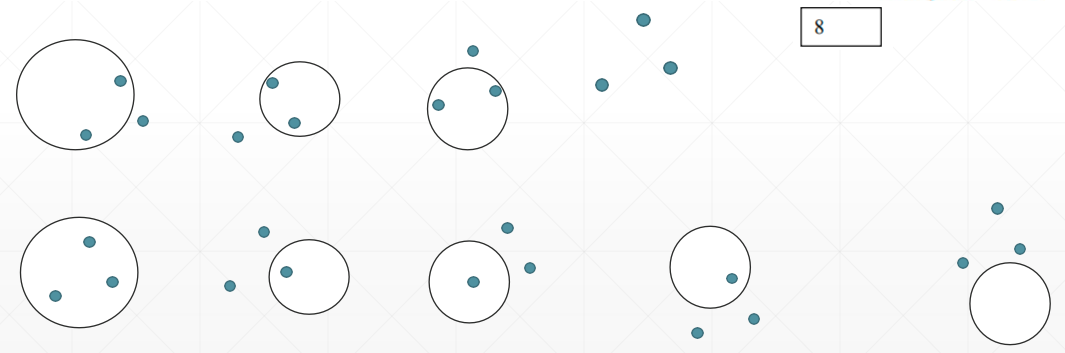
**מבוא ללמידה ממוכנת: שיעורי בית 3 –**

**מגישים:**   
נעם לביא - 325674513  
ליאון גורין – 214511214

**שאלה 1:**

Section a: ניעזר בציור מתוך התרגול לקראת המבחן כדי להראות שה (נקודות במעגל יסווגו עם +, אחרת עם -):



נראה דוגמה המפריכה את הטענה :

לא ניתן להכיל את 2 הדגימות החיוביות במעגל ללא עוד דגימה שלילית שתיכנס ותסווג לא נכון.

לכן .

Section b: ניעזר בנוסחה:

*נציב את הנתונים בנוסחה:*

כלומר אנחנו צריכים סדר גודל של דגימות כדי ש יהיה AGNOSTIC PAC Learnable.

Section c: תחת ההנחת ה Realizability ועם הנתון ש נוכל להשתמש בנוסחה:

*נציב את הנתונים בנוסחה:*

ולכן צריך לפחות: דגימות.

**שאלה 2:**

Section b: קיבלנו שאחוז ההצלחה הוא 100 כי עץ החלטה תמיד יעשה overfit לפי ה data הנתון לו במהלך ה training אם לא נגביל לו את הגובה.

Section c:

כאשר maxdepth=2 קיבלנו שאחוז ההצלחה באימון היה: 77.13% וטסט: 80.35%.  
במקרה זה קיבלנו תוצאה של underfitting כי סיבוכיות המודל הייתה קטנה מדי – שגיאות אימון גבוהה.

כאשר ה maxdepth=5 אחוז ההצלחה באימון היה: 96.89% ובטסט: 92.49% לעומת  
במקרה זה קיבלנו תוצאה של overfitting, שגיאת האימון הייתה נמוכה אבל עבור data שעדיין לא ראינו נקבל שגיאה גבוהה יותר.

Safety

high

med

low

Persons

unacc

Persons

2

2

4

more

4

more

acc

unacc

acc

unacc

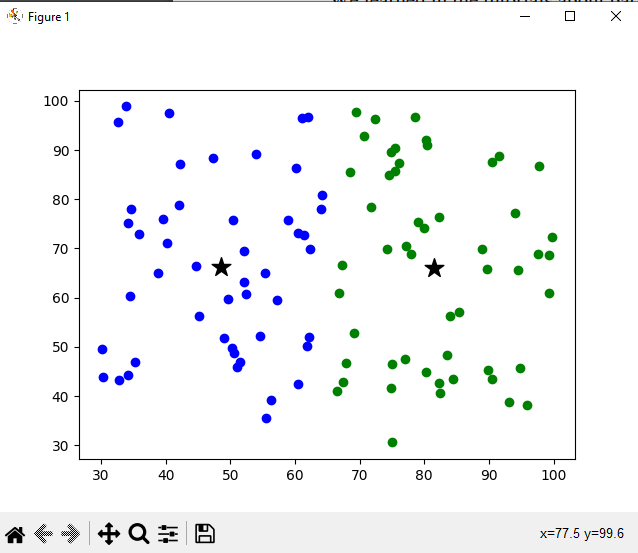
acc

acc

**שאלה 4:**

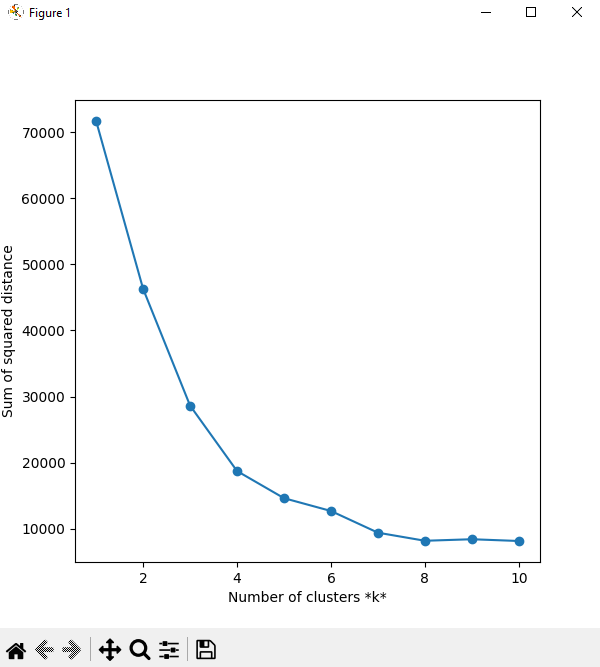
Section a: נרצה למצוא את הנגזרת של הביטוי כדי לראות מה הערך שמוביל למינימום עבור :

Section b:



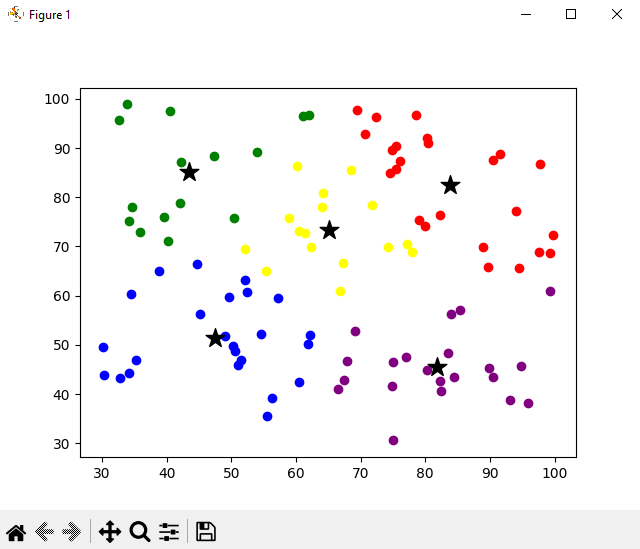
Section c:

**Elbow plotting –**



לפי הגרף, בחרנו ב k=5 כי ישר אחריו השיפור מתחיל להשתטח והשיפור נהיה מזערי.

**סיווג הנקודות לפי k=5**



Section d:

