

בחירת hyper params:

קצב למידה,

ראיתי תוצאות טובות מאפסילון שמתחיל מ 0.01 וכל אפוק הכפלתי ב0.9 כדי למתן את הגרדיאנט

מבחינת משקלים, שכבה סודית אחת, שמכילה 128 ערכים.

מספר האפוקים, 10 נתן תוצאה טובה.

מבחינת התוכנית,

ביצעתי נורמליזציה ו shuffle לפני תחילת הלמידה, לאחריה שלחתי שורה שורה עם התוצאה הצפויה באימון  
לprop ואז הערכים שהתקבלו לprop מצאתי נגזרות ובעזרתם עידכנתי משקלים. כך רצתי מספר אפוקים  
מבוקש, ובעזרת התוצאות מצאתי את  $y_{test}$  הצפוי.

205812290