למידת מכונה – תרגיל 3

***ראגב גאזי***

***314892506***

***בהתחלה עשיתי 2 פונקציות אתחול:***

* Initialize\_data : פונקציה שמאתחלת קבצי ה train\_x , train\_y
* Initialize\_parametrs : פונקציה שבוחרת ערכים רנדומליים ל w1,b1,w2,b2

***פונקציית bprop :***

פונקציה שמחשבת transpose לכמה משתנים ואז מבצעת נגזרת בשלבים .

***פונקציית fprop\_train :***

בהתחלה עושים ערבוב לנתוני x ו y ואז עוברים על כל זוג (x,y) ומחשבים את השכבה נסתרת ואז מחשבים ה-loss ובסוף מעדכנים את הפרמטרים ומחזירים אותם.

***פונקציית predictions :***

פונקציה שמקבלת את כל הנתונים הסופיים ומייצרת קובץ וכותבת את ה-predictions בו.

והשתמשתי בכמה פונקציות עזר (שמימשתי אותם):

.Normalize, soft\_max, sigmoid

***בחירת learning\_rate :***

התחלתי ב-לאמן רשת החל מ-learning\_rate נמוך (0.01 שזה הערך המסורתי) ואז הייתי בודק את ה-loss ומגדיל את ה-learning\_rate ותוך כדי הייתי בודק את אחוז ההצלחה עבור כל ערך של ה- learning\_rate.

עשיתי הרבה בדיקות ובסוף הגעתי להחלטה לבחור את ערך הlearning\_rate להיות **0.1** שזה הביא את אחוז ההצלחה להכי גבוה (0.95).

***בחירת מספר ה-epochs :***

הסתכלתי על ה- validation וה- training losses ועקבתי אחר הערכים שלהם. אם הloss validation הולך לגדול זה אומר התאמת יתר. ואז הבנתי שכדאי להגדיר את מספר ה epochs גבוה ככל האפשר ולהימנע מהתאמת יתר אבל אם אני מגדיל את זה ככל האפשר אז הקוד לוקח הרבה זמן לרוץ ובסוף הסתפקתי במספר epochs = 7 שזה בינוני.