

### הסבר תרגיל 3

שם: אליעד ארזואן

תז: 206482622

- את הבדיקות אצלי ביצעתי על תת סט מסט האימון כדי שהטעינה והאימון עצמו יהיו יותר מהירים. לכן שם בדקתי בעיות כלליות כדי לתקן אותן במהירות.
- בשביל שאבחר היפר פרמטרים מתאימים (פונקציה מתאימה לסיווג, מיו, מספר אפוקים...) אשתמש ב-20% מה data set כ validation set בשביל שאוכל לבדוק את loss שלי כיצד הוא משתנה כאשר אני משנה את אותם היפר פרמטרים.
- בהתחלה בקובץ לפני ההגשה עשיתי במהלך כל אפוק בדיקת שגיאות ב validation. היו מקרים שאחרי יותר מדי אפוקים זה "נדפק" ואחרי קצת זה לא היה מספיק. אחרי מספר בדיקות ראיתי שהכי מתאים לי 20 אפוקים.
- בנוסף בדקתי כיצד משתנה אם אני משתמש ב Relu או tanh או סיגמואיד לשכבה האמצעית. ראיתי שהכי מתאים לי סיגמואיד.
- בסופו של דבר ניסיתי להשתמש ב softmax בשביל לקבל את השכבה האחרונה ואז משם נקח את ה argmax.
- גזרתי בהתאם את הפונקציות לפי כלל השרשרת וקיבלתי את ה backprop. אחרי כל איטרציה עדכנתי.
- בדקתי ערכים שונים של eta בהתחלה 0.5, 0.1 והכי עבד לי טוב עם 0.01.
- בסופו של דבר אחרי בחירת הפרמטרים בחנתי על ה data set המלא ועבד טוב.
- תיקונים נוספים שהיו לי:
  - (1) נרמול הערכים – סיגמואיד נעשית בעייתית בערכים גדולים לכן חילקתי את כל הערכים של הפיקסלים ב 255.
  - (2) בחירת הפרמטרים של  $W, b$  מתפלגים יוניפורמית – במקום להגריל מספר בחרתי התפלגות יוניפורמית בין -0.8-0.8.
- בנוסף ראיתי איזה גודל hidden layer הכי עובד בשבילי וקבלתי ש 100 הכי עובד טוב