



תרגיל מעבדה 4

את שאלות 1-3 יש לבצע **ללא** נורמליזציה:

1. כתבו פונקציה בקובץ main היוצרת את המסווג 1-NN עם כל נתוני הסט ולאחר מכן מסווגת את כל נתוני הסט. דווחו accuracy. הסבירו מה קיבלתם ולמה.
2. כתבו פונקציה בקובץ main היוצרת בלולאה מסווגים עם k שכנים כאשר $1 \leq k \leq 30$. לכל מסווג הפונקציה מריצה leave-one-out cross-validation ומדווחת תוצאות accuracy. מיהו המסווג הטוב ביותר על סמך התוצאות?
3. בפונקציה המבצעת cross-validation הוסיפו הדפסה של איכות ה testing עבור כל אחד מה testing folds. עבור המסווג הנבחר בשאלה הקודמת הריצו:
 - (a) 2-fold-cross-validation והדפיסו את הפלט
 - (b) 10-fold-cross-validation והדפיסו את הפלט
 - (c) 20-fold-cross-validation והדפיסו את הפלטהיכן התקבלו ביצועים עקביים (בעלי שונות נמוכה) על החלוקות השונות (folds)? מהי כמות ה folds שהייתם ממליצים עליהם בהינתן שאתם עובדים עם קובץ נתונים קטן יחסית, ומדוע?
4. כתבו פונקציה בקובץ main המריצה 2-fold-cross-validation עבור 5-NN ו 7-NN מחשבת ומדפיסה את הפלט של תוצאת ה Accuracy:
 - (a) ללא נרמול
 - (b) עם נרמול sum (L1 norm)
 - (c) עם נרמול min-max
 - (d) עם נרמול Z-Normמהי המסקנה שלכם לגבי התוצאות?

מצורף פלט לדוגמה.

חשוב לציין ששאלות 1-2 צריך לממש ולדווח תשובות, אבל הם לא צריכים לרוץ ולהדפיס תוצאה בתוכנית סופית. כלומר, הקוד שלכם אמור להדפיס רק את התשובות עבור שאלות 3-4.



דגשים נוספים :

- (1) ניתן להשתמש בקוד שכבר מימשתם בתרגילי הבית.
- (2) ניתן להוסיף פונקציות נוספות, במידה ותמצאו לנכון.
- (3) עליכם לכתוב את הקוד בהתאם לדרישות של PEP8. יותר על זה [כאן](#) או [כאן](#). אתם לא חייבים לעמוד בכל הדרישות של PEP, אבל יש כמה חשובות לנו. יותר על זה במסמך במודל.

- אורך שורה לא יותר מ-120
- שמות של משתנים\טיפוסים\קבצים
- Imports
- Blank lines

(4) יש לתת שמות בעלי משמעות לכל משתנה.

(5) חובה לתעד את הקוד באנגלית.

הוראות הגשה

- התרגיל להגשה בזוגות או ביחידים (נדרש אישור להגשה ביחידים).
- לפני ההגשה, חובה לוודא שהתוכנית עובדת במעבדת ההוראה ולא בסביבה אחרת.
- ההגשה חייבת להכיל קובץ אחד (קובץ zip) :
 1. שם הקובץ חייב להיות hw4_lab_xxxxxxxxx_yyyyyyyyyy.zip כאשר xxxxxxxx ו-yyyyyyyyy הם מספרי תעודות הזהות של המגישים, כולל ספרת ביקורת. הקובץ מכיל את כל קבצי הקוד וקובץ דו"ח שלכם עם תשובות לשאלות. אין להכיל תיקייה ובתוכה קבצי הקוד, אלא את קבצי הקוד עצמם.**הערה :** עליכם לוודא שהתוכנית מתחילה לפעול מקובץ "main.py" בלבד.
- ההגשה היא אלקטרונית בלבד, דרך אתר ה-moodle של הקורס. תרגילים שיוגשו בכל דרך אחרת לא ייבדקו.
- אין להגיש את אותו הקובץ פעמיים. התרגיל יוגש ע"י אחד מבני הזוג.
- שימו לב שההגשה תיחסם בדיוק בשעה 23:55 ביום ההגשה. מומלץ להגיש לפחות שעה לפני המועד האחרון.
- ניתן להגיש כמה פעמים. רק ההגשה האחרונה תישמר.
- תרגיל בית שלא יוגש לפי הוראות ההגשה – לא ייבדק (כלומר יקבל ציון 0).
- לצורך תרגיל הבית יפתח פורום. ניהול שאלות ומתן תשובות בנושא התרגיל יתבצע דרך הפורום בלבד.

בהצלחה!