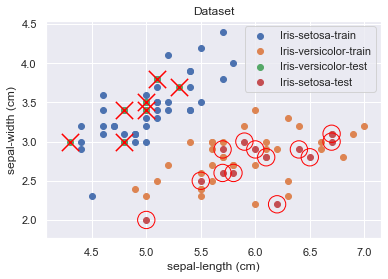
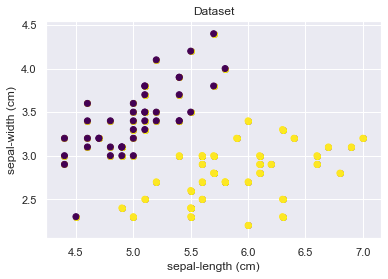
**מעבדה 7 – כריית נתונים**

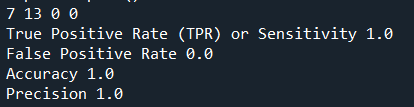
שם מגיש: אילון בן סימון ת"ז: 312162951

שם מגיש: סער ויקטור ת"ז: 312392822

שאלות 5-7:

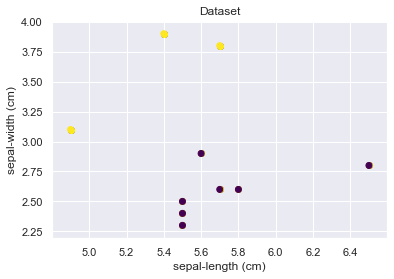
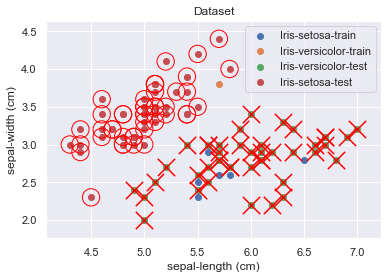
להלן החלוקה המקורית אל מול החלוקה שבצע אלגוריתם SVM. החלוקה הינה 20 אחוז מהנתונים הם נתונים לבדיקה ו-80 אחוז מהנתונים הם נתונים "מאומנים" ע"י אלגוריתם SVM. ניתן לראות שהאלגוריתם חילק את הנקודות למיקומים נכונים כלומר בסמיכות גבוהה לחלוקה המקורית.

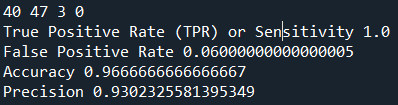


להלן תוצאות חלוקת האלגוריתם SVM (tp=7,tn=13,fp=0,fn=0) וחישוב טבלת הערכים:

שאלות 5-7 (נתונים אחרים):

להלן החלוקה המקורית אל מול החלוקה שבצע אלגוריתם SVM. החלוקה הינה 90 אחוז מהנתונים הם נתונים לבדיקה ו-10 אחוז מהנתונים הם נתונים "מאומנים" ע"י אלגוריתם SVM. ניתן לראות שהאלגוריתם חילק את הנקודות למיקומים נכונים כלומר בסמיכות גבוהה לחלוקה המקורית.

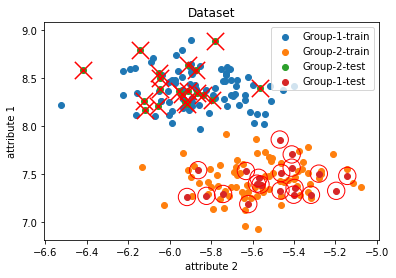
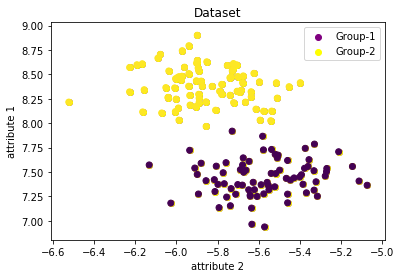
להלן תוצאות חלוקת האלגוריתם SVM (tp=40,tn=47,fp=3,fn=0) וחישוב טבלת הערכים :

שאלות 8-12 (בהתאם למעבדה 5.02):

להלן החלוקה המקורית אל מול החלוקה שבצע אלגוריתם SVM. בסעיפים אלו יצרנו 200 דגימות עם 2 נקודות מרכז (centers) בעזרת הפונקציה make\_blobs כלומר ניתן להתייחס לנתונים כ"שני פרחים, שתי תכונות ו-200 דגימות"

make\_blobs(n\_samples=200, centers=2, cluster\_std = 0.20, random\_state = 5)

לאחר מכן בצענו חלוקה כאשר 20 אחוז מהנתונים הם נתונים לבדיקה ו-80 אחוז מהנתונים הם נתונים "מאומנים" ע"י אלגוריתם SVM. ניתן לראות שהאלגוריתם חילק את הנקודות למיקומים נכונים כלומר בסמיכות גבוהה לחלוקה המקורית.



להלן תוצאות חלוקת האלגוריתם SVM (tp=20,tn=20,fp=0,fn=0) וחישוב טבלת הערכים:

