**תרגיל מספר 3 – עצי החלטה**

**תרגיל בית 03**

**תיאור הנתונים**:

קובץ הבסיס של התרגיל הוא הקובץ iris.

לקובץ זה עשינו preprocessing, חילקנו אותו לשני קבצים: iris-testing, iris-training המצורפים למטלה.

מצ״ב קובץ decisionTreeHandout.py שעליו נבצע את התרגיל.

המטרה: לבנות עץ החלטה לסיווג האירוסים לשלושה סוגים: Iris-setosa, Iris-versicolor, Iris-virginica.

**המשימה**:

1. כתבו פונקציה gini המקבלת שלושה ארגומנטים: ds, classes, mid ומחזירה את מדד gini של חלוקת הנתונים.
2. כתבו פונקציה chooseSplit המקבלת שני ארגומנטים: ds, classes ומחזירה את האינדקס של התכונה שלפיה מתבצעת החלוקה ואת ערך התכונה שלפיה מתבצעת החלוקה.
3. כתבו פונקציה buildTree המקבלת שלושה ארגומנטים: ds, classes, leafSize = 1 ומחזירה עץ החלטה.

**חשוב:**

**הפונקציות יהיו בקובץ ששמו hw3.py שאותו יש לשלוח לתיבת ההגשה**