# שאלה 1 סעיף ה'

שלומי דומננקו 318643640

נגדיר את המודל שלנו:

זה אקטיבציה. כעת נגדיר ערכי כך:

*נזכור שעץ הגזירה נראה כך:*

A picture containing text, device, gauge

Description automatically generated

נגדיר מה נרצה קודם כל לחשב:

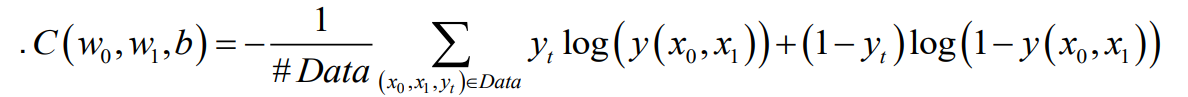
כשגוזרים את המשוואה עבור b למשל, נקבל בכל האיברים 0 חוץ מהמקדם של b, שהוא 1.

*בגלל ש נתון, נותר לחשב רק .*

*בגלל ש Z היא פונקציה שהקלטים שלה הם אזי צריך לגזור לפי כל אחד.*

פונקציית המחיר שבחרתי היא Cross Entropy כללי (עבור k פרמטרים).

בהרצאה 3 הוצג Cross Entropy cost function יותר ספציפית, עבור 2 פרמטרים:



A picture containing text, clock

Description automatically generated

וראינו גזירה:

Text, letter

Description automatically generated

עבור Sigmoid activation ולא ReLU:

A picture containing text, clock

Description automatically generated

נתחיל לגזור:

נשים לב שזה תואם את הנוסחה ל sigmoid gradient שמצאתי באינטרנט:

A picture containing diagram

Description automatically generated

כאשר מסמן sigmoid

<https://towardsdatascience.com/derivative-of-the-sigmoid-function-536880cf918e>

ולכן:

כעת נחשב נגזרת:

הנגזרת הראשונה – נתון לנו לפי השאלה

הנגזרת השנייה – חישבנו מקודם

נחשב את הנגזרת השלישית:

ולכן לסיכום: