

למידה חישובית תרגיל בית מספר 4

מטרת התרגיל היא מימוש ושימוש של רשת עצבית פשוטה לפי הנלמד בהרצאות, וכן בנייה ושימוש ברשתות באמצעות tensorflow ו-keras כדי להבחין בתמונה בין פנים של אנשים לעומת פרטים אחרים שלא מכילים פנים (face detection), וכן סווג של נתוני MNIST ו- Fashion_MNIST.



1. (70 נקודות) בתרגיל זה נשתמש ברשת tensorflow FCN.

א. (2 נקודות) טענו את נתוני האימון מתוך face_data.xlsx.

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
df = pd.read_excel('face_data.xlsx')
M = df.to_numpy()
MT = M.T
X = MT[:1,1000:5000]
Y = MT[784,1000:5000]
Y[Y==-1]=0
Y = Y.reshape(1,Y.shape[0])
```

כפי שאפשר לראות בקובץ ה-Excel, בקבוצת הנתונים המקורית יש 10000 תמונות, מתוכן 3000 הראשונות הן תמונות של פנים, והשאר (7000) הן תמונות שאינן פנים. לתמונות המקוריות נעשה עיבוד מקדים, הן נשמרו כתמונות ריבועיות של 28x28, וכן נמצאו תכונות מבחינות על-ידי שימוש בשיטת HoG. לאחר מכן נעשה שרשור העמודות, ליצירת וקטור שורה מכל תמונה של 784 תכונות.

ב. (3 נקודות) השתמשו ב-reshape וכתבו פונקציה המאפשרת להציג כל אחד מהוקטורים כתמונה.

ג. (5 נקודות) באמצעות פקודת imshow הציגו 25 תמונות של "פנים" ו-25 תמונות של "לא פנים", על-ידי שינוי הממדים של כל וקטור 784 ממדי למטריצה של 28x28.



ד. (5 נקודות) באמצעות Model Sequential של Keras והקוד שלמדנו לבניית רשת לזיהוי ספרות בכתב יד, יצרו רשת עם שכבת hidden אחת והתאימו את הפרמטרים של הרשת עבור ה-data.

ה. (5 נקודות) אמנו את הרשת באמצעות נתוני האימון. השתמשו ב-cross-validation וקבעו את מספר היחידות בשכבת ה-hidden, את מקדם הלמידה alpha, ואת דרגת הרגולריזציה.

ו. (5 נקודות) עבור הרשת עם הפרמטרים עבורם הביצועים הם הטובים ביותר בחנו את דיוק הסווג על נתוני המבחן בקובץ face_test.mat חוו את דעתכם.
כדי להעלות mat file ל-python אפשר להשתמש ב:

