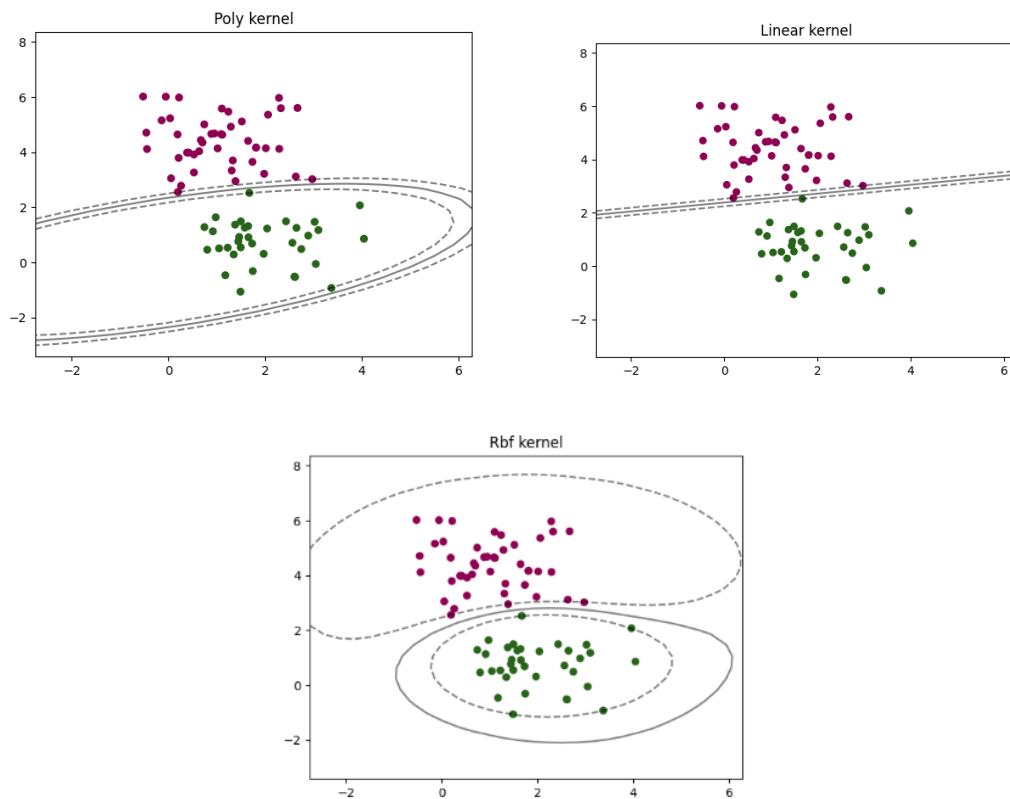


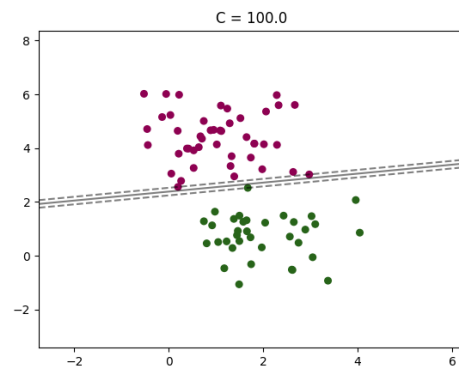
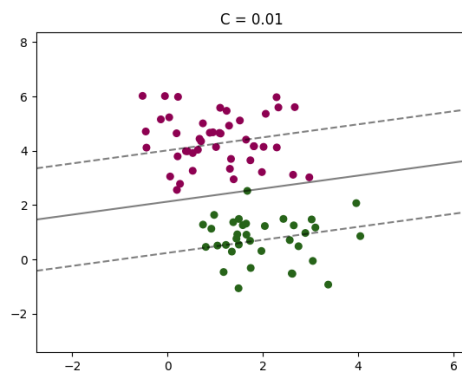
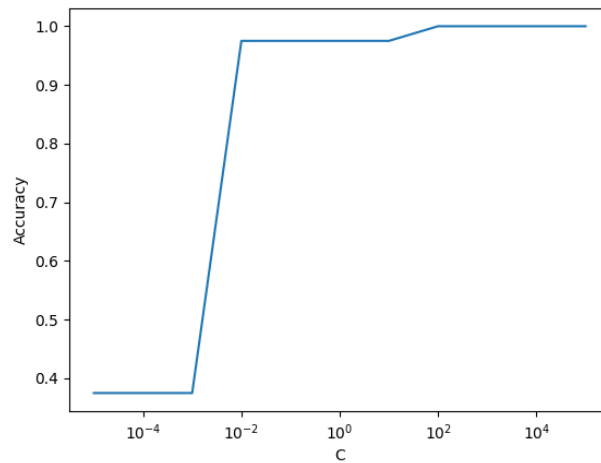
Programing Assignment

- (a) כפי שניתן לראות בגרפים המצורפים מטה, הצורה הגאומטרית של גבולות ההחלטה במישור משתנים עבור kernels שונים, וכך גם מספר ה-Support Vectors:
- עבור Linear Kernel: מפריד לינארי עם Support Vectors (2,1) 3.
 - עבור Quadratic Kernel: אליפסואיד עם Support Vectors (2,2) 4.
 - עבור RBF Kernel: קומבינציה לינארית של גאוסיינים מסביב ל-Support Vectors (3,3) 6.



נשים לב שהמורכבות של RBF התאימה את המפריד להתפלגות המדגם יותר מאשר ה-kernels האחרים והביא ל-margin הגדול ביותר. אכן הדבר נדמה כיתרון במבט ראשון, אך זו גם סכנה ל-over-fitting (לדוגמה לא נצפה שנקודה בתחתית הימנית של הגרף תהיה אדומה אך זו התוצאה לפי המפריד שהוגדר עם kernel זה).

(b) עבור $C \geq 100$ קיבלנו דיוק של 100% על ה-Validation Set, ומכאן ש- C בתחום זה יהיה אופטימלי. C הוא היפר-פרמטר לרגולציה המשפיע על המשקל הניתן לשגיאות, וככל שהוא גדול יותר האלגוריתם יחפש מפרד עם margin נמוך יותר. ערכים נמוכים יותר של C מכריחים את המפרד להיות עם margin גדול יותר גם במחיר של קלסיפיקציה שגויה של דוגמאות (דבר שייתכן שנרצה במקרים מסוימים).



(c) הפרמטר γ קובע את מידת השפעה שתהיה לכל נקודה כתלות במרחק, והוא אופטימלי בתחום [1,10] (וסביבתו), שכן בתחום זה קיבלנו 100% דיוק על ה-Validation Set. γ גדול מידי יוביל לדיוק מושלם על ה-Training Set אך יפגע בדיוק של ה-Validation Set כיוון שבמצב זה המפריד סובל מ-over-fitting. כאשר γ נמוך מידי, הדיוק הן על ה-Training Set והן על ה-Validation Set נפגע שכן המפריד לא מצליח לבטא מספיק טוב את המורכבות של ההתפלגות ולכן הוא סובל מ-under-fitting.

