

למידת מכונה - תרגיל 2

בעיבוד המקדים לנתונים התבצע שינוי של התכונה השמית, כאשר במקום W הוכנס 0 ובמקום R הוכנס 1. זה היה מספק ו-one hot encoding לא נתן ביצועים טובים יותר. בנוסף, לכל התכונות בוצעה נורמליזציית מינימום-מקסימום בין 0 ל-1 כאשר שאר שיטות הנורמליזציה נתנו ביצועים דומים, הוחלט להשאיר את נורמליזציית מינימום-מקסימום האינטואיטיבית.

לגבי אתחול המשקולות בשני האלגוריתמים הרלוונטיים, איתחול משקולות אקראי בין 0 ל-1 תיפקד טוב יותר מאשר איתחול עם משקולות שהן 0.

KNN

על מנת לבחור מהו ה-K אשר נותן את התוצאות המיטביות בוצע KFOLD עם K של 5 ו-10. כיוון שקבוצת האימון הנתונה לא גדולה במיוחד, המיקוד היה יותר עם FOLD של 5. FOLD של 10 יוצר קבוצות ולידציה קטנות למדי ויכולות אולי לא מספיק לייצג.

בוצעה בדיקה לביצועי דיוק לכל K בטווח של 1 עד 19. עבור 5 FOLDS להלן התוצאות:

K = 1	0.771830985915493
K = 2	0.7295774647887323
K = 3	0.7577464788732394
K = 4	0.7830985915492958
K = 5	0.7830985915492958
K = 6	0.7774647887323944
K = 7	0.7943661971830986
K = 8	0.7971830985915492
K = 9	0.7830985915492958
K = 10	0.771830985915493
K = 11	0.7661971830985914
K = 12	0.7718309859154929
K = 13	0.7661971830985915
K = 14	0.7661971830985915
K = 15	0.7436619718309859
K = 16	0.7577464788732393
K = 17	0.752112676056338
K = 18	0.7464788732394365
K = 19	0.7380281690140844

נראה כי 7 ו-8 נותנים ביצועים טובים ודומים, אך למרות שיש פונקציה שמכריעה במקרה של שיוויון בהצבעה בין השכנים, הוחלט לקחת את הערך 7 שהוא לא מתחלק ב-2 ולא ב-3 (מספר הסיווגים האפשריים) ולכן לא יכול להיות בו שיוויון.

Perceptron

כאן היה צורך לכייל שני היפר-פרמטרים. הפרמטר של קבוע הלמידה שונה ידנית החל מ-1 כאשר כל פעם בוצעה חלוקה ב-10 (1, 0.1, 0.01...), והפרמטר של מספר האפוקים ללמידה נבחן ע"י לולאה.

סיכום:

קבוע הלמידה – מספר אפוקים עם הדיוק המקסימלי - דיוק

0.0001 - 36 - 0.58

0.001 - 34 - 0.69

0.01 - 26 - 0.73

0.1 - 20 - 0.74

1 - 10 - 0.74

על מנת לא לבחור בקבוע למידה גדול מדי של 1 ולעשות את הצעדים יותר הדרגתיים, הוחלט על קבוע למידה 0.1, ועל אחד ממספר האפוקים השונים אשר הביא לתוצאה הגבוהה ביותר והוא 20 אפוקים, כאשר הייתה מחשבה גם על זמן ריצת התוכנית.

PA

כיוון שאלגוריתם הפאסיב-אגרסיב מבצע תיקונים גדולים בכל דוגמה ובנוסף מוסיף מרווח טעות מסוים, מספר האפוקים לא היה אמור לשנות את ביצועי האלגוריתם. כיוון שגם אין קבוע למידה שנקבע על ידינו, אין לאלגוריתם זה היפר פרמטרים בקוד.