

### למידת מכונה – רשתות נוירונים, דוח תרגיל 3:

#### מבנה הרשת:

- שכבת הקלט : מקבלת תמונה בגווי אפור שגודלה  $28 \times 28$  (784 פיקסלים)
- שכבת נסתרות : שימוש בפונקציית האקטיבציה
- שכבת הפלט: ע"י שימוש בפונקציית softmax לסיווג הקלט בין 10 הקטגוריות.

התוכנית קיבלה 3 קבצים שהם `train_x, train_y, test_x`, `data.n` חולק ביחס של 80% אימון ו 20% ל `validation`, מכיוון שאנו נרצה לבדוק את איכות האימון שלנו ובנוסף `train_x` עצום לכן חלוקה זאת אפשרי.

אתחלנו 4 פרמטרים שהם משקולות,  $w_1$  בגודל של  $(784 \times \text{hidden\_layer})$ ,  $w_2$   $(\text{hidden\_layer} \times 10)$  ו  $b_1$  ו  $b_2$  בגודל של  $10 \times 1$  ובנוסף כל הפרמטרים רנדומלית בין הערכים -0.08 ל 0.08.  
הערכים נורמלו בחלוקה על ידי 255 כך שהערכים ינועו בין 0 ל 1.

#### היפטר פרמטרים:

- גודל השכבה הנסתרת : בחרתי לאחר מספר בדיקות את גודל השכבה הנסתרת להיות 150, בדקתי מספר אופציות ובחרתי את המספר הזה כי הוא נתן את הביצועים הכי טובים ואחוזי ההצלחה השתפרו
- קצב למידה ( $\eta$ ) – 0.01 בחרתי קצב למידה נמוך יחסית על מנת שנוכל להמשיך תמיד להתקדם לעבר הירידה ולא נעבור אותה
- מספר מעברים (epochs) – 30
- הפונקציה שנבחרה היא Relu על פי הבקשה בתרגיל