

## למידת מכונה תרגיל 2

### המשותף לכל האלגוריתמים:

- פיצלנו את קובץ הtrain שקיבלנו ל-80% train ו-20% cross validation, כך שלאחר שהאלגוריתם למד את ה-80% הרצנו אותו גם על ה-20% בכדי לוודא שאחוז השגיאה דומה בין שניהם (כלומר, ווידאו שהאלגוריתם שלנו לא למד "לשנן" את ה-training set אלא באמת למד לסווג כל data מתאימה שהוא יקבל). ביצענו זאת 5 פעמים - כל פעם על 80% ו-20% אחרים.
- אתחלנו את מטריצת המשקולות  $w$  להיות בגודל  $3 \times 8$  וכולה אפסים. גודל זה הוא מספר הפיצורים (משקל, מין, אורך וכו') כפול מספר הסיווגים האפשריים (גיל ה abalone האפשרי: 0, 1, 2)
- בכל איטרציה של הרצה של סט הדוגמאות אנחנו מערבבות את הדוגמאות על מנת שהאלגוריתם לא ילמד את הסדר של הדוגמאות ו"ישנן" אותן, אלא כל פעם יאמן אותו לפי סדר שונה.
- בדקנו עבור כל האלגוריתמים כי ממוצע השגיאות אינו עולה על 40%. (למעשה, מלבד ב PA הוא היה אף טוב יותר), באלגוריתם PA ממוצע השגיאות הגיע לעיתים ל-45% לדוגמה:

```
preceptron err = 0.315271435819381  
svm err = 0.3088280060882801  
pa err = 0.45568657284138453
```

- עבור כל האלגוריתמים, כאשר נרמלנו את הערכים של הtrain set ע"י פונקציית zscore המובנית בפיתון וכן כאשר נרמלנו אותם ע"י פונקציית שאנחנו כתבנו (normal\_x\_train), ערך השגיאה יצא גבוה יותר ולכן הערכים שקיבלנו יצאו פחות מדויקים. לכן החלטנו שלא לנרמל את ה data שקיבלנו. (הפונקציות נשארו בקוד בתור הערות)

### :Perceptron

באלגוריתם זה היפר-הפרמטרים שלנו הם:

- eta: לצורך כלל העדכון, אתחלנו את eta להיות 0.1 ובכל איטרציה אנו מקטינות את הערך שלה. הבנו שהערך של eta צריך להיות קטן וכדי להגיע לתוצאה מדויקת יותר, הערך שלה צריך להיות קטן יותר בכל איטרציה. גם בעדכון של eta ניסינו להגדיל ולהקטין את הערכים בהם חילקנו את eta, התחלנו ב1, הגדלנו ל10, 100 ואף הקטנו ל0.1, 0.001 עד שמצאנו את האופטימלי כאשר נחלק את eta ב-100 e+100 ראינו זאת ע"י הרצה של האלגוריתם עם ערכי eta שונים. ניסינו בהתחלה עם eta=1, הגדלנו והקטנו את הערכים, עד שמצאנו את הערך שהוא האופטימלי.
- epochs: מס' הפעמים שהאלגוריתם מריץ את הדוגמאות. הרצנו את האלגוריתם מס' פעמים שונה עם קבצים שונים ובדקנו אחרי כמה איטרציות ה w לא משתנה ואחוז השגיאה מתכנס למספר קבוע. הבחנו כי האלגוריתם מבצע את מה שרצינו לאחר 8-13 איטרציות ולכן, כדי להיות בטוחות שהוא יתכנס ויבצע את מה שאנחנו רוצות, קבענו את מס' האיטרציות להיות 15.

### :SVM

באלגוריתם זה היפר-הפרמטרים שלנו הם :

- eta : נקבע כפי שקבענו באלגוריתם הקודם, ולכן אותחל להיות 0.1 עם אותו כלל עדכון.
- epochs : נקבע באופן דומה לאופן שקבענו באלגוריתם הקודם, רק שבאלגוריתם זה, ראינו כי האלגוריתם מתכנס אחרי 5-7 איטרציות ולכן קבענו את מס' האיטרציות להיות 10.
- lamda : על מנת לקבוע את הערך של lamda ניסינו בהתחלה עם ערכים גדולים מ-1 (5, 10, 50) וראינו שככל שהמספר גדל ערך השגיאה עולה. אז ניסינו עם ערכים קטנים מ-1 וראינו שהערך השגיאה קטן. ניסינו ערכים קטנים (0.1, 0.001, 0.01 וכו') עד שמצאנו את הערך שהביא לנו את התוצאה האופטימלית : 0.0001.

### :PA

באלגוריתם זה היפר הפרמטרים שלנו הם :

- epochs : האלגוריתם לא מתכנס, לא משנה כמה פעמים נריץ אותו. לכן קבענו מספר גדול מספיק של הרצות אך בלי לבזבז יותר מידי זמן ריצה ולכן קבענו אותו להיות 30.
- tau : הערך של tau נקבע עפ"י מה שקיבלנו בהוראות התרגיל. בכל איטרציה, חישבנו את ה hinge loss לפי ערכי  $x, y, w$  המתאימים, וכן את המכפלה של 2 כפול הנורמה של  $x$  בריבוע וכך עדכנו את ערך tau.