חלק ב' – פיתוח KNN

מימוש KNN- מצורפי קבצי הקוד להגשה.

הרצת ניסויים:

**תרגיל 5:**

1. מצורף *experiments*6.*csv* להגשה.
2. מבין המדידות שעשינו עבור k=3 התקבל הערך המקסימלי של דיוק 0.95 כלומר 95%.
3. ערך המינימום הינו: 0.934 ומתקבל כאשר k=13, ערך המקסימום הינו 0.95 ומתקבל כאשר k=3. כפי שניתן לראות מהגרף ישנה מגמת עלייה עד k=3 וישנה מכמת ירידה החל מk=3. מגמת העלייה בדיוק הראשונית נובעת מכיוון שהגדלת הערך של k מונעת תופעות של overfitting אמנם הגדלת יתר של k כפי שאנו שמים לב דווקא גורמת לירידה בדיוק מכיוון שהאלגוריתם מסתמך על שכנים "רחוקים" יותר, נוסיף כי העלייה בסוף הגרף מערך 7 לערך 13 היא מזערית ולא יכולה להעיד כי הדיוק יעלה אם נמשיך להגדיל את k.  
   כפי שניתן לראות הצלחנו למצוא k אופטימלי בעזרת k – cross - fold validation .

**תרגיל 7:**