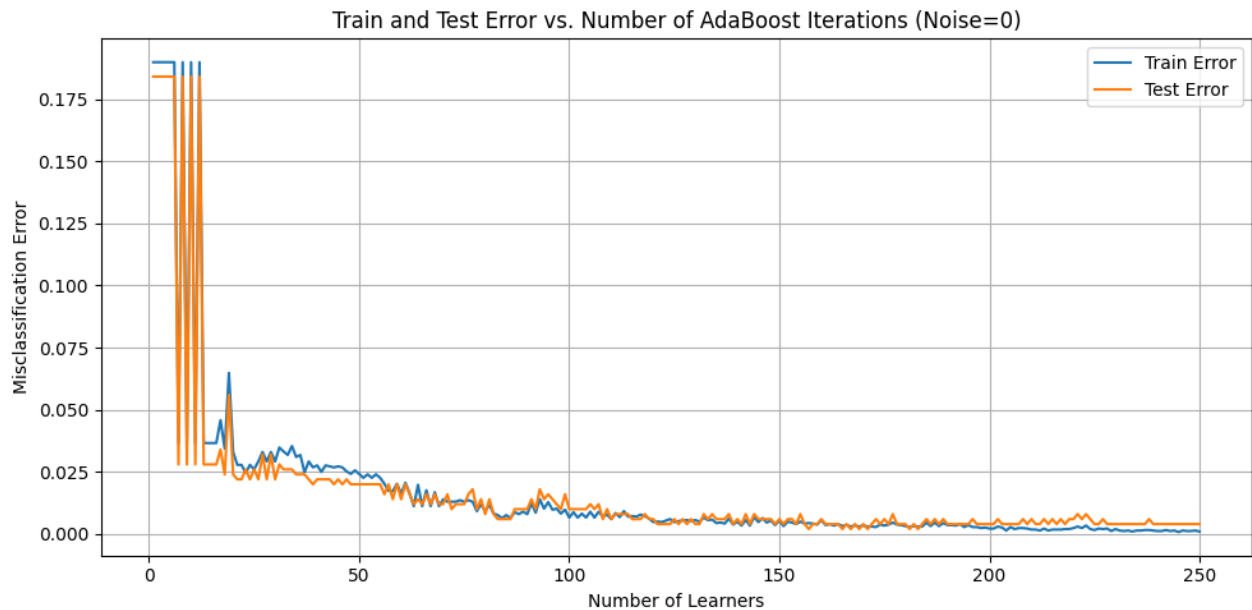


Practical part Answers:

Boosting:

0 noise:

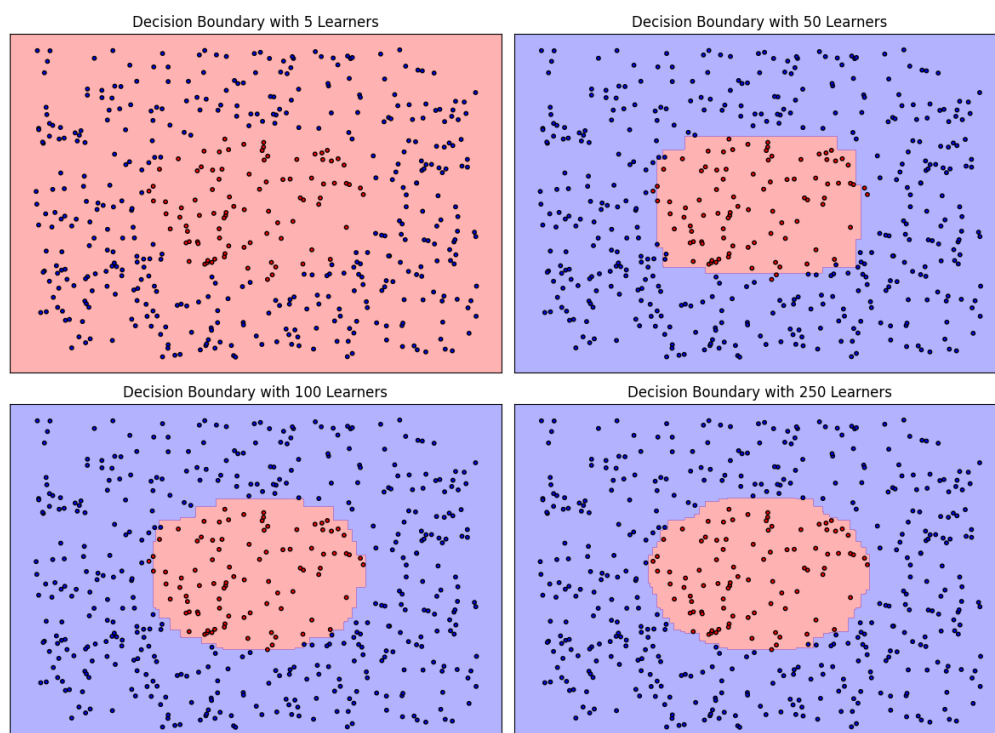
1)



