שאלה 4:

1. למדנו בהרצאה כי w ניתן לייצוג על ידי $w = \sum_{i=0}^{m} \alpha_i x_i$ נשים לב שלפי אלגוריתם הפרספטרון בכל צעד עדכון של האלגוריתם אנחנו מעדכנים את הכניסה ה-i ב-w ומוסיפים לה את הוקטור $y_i x_i$, לכן לדוגמה, בעדכון השלישי של כניסה i באלגוריתם (אם קיים), ערכה של הכניסה ה-i ב-w תתעדכן להיות $3y_i x_i$. מתוך הבנה שלכל i, הביטוי $y_i x_i$ הוא קבוע לכל אורך הרצת האלגוריתם (מפני שהנק' x והתיוגים y נתונים ואינם משתנים), במקום לעדכן את w נוכל לעדכן רק את המקדם שסופר את מס' העדכונים של האלגוריתם בכל כניסה. נסמן את המקדם ב-c ואז יתקיים לכל כניסה i: $w = \sum_{i=0}^{m} \alpha_i x_i\ ,\ \alpha_i = c_i y_i$.

האלגוריתם ייראה כך:

Input: a training set $(x_1, y_1),..., (x_m, y_m)$
Initialize: $c = (0,..., 0)$, vector of length m
for $t = 1, 2,...$:
    if $\exists i$ such that $\widehat{y}(x_i) \neq y_i$
        $c_i = c_i + 1$
    else:
        return c

כעת, נתייחס לדרישה לכתוב את אלגוריתם הפרספטרון לאחר שהווקטורים x עברו מיפוי ψ. כפי שלמדנו בכיתה, ניתן לבטא את כלל ההחלטה לפיו אנחנו משייכים נקודה לתיוג חזוי גם על ידי פונקציית קרנל בצורה הבאה: $\widehat{y}(x) = sign\left(\sum_{i=1}^{m} \alpha_i \langle\psi(x_i), \psi(x)\rangle\right) = sign\left(\sum_{i=1}^{m} \alpha_i K(x_i, x)\right) = sign\left(\sum_{i=1}^{m} c_i y_i K(x_i, x)\right)$

לכל נק' x, $\widehat{y}(x)$ תהיה הפרדיקציה שלה כחלק מהשלבים באלגוריתם שכתבנו (בשורה של התנאי).

כמענה של שתי השאלות המכווינות, הפלט של האלגוריתם הוא הווקטור c שעבור כל נק' x סופר את כמות העדכונים שהאלגוריתם נאלץ לבצע עבורה עד להתכנסות שלו, בהתאם לכניסה המתאימה בווקטור c (כלומר בכניסה ה i יש את כמות הפעמים שהאלגוריתם עדכן את התיוג של הוקטור x_i).

2. מאחר שבכל צעד עדכון של כניסה i באלגוריתם פרספטרון המקורי שלמדנו בכיתה מתווסף לכניסה זו בווקטור w הביטוי $y_i x_i$, בסוף הריצה של אלגוריתם זה הערך של כל כניסה i יהיה

$w_i = \sum_{i=0}^{c_i} y_i x_i = c_i y_i x_i$, כאשר c_i הוא מס' העדכונים שנדרשו לכניסה ה-i במהלך הריצה. נשים לב שבמקום לעדכן את w_i נוכל לעדכן רק את c_i ונקבל את אותה התוצאה ונבצע את אותם עדכונים. לכן נכונות האלגוריתם שלנו נובעת באופן ישיר מנכונות אלגוריתם הפרספטרון.

כפי שנלמד בשיעור אלגוריתם הפרספטרון מתכנס אם ורק אם הדאטא פריד ליניארית. במקרה הנוכחי אנו מקבלים פונקציית קרנל על המרחב של המכפלה הפנימית של המיפוי ψ כלומר פונקצית הקרנל הינה $\langle\psi(x_i), \psi(x)\rangle$.

מאחר ואנו רוצים שהאלגוריתם יתכנס גם עבור הדאטא אחרי שהוא מופה מחדש, ומאחר ולא בוצע שינוי באלגוריתם הפרספטרון שלנו אלא רק ספירה של עדכונים ושינויים שבוצעו ל-w, לאלגוריתם שלנו תהיה אותה דרישה שיש לאלגוריתם המקורי על מנת שיתכנס: הדאטא S^{ψ} צריך להיות פריד לינארית בכדי שהאלגוריתם יתכנס, כלומר מנכונות אלגוריתם הפרספטרון ההתכנסות תקרה אם ורק אם S^{ψ} פריד ליניארית.