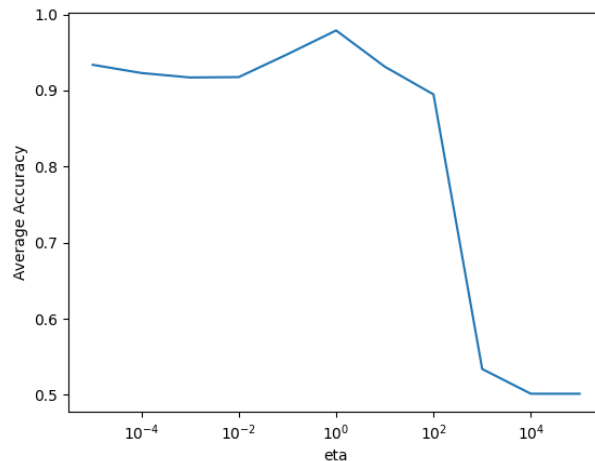


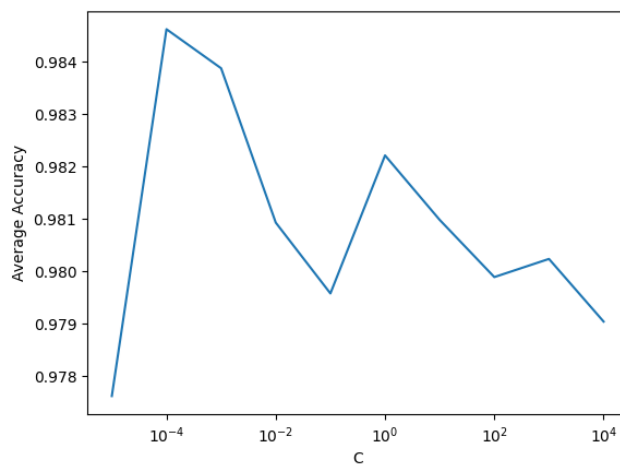
# Programing Assignment

## שאלה 1 - SGD for Hinge loss

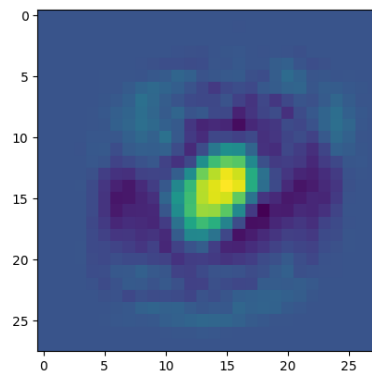
(a) בגרף ניתן לראות את הדיוק הממוצע על ה-validation set כתלות ב- $\eta_0$ . ניתן לראות שה- $\eta_0$  האופטימלי מתקבל בסביבת המספר 1 בסקלה הלוגריתמית שצוינה בשאלה. קונקרטי,  $\eta_0$  האופטימלי נמצא בטווח שבין 0.1 ל-1, ובהגדלת רזולוציה שבוצעה התקבל ש- $\eta_0$  האידיאלי הינו 0.816.



(b) הצבנו את הקירוב ל- $\eta_0 = 0.816$  האופטימלי שמצאנו בסעיף הקודם, ובאופן דומה ניתן לראות את גרף הדיוק הממוצע כתלות ב-C שמתקבל עבור אותו טווח ערכים בסקלה לוגריתמית. קונקרטי, ניתן לראות שה-C האופטימלי נמצא בסביבת  $10^{-4}$ , ובהגדלת רזולוציה שבוצעה התקבל ש-C האופטימלי הינו 0.000144.



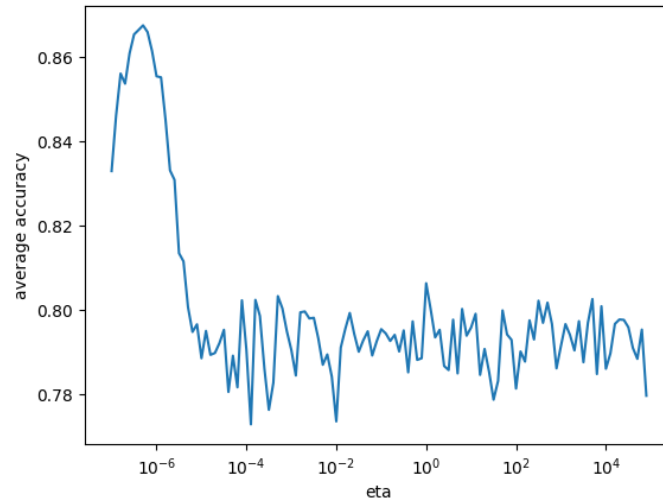
(c) נציג את  $w$  כתמונה. ניתן להבחין כי המשקל הרב ביותר יושב במרכז התמונה בצהוב, וזו אכן מקום שמפריד מהותית בין ספרות בכתיבתן וביניהן 8. מסביב למרכז הצהוב, ניתן להבחין במשקל יחסית רב ויחסית אחיד בצורה של מעגל בצבע כחול כהה, שייתכן שמחקה את ההיקף של הספרה 0.



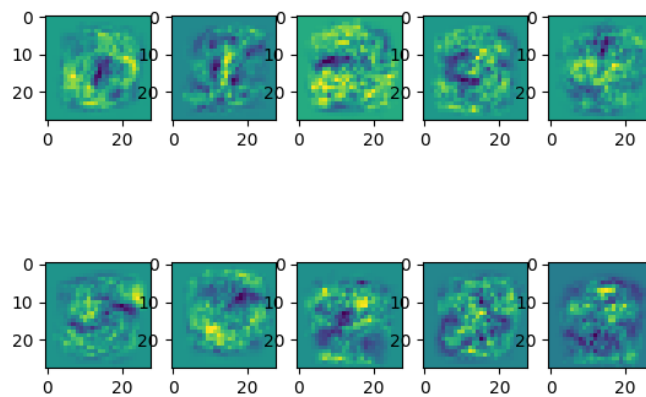
(d) בהרצת האלגוריתם עם הפרמטרים האופטימליים שמצאנו ( $\eta_0 = 0.816, C = 0.000144$ ), על ה- test set התקבל דיוק של 0.9928352098259979.

## שאלה 2 – SGD for multi-class cross-entropy

(a) בגרף ניתן לראות את הדיוק הממוצע על ה-validation set כתלות ב- $\eta_0$  בסקלה לוגריתמית קצת יותר מורחבת מאשר זו בשאלה הקודמת. קונקרטית, בהגדלת רזולוציה ניתן לראות שה- $\eta_0$  האופטימלי מתקבל בסביבת המספר  $10^{-6.2}$ .



(b) ניתן לראות בתמונות צורות "מטושטשות" לפי המשקלים שיכולות להזכיר את הצורה של הספרות עצמן.



(c) בהרצת המסווג האופטימלי שמצאנו על ה-test set-ה-0.8835 דיוק של 0.8835.