5. נניח שדוגמת אימון (x, y) בעלת תיוג $y = +1$ מסווגות באופן שגוי ידי פרספטרון עם וקטור משקלים w^t, כלומר $\hat{y} = \varphi_{HL}\left(w^{tT}x\right) = -1$. הראו כי עבור $\alpha > 0$ גדול מספיק, וקטור המשקלים המעודכן w^{t+1} יוביל לסיווג נכון, כלומר $\hat{y} = \varphi_{HL}\left(w^{t+1T}x\right) = +1$.

נתונה דוגמת אימון (x, y) בעלת תיוג $y = 1$ המסווגת באופן שגוי עבור וקטור המשקלים w^t, כלומר:

$$\hat{y} = \varphi_{HL}(w^{tT}x) = -1$$

כמו שראינו בחומר ההכנה מתקיים:

$$y = \varphi_{HL}\left(w^T x\right) = \begin{cases} -1: & w^T x < 0 \\ +1: & w^T x \geq 0 \end{cases}$$

ולכן עבור הנתון

$$\hat{y} = \varphi_{HL}(w^{tT}x) = -1$$

מתקיים כי:

$$w^{tT}x < 0$$

לפי שלב 3 באלגוריתם הלמידה – עדכון וקטור המשקלים w^{t+1}:

$$w^{t+1} = \begin{cases} w^t + \alpha \cdot (y - \hat{y}) \cdot x & , y \neq \hat{y} \\ w^t & , y = \hat{y} \end{cases}$$

עבור דוגמת האימון שלנו:

$$w^{t+1} = w^t + \alpha \cdot (y - \hat{y}) \cdot x = w^t + \alpha \cdot \big(1 - (-1)\big) \cdot x = w^t + 2\alpha x$$

$$\Longrightarrow \quad w^{t+1} = w^t + 2\alpha x$$

ובשלב הבא:

$$\hat{y} = \varphi_{HL}((w^{t+1})^T x\,) \overset{w^{t+1} = w^t + 2\alpha x}{\cong} \varphi_{HL}((w^t + 2\alpha x)^T x) = \varphi_{HL}(w^{tT}x + 2\alpha x x) = \varphi_{HL}(w^{tT}x + 2\alpha x^2)$$

נדרוש כי התקיים:

$$(w^{tT}x + 2\alpha x^2) \geq 0$$

$$2\alpha x^2 \geq -w^{tT}x$$

$$\alpha \geq -\frac{w^{tT}x}{2x^2} = (-1)\frac{w^{tT}}{2x}$$

בפרט מתקיים עבור

$$\alpha = (-1)\frac{w^{tT}}{2x}$$

$$\hat{y} = \varphi_{HL}((w^{t+1})^T x\,) = \varphi_{HL}(w^{tT}x + 2\alpha x^2) = \varphi_{HL}\left(w^{tT}x + 2\left[(-1)\frac{w^{tT}}{2x}\right]x^2\right) = \varphi_{HL}(0) = 1$$

ובפרט עבור $\alpha \geq 0$

6. הראו כי עבור האלגוריתם של רוזנבלט, בחירה של משקלים התחלתיים w^0 וגודל צעד α שקולה לבחירת משקלים התחלתיים w^0/α וגודל צעד 1. במינוח "שקולה" הכוונה לכך ששני המסווגים יתנו תוצאה זהה עבור כל סדרת בוחן.

קראנו בהכנה כי אנו עושים שימוש בפונקציה:

פונקציית ההפעלה φ שבה נבחר בניסוי זה היא פונקציית הפעלה מסוג Hard Limiter, והיא מותאמת *לסיווג בינארי* בין שתי מחלקות:

$$y = \varphi_{HL}\left(w^T x\right) = \begin{cases} -1: & w^T x < 0 \\ +1: & w^T x \geq 0 \end{cases}$$

וראינו בהכנה כי מתקיים $\alpha > 0$ שכן מייצג את גודל הצעד/התקדמות.

