**הוראות לדו'’ח בנושא Unbalanced Datasets**

* חקור את הנושא, הבן את השיטות המרכזיות להתמודדות עם בעיה זו.
* השתמש ב-Heart Disease Dataset כמקרה בוחן לבעיה (נדרש לחזות האם אדם סובל ממחלת לב).
* שים לב כי המידע המצורף הינו יחסית מאוזן (מבחינת יחס הדוג' חיוביות לשליליות). תחילה השתמש במודל מסוים ובחן תוצאות עם היחס הדוג' חיובי:שלילי מאוזן 1:1.
* לאחר מכן, בחר דוגמאות חיוביות, כך שהיחס הינו 5:1. שמור על אותו מס' דוגמאות כולל.
* בחן שלוש שיטות שונות להתמודדות בעיית ה-unbalanced.
* סכם בדו'’ח מחקרי את תוצאותייך.
* עלייך לכלול גרפים המדגימים את הבעיה ופתרונה. ניתן להדגים על מס' מודלי קלסיפיקציה שונים.

**מבנה הדו'’ח**

* מטרה
* תקציר
* רקע תיאורטי
* מהלך הניסוי
* תוצאות הניסוי
* מסקנות ודיון