**עדכון התקדמות**

1. הוחלף ה dataset האדום, והוספו 3 datasets חדשים שלקחתי מהרשת, כך שיש ברשותי 5 datasets שאני עובד עליהם, ולא קלים מבחינת חיזוי:

|  |  |  |
| --- | --- | --- |
| Num Samples | Num features | Dataset |
| 1000 | 20 | German credit |
| 1340 | 19 | NBA logreg |
| 768 | 8 | Diabetes |
| 2785 | 9 | Water potability |
| 1600 | 11 | Wine quality |

1. עדכנתי את השינוי שהצעת באלגוריתם ה anytime. כרגע בנוסף הוא גם שומר את קבוצת התכונות שנתנה את התוצאה הטובה עד עכשיו.
2. תוצאות:

לכל dataset שני גרפים:

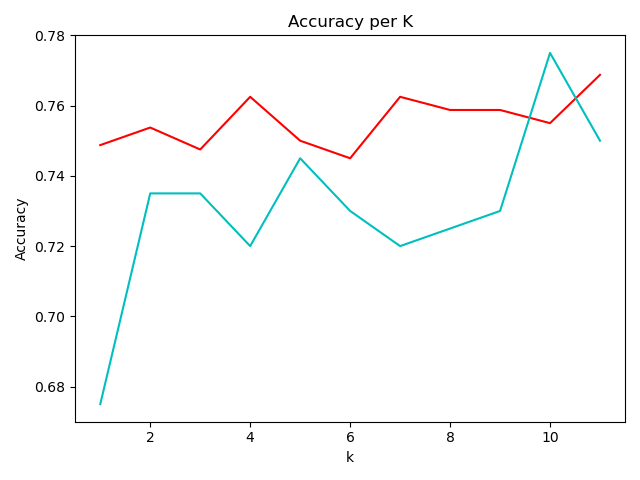
* גרף 1: שני plots, אחד של דיוק המודל עם ה features שנבחרו ע"י אלגוריתם ה anytime על קבוצת האימון (ממוצע על CV), והשני, דיוק המודל עם ה features שנבחרו ע"י אלגוריתם ה anytime על קבוצת המבחן.
* גרף 2: מיצוע הדיוק על קבוצת המבחן כפי שהצעת במייל האחרון (הפרש מהדיוק הראשון).

German credit dataset:

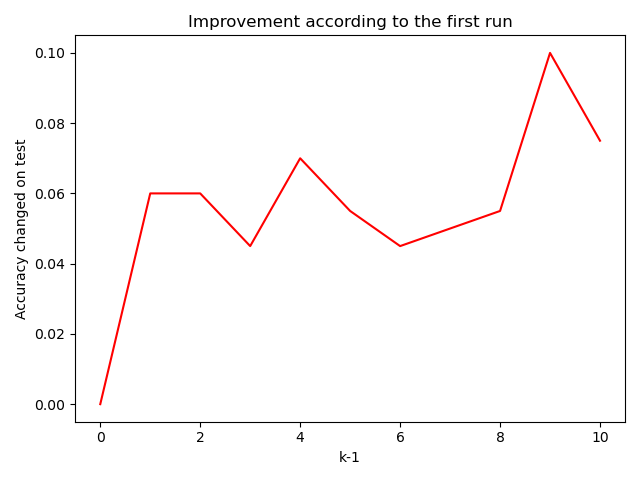
Graph 1:

Red – train accuracy

Blue – test accuracy



Graph 2:

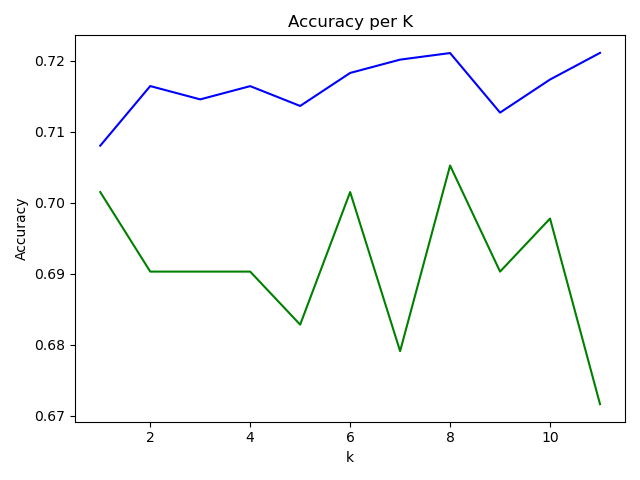


NBA logreg dataset:

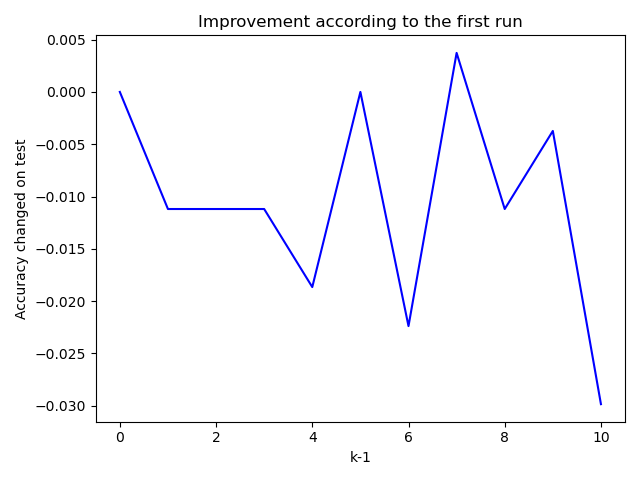
Graph 1:

