**שאלה 2-ג**

בהתבסס על מה החומר שנלמד, מה אנחנו מצפים שהתוצאות יהיו והאם התוצאות תואמות את הציפיות שלנו?  
נחלק את התשובה שלנו ל3 חלקים ובכל תשובה נרחיב על הציפיות והתוצאות.

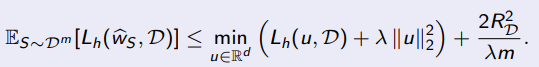
1. איזה גודל מדגם אמור לקבל שגיאת אימון קטנה יותר? מה לגבי שגיאת על הtest?
2. מה צריכה להיות המגמה בשגיאת האימון כפונקציה של λ? למה?  
   האם התוצאות עבור הניסוי הראשון תואמות את הציפיות שלך?
3. מה צריכה להיות המגמה בשגיאת הtest כפונקציה של λ? למה?  
   האם התוצאות עבור הניסוי הראשון תואמות את הציפיות שלך?

**פתרון:**

1. ציפיות:

שגיאת האימון – בניסוי הראשון גודל המדגם היה בגודל 100 ובניסוי השני גודל המדגם היה בגודל 1000.  
הציפיות שלנו הן ששגיאת האימון (בממוצע) תהייה קטנה יותר בניסוי הראשון מאשר בניסוי השני מאחר וקל יותר למצוא מפריד עבור מספר קטן יותר של דוגמאות (פחות דוגמאות לטעות עליהן), כלומר הסיכוי לטעות על מדגם האימון קטן כאשר גודל המדגם קטן.

שגיאת הtest – הציפיות שלנו במקרה זה יהיו הפוכות. מאחר שהתאמנו על יותר דוגמאות בניסוי השני, אנו מצפים שהשגיאה על הtest בניסוי השני תהייה קטנה יותר מאשר השגיאה על הtest בניסוי הראשון.  
לפי הנוסחה שלמדנו בכיתה ניתן לראות שככל שכמות הדוגמאות גדלה השגיאה קטנה:



תוצאות:  
שגיאת האימון – כמו שצפינו, ניתן לראות בגרף שאכן שגיאת האימון בניסוי הראשון קטנה יותר מאשר בניסוי השני.

שגיאת הtest –כמו שצפינו, התוצאות במקרה זה הפוכות לשגיאת האימון.  
ניתן לראות בגרף שאכן שגיאת האימון בניסוי הראשון גדולה יותר מאשר בניסוי השני.

1. ציפיות: למדנו בכיתה כי שככל ש- λגדלה כך השגיאה על האימון תגדל גם.

הסיבה לכך היא כשאנחנו מגדילים את λ אנחנו נותנים משקל גדול יותר למזעור הנורמה של u, כלומר נבחר לבסוף u בעל נורמה קטנה יותר וכתוצאה מכך ה soft margin יגדל.  
כאשר ה soft margin גדול יותר אנחנו פחות רגישים ל sample הספציפי ולכן השגיאה על האימון תגדל.

תוצאות: כמו שצפינו, ניתן לראות בגרף שכאשר λ גדלה גם שגיאת האימון גדלה.

1. ציפיות: למדנו בכיתה שישנו טרייד אוף בבחירת ה- λ:  
   כאשר נבחר λ גדולה – נעניש יותר עבור נורמה גדולה של u (הגורמת לsoft-margin קטן יותר), דבר העלול לגרום לאובר פיטינג.

כאשר נבחר λ קטנה – נעניש יותר עבור גודל מדגם קטן דבר שיגרום למרחב האפשרויות של האלגוריתם לרגישות לשגיאה שנוצרת ממדגם קטן. כתוצאה מכך לא ימצא מפריד לינארי בצורה טובה.  
לכן אנחנו מצפים למגמת ירידה בשגיאה על על ה-test עד שנגיע ל- λ האופטימלית ולאחר מכן השגיאה תחל לעלות.

תוצאות: כמו שצפינו, ניתן לראות בגרף שהשגיאה על הtest הקטנה ביותר מתקבלת לא ע"י λ גדולה ולא ע"י λ קטנה.  
בנוסף, ניתן לשים לב שאנו מקבלים שגיאה גדולה יותר ככל שאנו מגדילים את λ וגם שגיאה גדולה יותר כאשר λ קטנה מאוד.