

Report

רקע

בתרגיל זה התבקשנו לבנות רשת נוירונים שיודעת לסווג כל תמונה לסוג הבגד המופיע בה ע"י שימוש בספריות numpy & scipy בלבד. בסט האימון התקבלו 55,000 תצפיות, כאשר כל תצפית מיוצגת כתמונה של פריט אופנה בעל 28×28 פיקסלים. עבור כל דוגמא בסט האימון קיבלנו את התיוג המתאים לו מבין התיוגים 0,1,2,3,4,5,6,7,8,9, ובסט המבחן התקבלו 5,000 דוגמאות חדשות.

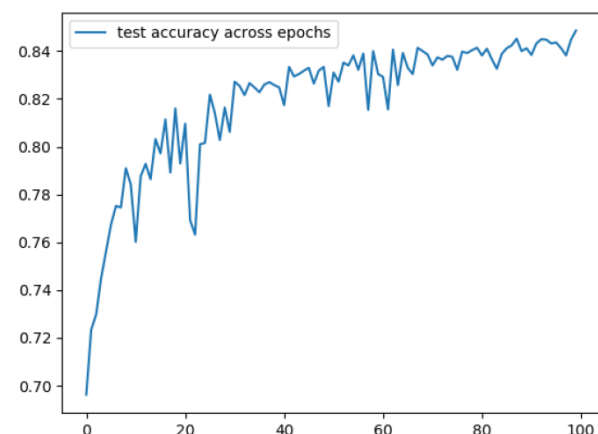
Pre-Processing

לפני הפעלת הלמידה על סט האימון נרצה לבצע נרמול לדאטא על מנת לקבל תוצאות טובות יותר. כל פיקסל מיוצג כמספר בין 0 ל-255, ולפיכך הנרמול בוצע ע"י חילוק כל פיקסל ב-255 וכך כל הערכים המספריים יהיו בטווח [0,1].

קביעת משקלים והיפר פרמטרים

ראשית, וקטור ה-BIAS וערכי המשקולות נבחרו רנדומלית בהתפלגות נורמלית. רשת הנוירונים הינה בעלת שלוש שכבות: שכבה ראשונה בעלת 784 נוירונים, שכבת חבוייה בעלת 40 נוירונים ושכבת הפלט בעלת 10 נוירונים כמספר התיוגים. לאחר בניית הרשת נעשו בדיקות לקביעת ערכי ההיפר פרמטרים שיתנו אחוזי דיוק גבוהים יותר. היו קבועים איתם הגעתי לאחוזי דיוק גבוהים תוך 30 איטרציות עם מקדם למידה גבוה, אך התוצאות עבור סט המבחן היו מאוד שונות מהרצה להרצה, ולא נתנו אחוזי דיוק מספיק טובים. לכן, נעשו בדיקות עבור מקדמי למידה קטנים יותר ומספר איטרציות גבוה יותר (300). לבסוף התברר כי לאחר 100 איטרציות ישנה התכנסות מספיק טובה ואחוזי הדיוק נשמרים (אין קפיצות גדולות בין הבדיקות). סה"כ ההיפר פרמטרים שנבחרו הם:

Eta = 0.04, Mini_batch_num = 28, Layers_size = [784,40,10], Epochs = 100



להלן גרפים המציגים את עקומת הלמידה של הרשת בהתאם להיפר פרמטרים שנקבעו. ניתן להיווכח בכלל הגרפים שלאחר 60 איטרציות ישנה התכנסות ואין קפיצות גדולות בין הבדיקות.