ולכן עבור $\alpha > 0$ מתקיים שגם $\alpha^{-1} > 0$

עבור המסווגים שלנו השוני היינו בהכפלה בקבוע (גדול מ-0) – עבור הצעד הראשון נקבל תוצאה זהה.

עבור הצעד הבא

$$w_a^1 = w_a^0 + 2\alpha x$$

$$w_b^1 = \frac{w_a^0}{\alpha} + 2x = (w_a^0 + 2\alpha x)\alpha^{-1} = \frac{w_a^1}{\alpha}$$

עבור הצעד הבא

$$w_a^2 = w_a^1 + 2\alpha x$$

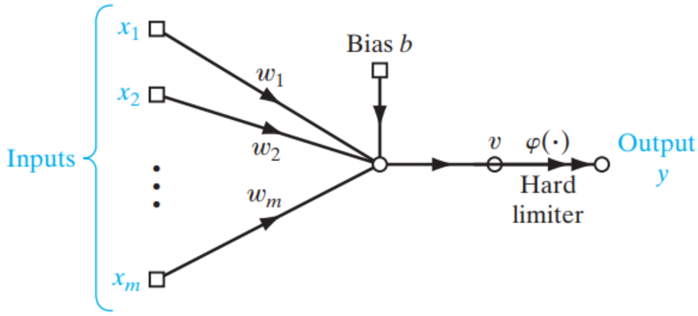
$$w_b^2 = \frac{w_a^1}{\alpha} + 2x = (w_a^1 + 2\alpha x)\alpha^{-1} = \frac{w_a^2}{\alpha}$$

עבור המקרה הכללי (בהנחה כי $y \neq \hat{y}$)

$$w_a^{t+1} = w_a^t + 2\alpha x$$

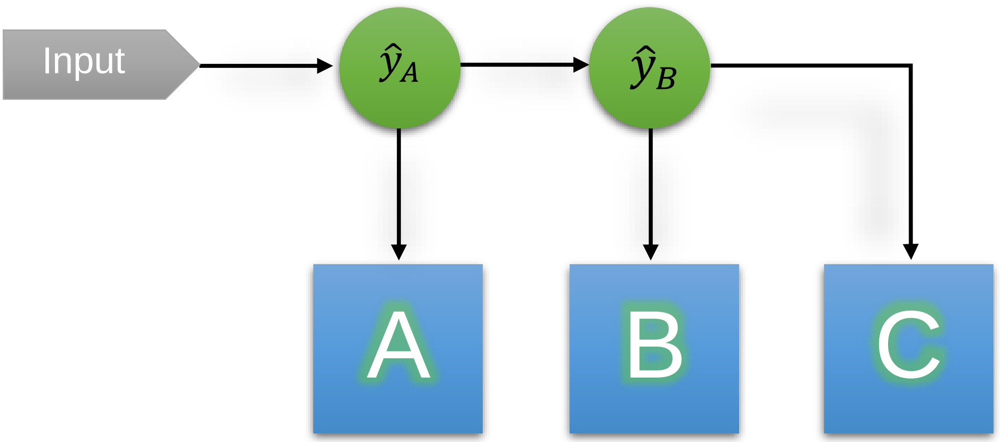
$$w_b^{t+1} = \frac{w_a^t}{\alpha} + 2x = (w_a^t + 2\alpha x)\alpha^{-1} = \frac{w_a^{t+1}}{\alpha}$$

ולכן עבור $\alpha > 0$ ועבור פונקציה $\varphi_{HL}(w^t)$ שהיא $hard\_limiter$ נקבל עבור על צעד בסדרת האימון והבוחן שוויון. (שכן הסימן לא משתנה)

