

## תרגיל מס' 4

### למידה שאיננה מונחית

בתרגיל זה אתם תיצרו את הדאטה בעצמכם – ותראו איך נראית התפלגות נורמלית רב מימדית. לאחר מכן תפעילו אלגוריתם KMEANS (או גם GMM לפי בחירת רמת קושי כרצונכם).

- (1) עליכם לייצר דאטה שמורכב מ-2 תכונות ושלוש מחלקות. כל מחלקה תיוצר על ידי התפלגות נורמלית רב מימדית בעזרת הפונקציה: `numpy.random.multivariate_normal` : שימו לב לבחור מטריצת covariance חוקית (סימטרית! וערכים עצמיים גדולים מאפס)
- (2) בדקו עבור אחת המחלקות שאתם מסוגלים לשחזר את המטריצה והממוצע מהקלט.
- (3) ציירו את המחלקות, כל מחלקה בצבע אחר נסו שתהיה קצת חפיפה בין המחלקות
- (4) השתמשו באלגוריתם k-means כדי למצוא את המחלקות השונות. השתמשו ב-k שונים ובדקו האם באמת "גיליתם" שבדאטה מתחבאים 3 מחלקות. (שימו לב שכבר כתבנו קוד ב KNN שפורסם לכם שמחשב מרחקים.
- (5) רק למי שרוצה ציון גבוה ממש – ממש את אלגוריתם GMM עם 3 מחלקות ובדוק התאמה לדאטה המקורי

כרגיל ההגשה היא בזוגות, לינק למחברת, עד ה 15.8, רשמו מסקנותיכם מהתרגיל בתוך המחברת בקטע טקסט.

בהצלחה!