

חישוב מדדי הצלחה שונים לאלגוריתם סיווג

לספריית sklearn יש תת-ספרייה של data. אנחנו ניקח משם את הדאטה של digits. (ראה למשל: https://scikit-learn.org/stable/auto_examples/cluster/plot_kmeans_digits.html)

על דאטה הזו (עשר מחלקות - עשר ספרות - ו-64 תכונות)

עליך לבחור שתי מחלקות כרצונך ולנסות להפריד ביניהן על ידי logistic_regression. (השתמשו באלגוריתם המובנה ב sklearn)

שימו לב לנרמל, לחלק לשתי קבוצות train/test.

הדפיסו את ה confusion matrix. חשבו ערכי המדדים השונים שלמדנו.

עבור ערכי סף שונים חשבו את TPR ו-FPR וצירו את גרף ה ROC. כמו כן חשבו את ה AUC.

כרגיל יוגש לינק למחברת python דרך ה colab – לתרגיל תקראו ex2.