

## למידת מכונה תרגיל 1

1. מכיוון שהעיגול מכיל את כל הדוגמאות החיוביות, ייתכן והוא יהיה חייב להכיל גם הרבה דוגמאות שליליות (למשל עיגול של דוגמאות שליליות המוקף בעיגול של דוגמאות חיוביות). במצב זה העיגול יכיל את כל הדוגמאות השליליות בנוסף לחיוביות והלמידה בהכרח לא תהיה טובה בסתירה ל-ERM). במצב זה, היה עדיף להסתפק בעיגול קטן יותר המכיל רק חלק מהדוגמאות החיוביות עם מספר קטן של דוגמאות שליליות מה שהיה מקטין את השגיאה בלמידה.

הצורה הגאומטרית המינימלית שתענה על הגדרת האלגוריתם היא מלבן. עיגול מוגדר על ידי רדיוס שכולל המון "שטח" לא נצרך. עבור מלבן, משתמשים בדיוק ב"שטח" הנצרך לכסות את כל הדוגמאות החיוביות ולכן ייתן את הכיסוי המינימלי עבור הגדרת המסווג.

2. צריך להוכיח:

$$\mathbb{E}_{S \sim D}[L_S(h)] = L_D(h)$$

הוכחה:

$$\begin{aligned}\mathbb{E}_{S \sim D}[L_S(h)] &= \mathbb{E}_{S \sim D}\left[\frac{1}{m} \sum_{m=1}^1 \mathbb{I}[h(x) \neq f(x)]\right] = \frac{1}{m} \sum_{m=1}^1 \mathbb{E}_{S \sim D}[\mathbb{I}[h(x) \neq f(x)]] \\ &= \frac{1}{m} \sum_{m=1}^1 p_{X \sim D}[h(x) \neq f(x)] = L_D(h)\end{aligned}$$

כאשר השוויון השני נובע מלינאריות התוחלת, והשוויון השלישי נובע מכך שהתוחלת של אינדיקטור היא ההסתברות של תנאי האינדיקטור.