כפי שניתן לראות התוצאות תואמות את הציפיות- ככל שמספר הלומדים עולה שגיאת הקלאסיפיקציה יורדת, בערכים הקטנים יחסית השגיאה מאוד תנודתית אך ממספר מסוים באזור ה-25 לומדים השגיאה מתייצבת ומשתפרת באופן איטי מאוד כפונקציה של הגדלת מספר הלומדים. מגמה זאת הגיונית מאחר שממספר לומדים מסוים התוצאות כבר די מתמחרות והלומדים חוזרים על עצמם ולכן מדייקים ברמה מאוד קטנה את השגיאה.

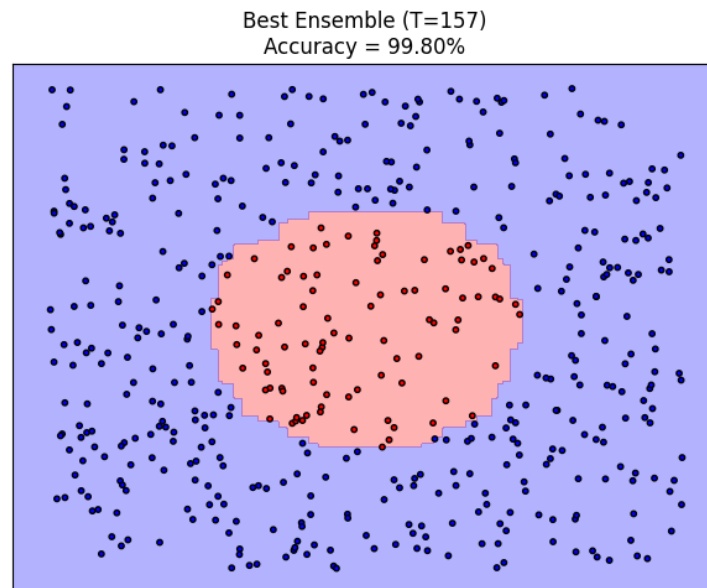
2)