Blue – train accuracy

Green – test accuracy



Graph 2:

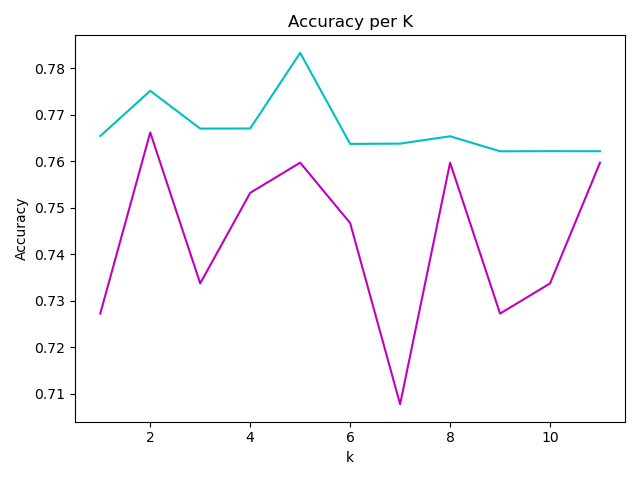


Diabetes dataset:

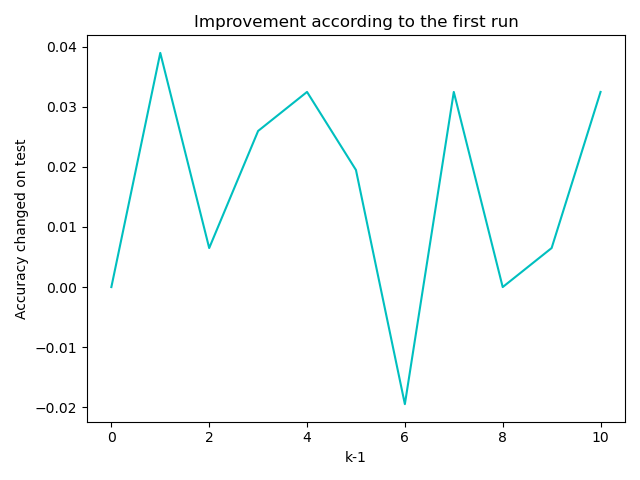
Graph 1:

Blue – train accuracy

Purple – test accuracy



Graph 2:

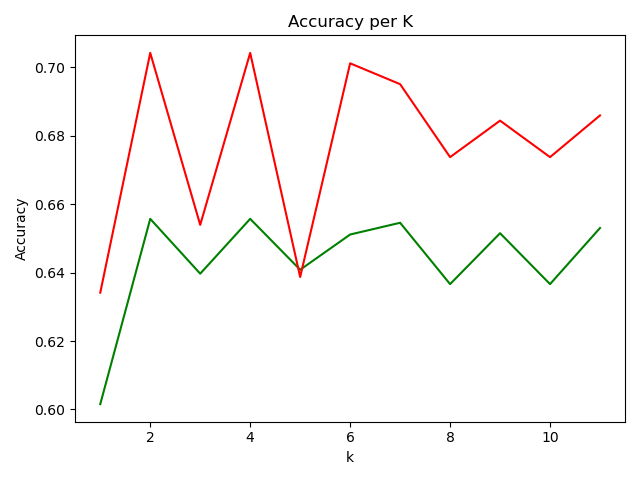


Water potability dataset:

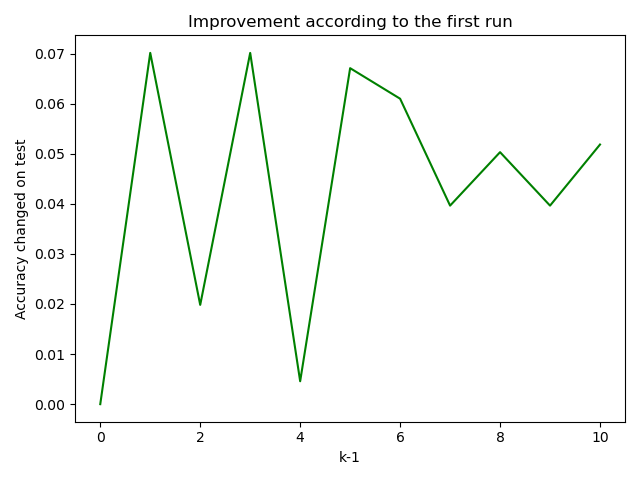
Graph 1:

Green – train accuracy

Red – test accuracy



Graph 2:

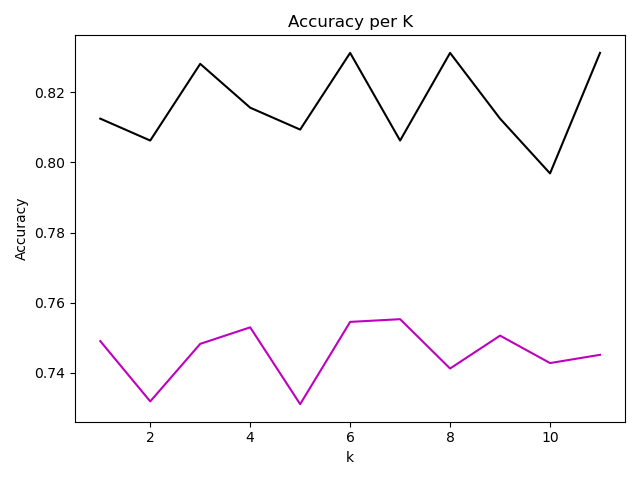


Wine quality dataset:

Graph 1:

Purple – train accuracy

Black – test accuracy



Graph 2:

