

תרגיל 3 – רשתות נוירונים

מבנה הרשת:

❖ **שכבת קלט –** שכבת הקלט מקבלת תמונה בגוויי אפור (ערך בודד) גודל התמונה הוא 28×28 אך אנו מקבלים אותה כווקטור ערכים (מאבדים את דו-הממדיות של התמונה)

❖ **שכבת נסתרת 1 –** שימוש בפונקציית האקטיבציה Relu

❖ **שכבת פלט –** ע"י שימוש ב SoftMax על מנת לקטלג בין 10 אופציות (פרטי הלבוש השונים)

היפר פרמטרים:

❖ **גודל השכבה הנסתרת –** בחרנו לאחר מספר בדיקות את גודל השכבה הנסתרת להיות **128**
בדקנו מספר אופציות ובחרנו מספר נוירונים זה מכיוון שהוא נתן את הביצועים הכי טובים, וזה הגיוני בהתחשב בכך שאנחנו רוצים כמות נוירונים פרופורציונלית ביחס לגודל השכבה הקודמת והבאה (748,10).

❖ **קצב למידה (eta) – 0.05** בחרנו קצב למידה נמוך יחסית על מנת שנוכל להמשיך תמיד להתקדם לעבר הירידה ולא נעבור אותה.

❖ **מספר מעברים (epoch) –** בחרנו מספר גבוה של מעברים **100** מכיוון שראינו כי כל מעבר נוסף מקרב אותנו לאחוז שגיאה נמוך יותר, אך לא נוצר overfitting (ע"י בדיקה בהגשה של התרגיל).

❖ **גודל מקבץ (batch) –** בחרנו להשתמש ב **SGD (עדכון לאחר כל דוגמה)** מכיוון שה training set מאוד גדול רצינו לבצע עדכון למשקלים גם בתוך epoch ולא רק בסיום כל אחד ובכך גם נקבל ביצועים מדויקים יותר אומנם על חשבון זמן אימון.