AdaBoost Decision Boundaries at Different Iterations

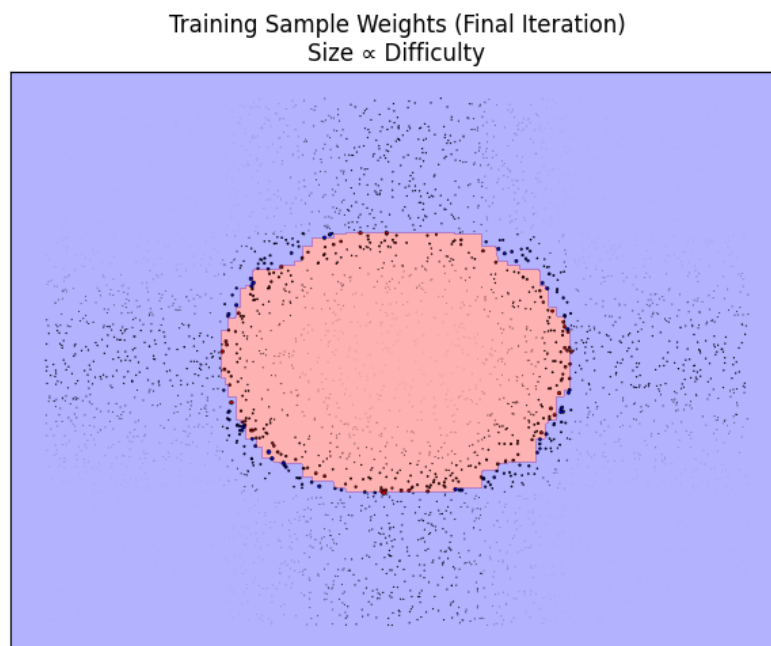


ניתן לראות שבמספר לומדים נמוך גבול ההחלטה היה כל הדאטה ולכן נתן שגיאה גדולה ולא מדויקת אך ככל שמספר הלומדים עולה הגבול מתהדק על הדאטה ונותן תוצאות איכותיות יותר. בנוסף ניתן לראות שהגבול עצמו נהיה יותר ויותר מעגלי בהתאם למספר הלומדים ופחות מרובע וזה הגיוני מאחר ויש למודל יותר גדמי החלטה לעבוד לפיהם כלומר יותר מלבנים לסמן על גרף ההחלטה ולכן הוא יכול לדייק יותר ויותר את הגבול האמיתי על הדאטה עצמו שנמצא בצורה מעגלית.

3)

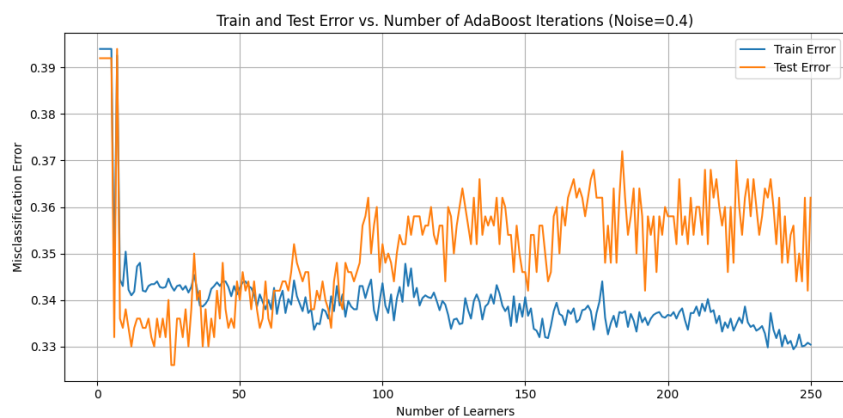


4)

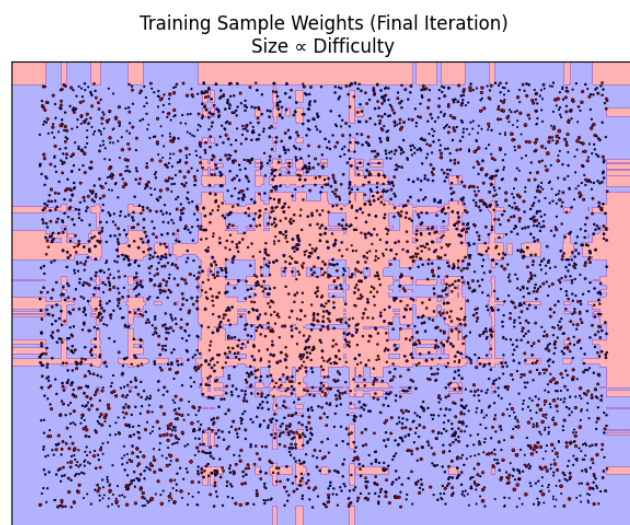
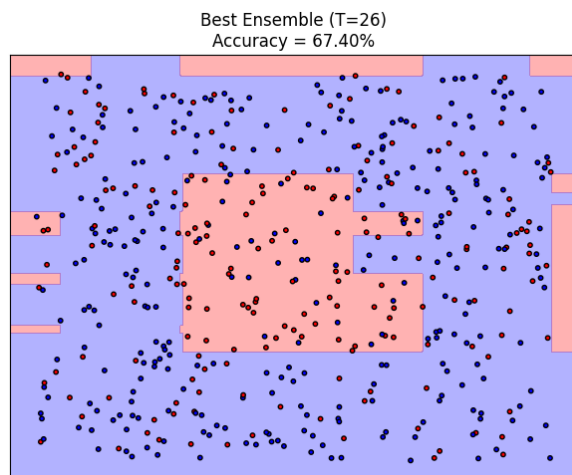
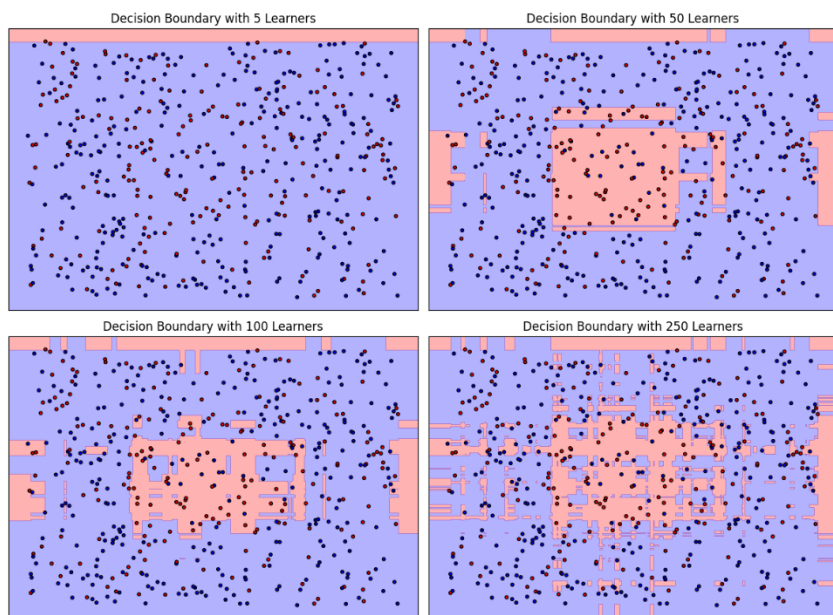


