המתקבל.
- ציירו (plot) את הדאטה עם המפריד הלינארי.
- ציירו גם את קווי השוליים.
- סמנו את ה-support vectors.

שאלה 2:

בשאלה זו תעבדו על הדאטהסט שבקובץ `simple_nonlin_classification.csv` (נמצא באתר).

השתמשו **בתכנית הדואלית של SVM עם שימוש ב-kernel** והריצו על הדאטהסט. השתמשו בקרנלים שונים עם פרמטרים שונים. למשל קרנל פולינומי עם מעלות שונות ו/או קרנל RBF עם ערכים שונים למקדם γ .

השוו את התוצאות עבור המודלים השונים (בהתאם לקרנל ובחירת הפרמטרים). כדי להשוות חלקו את הדאטה ל-`train` ו-`test`, אמנו על ה-`train` ובחנו על ה-`test`. ציירו גרף של הטעות על ה-`test` לפי המודלים השונים.

בנוסף, עבור מודלים נבחרים:

- ציירו (plot) את הדאטה עם המפריד (הלא לינארי).
- ציירו גם את קווי השוליים.
- סמנו את ה-support vectors.

שאלה 3:

ממשו מחלקה SVM עם המתודות הבאות:

- `__init__(kernel, degree, C, gamma)`
- `fit(X, y)`
- `predict(X)`
- `decision_function(X)`
- `score(X, y)`

ההתנהגות צריכה להיות דומה למחלקה SVC במודול ה-`svm` של `sklearn`, ראו [sklearn.svm.SVC](https://sklearn.org/stable/modules/generated/sklearn.svm.SVC.html).

אין צורך לממש את כל הפונקציונליות של מחלקה זו, רק את החלק שפירטנו כאן. ואין גם צורך לאפשר את כל הפרמטרים (ב-__init__), רק את הפרמטרים שציינו.

שאלה 4:

בשאלה זו תעבדו עם הדאטהסט **Processed Wisconsin Diagnostic Breast Cancer.csv** שכבר הכרתם בתרגיל הראשון.

הריצו SVM על הדאטהסט. השתמשו בקרנלים שונים עם פרמטרים שונים. השוו את התוצאות עבור המודלים השונים (בהתאם לקרנל ובחירת הפרמטרים). כדי להשוות חלקו את הדאטה ל-train ו-test, אמנו על ה-train ובחנו על ה-test. ציירו גרף של הטעות על ה-test לפי המודלים השונים.