```
import scipy.io as sio
X_face_test = sio.loadmat(face_test.mat)
נוצר מילון, ואפשר לרשום את וקטורי הפנים ואת התיוגים באמצעות:
Xtest = X_face_test['Xtest']
ytest = X_face_test['ytest']
```

ז. (5 נקודות) הוסיפו לרשת שכבת hidden נוספת וחזרו על הסעיפים הקודמים.
ח. (5 נקודות) בחנו איזה רשת תניב תוצאות יותר טובות – רשת עם שכבת hidden אחת עם 64 יחידות, או רשת עם 5 שכבות hidden של 16, 16, 16, ו-8 יחידות (כל אחת עם פונקציות אקטיבציה relu מלבד האחרונה עם פונקציית אקטיבציה סיגמואידית).
ט. (10 נקודות) עתה ניצור רשת נוספת באמצעות Model Sequential עבור נתוני fashion_mnist, המכילים עשרה סוגים של פרטי לבוש ואופנה. המטרה היא לבנות רשת שתוכל לסווג כל פריט לבוש לאחת מעשר מחלקות: חולצת T, מכנסיים, שמלה, סנדלים, מעיל וכו'.

טענו את הנתונים באמצעות שימוש ב-script הנמצא במחיצה Materials for exercise 4 – NN: Fashion_MNIST_load_data.py. תכננו רשת עם שתי שכבות hidden, אמנו את הרשת על נתוני האימון ובחנו קונפיגורציות שונות (כלומר מספר היחידות בכל אחת מהשכבות, סוג פונקציות האקטיבציה) על נתוני ה-validation. בחרו את הקונפיגורציה הטובה ביותר על נתוני ה-validation והפעילו את הרשת על

נתוני המבחן. הדפיסו את אחוז הזיהויים הנכונים עבור נתוני האימון, האימות (validation) ונתוני המבחן.



2. (30 נקודות) נתונה רשת עם שכבת קלט של שתי יחידות (כל דוגמת אימון עם שתי תכונות), שכבת Hidden עם שלוש יחידות ושכבת פלט עם שלוש יחידות לסווג מזג האוויר לשלושה מצבים אפשריים - "שמש", "מעונן" או "גשום". הניחו כי פונקציית האקטיבציה עבור כל היחידות בשכבת ה- hidden היא שכבת ReLU ובשכבת הפלט היא פונקציית sigmoid.

כתבו באופן מפורש את חישובי כל ה- $\delta_i^{(l)}$ בכל אחת מהשכבות, תחילה עבור כל i בכל אחת מהשכבות, ולאחר מכן הציגו את החישובים באופן וקטורי-מטריצי.

כתבו באופן מפורש את חישובי כל הגרדיאנטים, כלומר את חישובי $\frac{\partial J(\theta)}{\partial \theta_{ij}^{(l)}}$ לכל i ו- j עבור כל אחת מהשכבות הרלבנטיות, ולאחר מכן רשמו את החישובים באופן מטריצי-וקטורי.

3. (10 נקודות) בתרגיל זה נכתוב פונקציות למימוש רשת עצבית עם מספר שכבות Hidden. בהמשך לפונקציות שכתבנו במעבדה - השלימו את הרשת שהתחלנו לכתוב – כתבו פונקציה למימוש ה- backpropagation עם רגולריזציה.

א. עדכנו את הפונקציות feed_forward ו- back_propagation כך שיוכלו לייצר מספר כלשהו של שכבות hidden (ולא רק שכבת hidden אחת).

ב. הוסיפו לקוד שתי פונקציות אקטיבציה נוספות: tanh ו- ReLU. עדכנו את שכבות ה- hidden כך שהמשתמש יוכל לבחור בין שימוש בפונקציית אקטיבציה ReLU, sigmoid או tanh. יש לעדכן את שתי הפונקציות feed_forward ו- back_propagation (הנגזרת של ReLU ושל tanh שונה מהנגזרת של פונקציית ה- sigmoid).

ג. אמנו את הרשת עם שכבת hidden אחת, כאשר מספר היחידות בשכבה זו הוא 32, על 7200 דוגמאות אימון מתוך Mnist (מאגר המידע עליו עבדנו במעבדה עם רזולוציה של 28x28) ובחנו על 3000 דוגמאות מבחן, כאשר כל פונקציות האקטיבציה בכל השכבות הן סיגמואיד. חזרו על ההרצה עבור ערכי alpha ומספר יחידות hidden שונים ומצאו צירוף שיביא לתוצאה הטובה ביותר עבור קבוצת המבחן.

ד. חזרו על האימון והמבחן עבור פונקציית אקטיבציה ReLU בשכבת ה- hidden (פונקציית סיגמואיד ביחידות הפלט) והשוו את התוצאות לסעיף הקודם. שנו את ערכי ה- α ומספר יחידות ה- hidden לקבלת התוצאה הטובה ביותר עבור קבוצת המבחן.

ה. חזרו על האימון והמבחן עם שתי שכבות hidden כאשר כל היחידות בשכבות אלה הן עם פונקציית אקטיבציה ReLU ופונקציית סיגמואיד ביחידות הפלט.