ניתן לראות בגרף הממושקל שלנקודות על התפר של גבול ההחלטה היו משקלים יותר גדולים בהרבה מהנקודות הרחוקות מהגבול, זאת תוצאה הגיונית מאחר ונקודות אלה הרבה יותר משמעותיות לגבול ההחלטה מאחר והדאטה בהן פחות חד משמעי ויותר חשוב לדייק בהן בקבלת הסיווג הסופי.

5)



AdaBoost Decision Boundaries at Different Iterations



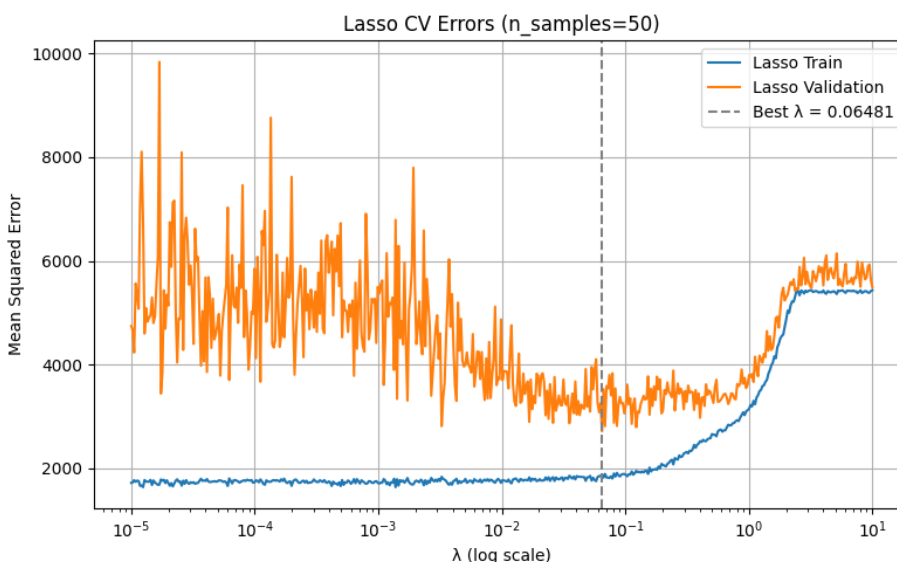
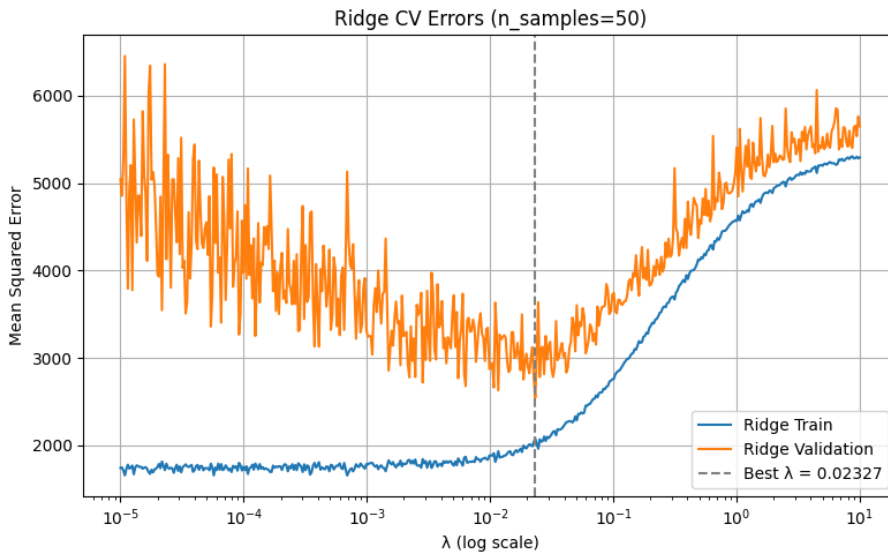
ניתן לראות מספר שינויים בגרפים כאשר הרעש הוא 0.4. ראשית המודל הרבה פחות מדויק באופן כללי כפי שניתן לראות בגרף הראשון והשלישי, שנית ככל שיש יותר לומדים מאחר והדאטה אינו מעגלי מושלם נוצרות יותר הפרדות בין הדגימות וגבול ההחלטה נהיה יותר ויותר מפוזר ומבולגן. שלישית בגרף הרביעי ניתן לראות שמשקל הדגימות מפוזר מאוד ולא בהכרח קרוב לגבולות ההחלטה כנראה בגלל הרעש שגורם לדגימות רבות לקבל משקל גבוה מהנדרש.

