**תכנית פעולה עבור מיני פרויקט- אנליזה של מידע**

אלגוריתם לבניית עץ החלטה:

1. נתונים:

- מדגם האימון **בקלט**, קבוצה של דוגמאות עבורם נבנה את עץ ההחלטה.

**L** - הפרמטר לקביעת גדלי העץ עבורם נריץ את האלגוריתם.

**P** - אחוז הדוגמאות ממדגם האימון בקלט אותם נפריד ממדגם האימון עליו נעבוד, ונבנה מהם את ממדגם הולידציה לצורך בדיקת הצלחת האלגוריתם המשוערת ע"י הפעלת מודל החיזוי על מדגם הולידציה.

1. חלץ את מדגם הולידציה מ- (מימוש החילוץ נפרד, שימוש בפרמטר P).

נסמנו .

1. נסמן את מדגם האימון ב-.
2. מצא את התווית הנפוצה ביותר במדגם האימון. נסמנו .
3. לכל :
   1. בנה עץ בינארי בעל צומת יחידה (שורש) שהיא .
   2. .
   3. כל עוד קטן מ-:
      1. בחר עלה מ- לפי "מדד האינפורמציה" (מימוש הבחירה בנפרד).
      2. החלף את בצומת חדשה עם תנאי שנבחר לפי שיקולי אנטרופיה (מימוש בחירת תנאי בנפרד).
      3. הוסף ל- 2 בנים עלים כך שלכל בן מתקיים:
         1. חלץ את קבוצת הדוגמאות המגיעות ב- לבן.
         2. **מתוך הקבוצה הנ"ל**, מצא את התווית המפוצה ביותר.
         3. בחר את ערך הבן להיות התווית הנ"ל.
      4. .
      5. חזור ל- (3.1).
   4. בשלב הזה יש לנו עץ החלטה בגודל : הרץ את על ושמור את אחוזי ההצלחה שלו.
4. בשלב זה סיימנו לעבור על כל האפשרויות עבור עצי ההחלטה: החזר את עץ ההחלטה (בפורמט כלשהו, JSON אולי) שנתן את אחוזי ההצלחה הגבוהים ביותר עבור .

**מודל החיזוי –** אלגוריתם שמקבל עף החלטה ומדגם בדיקה, מריץ את הדוגמאות מהמדגם על עץ ההחלטה (בעצם מתאים על דוגמא לתווית לפי עץ ההחלטה).