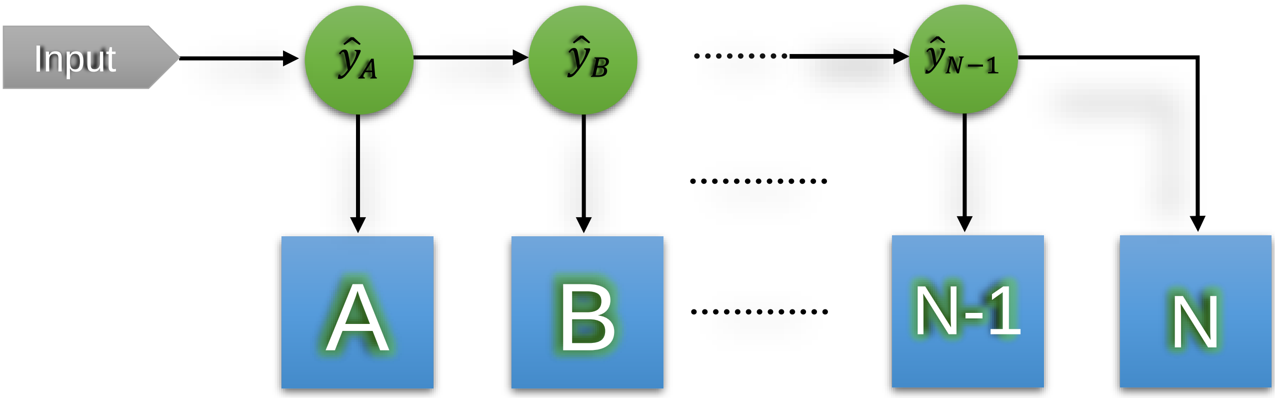


7. הציעו מימוש למסווג תלת מחלקתי הבנוי ממספר מסווגי פרספטרון בינאריים. לכמה מסווגים בינאריים תידרשו על מנת לממש מסווג n-מחלקתי?

- **מסווג תלת מחלקתי** יהיה בנוי משני מסווגי פרספטרון בינאריים, לדוגמא:



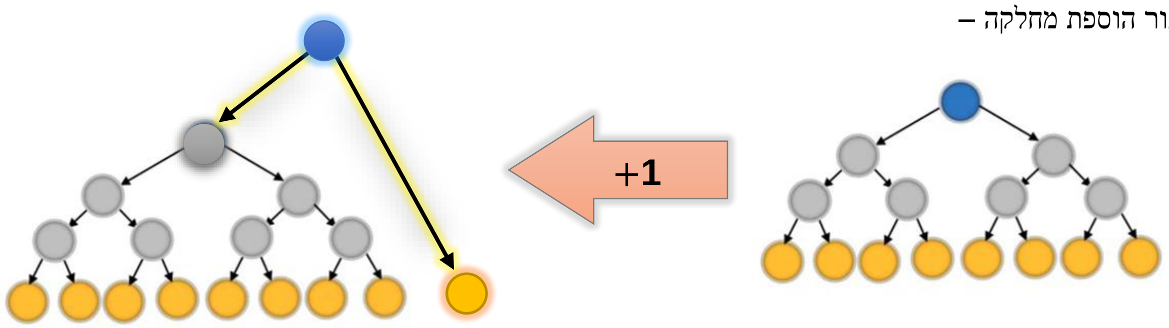
- **עבור המקרה הכללי** -



נסתכל על הבעיה כבנייה של עץ בינארי כאשר:

- ❖ כל עלה של העץ מייצג מחלקה – כלומר נבנה עץ בעל n עלים.
- ❖ כל צומת שהיא לפחות בעלת בן אחד (כלומר לא עלה) נספור כמסווג בינארי.
- ❖ שורש העץ יהווה המסווג אשר יקבל את וקטור הכניסה.

נקבל עבור n עלים (n מחלקות) – עץ בעל $(n-1)$ צמתים פנימיים. (טענה שהוכחה בקורס מבני נתונים)
נשים לב כי עבור הוספת מחלקה נוספת נצטרך להוסיף עוד מסווג אחד כלומר:
עבור הוספת מחלקה –



מסווג החדש יקבע שייכות למחלקות העץ הקודם או למחלקה החדשה שהוספנו.

ולכן למימוש מסווג n-מחלקתי נזדקק ל- $(n-1)$ מסווגי פרספטרון בינאריים