ניתן לראות בבירור בגרף הראשון את הביאס וארינס טרייד אוף, ככל שיש יותר לומדים השגיאה על דאטה האימון יורדת, כלומר הביאס יורד אך השגיאה על דאטה הטסט עולה- כלומר הוארינס עולה. דבר המייצג באופן מדויק את הטרייד אוף ואת הסכנה באובר פיטינג.

Cross Validation:

(2) תחילה ניסיתי למבדות בטווח של 0 עד 3 בקפיצות די גדולות וקיבלתי גרפים שלא כל כך הראו את המגמה כמו שצריך בנוסף למבדות אופטימליות מאוד קטנות ולכן החלתי ללכת על טווח למבדות בקפיצות מאוד קטנות ולייצג אותן בגרף בסולם לוגריתמי כדי להדגיש את ההבדל הקפיצות ולהראות את המגמה הנדרשת.

(3)



```
Best Ridge  $\lambda = 0.0233$ , Test Error = 3253.3949
Best Lasso  $\lambda = 0.0648$ , Test Error = 3515.1410
Least Squares Test Error = 3612.2211
```

ניתן לראות מספר דברים מהגרפים והתוצאות: ראשית בשני מודלי הרגולריזציה קיבלנו למבדה אופטימלי מאוד קטן שהביא לשיפור של בין 2-10 אחוזים בשגיאה לעומת רגרסיה לינארית ללא רגולריזציה. שנית ניתן לראות שלדאטה המסוים שלנו רידג' הביא לתוצאה יותר טובה כנראה בגלל שיש יותר פיצ'רים בעלי קישוריות לינארית לתוצאות (מאופן פעולת רידג'). בנוסף נשים לב למגמתיות בטווחי הלמבדה של שני מודלי הרגולריזציה, בשניהם בתחילת העלייה בערכי הלמבדה הקטנים ניכר שיפור בשגיאה בדאטה הוואלידציה כלומר המודל כולו השתפר אך החל מנקודת המינימום ניכרה עלייה כלומר המודל כבר לא משתפר ומתחיל להיות פחות טוב- זה מתכתב בדיוק עם החומר הנלמד בכיתה ומעיד על מציאת הלמבדה האופטימלי למודל. בנוסף נראה שבדאטה האימון ניכרת עלייה תמידית בשגיאה שזה הגיוני מאחר וככל שמשקל הרגולריזציה עולה הביאס עולה גם הוא ואיתו השגיאה על סט האימון. נשים לב שבמודל הלאסו החל מלמדה מסוים השגיאה מתייצבת שזה הגיוני מאחר ומודל זה מאפס פיצ'רים ולכן יכול לגרום להתייצבות השגיאה החל ממשקל מסוים.