

מטלת מנחה (ממ"ן) 12

הקורס: מבוא לראייה ממוחשבת – 22928

חומר הלימוד למטלה: יחידות 4-6

סמסטר: 2023א

הערות:

להגשה – קובץ ZIP הכולל:

- דו"ח בפורמט PDF עם תשובות לכל שאלה בתוספת איורים כנדרש.
- יש לתאר בדו"ח את כל תהליך העבודה עם הפרמטרים שהשתמשתם בהם.
- בנוסף, יש לצרף את תוכניות ה-Python - עם הערות בגוף התוכנית ועם קישור ברור לשאלה המתאימה.
- הקוד לא בא להחליף את הדו"ח ואין להדביק קטעי קוד בדו"ח.

חלק 1: MNIST –

בחלק הראשון של המטלה נתרגל אלגוריתמים לסיווג שהוצגו ביחידות 4-6, ע"י סיווג וניתוח של

תמונות של ספרות מתוך MNIST dataset.

(לקריאה נוספת - https://en.wikipedia.org/wiki/MNIST_database)

הקובץ עם התמונות נמצא באתר, הוא מכיל 3 קבוצות של תמונות, אחת כוללת 50K תמונות,

משמשת לאימון, השנייה והשלישית מכילות 10K ומשמשות ל-TEST ול-VLIDATION
במטלה נשתמש בסט האימון והבדיקה בלבד.

הקריאה ל dataset -נעשית ע"י:

```
with gzip.open('mnist.pkl.gz', 'rb') as f:
    u = pickle._Unpickler(f)
    u.encoding = 'latin1'
    train_set, valid_set, test_set = u.load()

# separate to data and labels:
x_train, y_train = train_set
x_test, y_test = test_set
```

סט האימון מכיל מטריצה בגודל 50000x768 שכל שורה בה היא תמונה של סיפרה אחרת, ווקטור באורך 50,000 שמכיל את ה label -או הסיפרה שמופיעה באמת בתמונה המתאימה. סט הבדיקה מכיל מטריצות בעלות מבנה דומה עם 10K שורות.

כדי להציג את התמונה ה- 100 מתוך סט האימון:

```
# show sample #100 from the train set
i = 100
im = x_train[i].reshape([28,28])
label = y_train[i]

plt.figure()
plt.imshow(im, cmap='gray', interpolation='nearest')
plt.title(label)
plt.show()
```

סט הנתונים מכיל תמונות של 10 ספרות. כתבו כמה תמונות יש מכל סיפרה? הציגו את 12 הספרות הראשונות מתוך ה train set - figure אחד (מחולק לשלוש שורות וארבע עמודות).

שאלה 1 - KNN

הדרך הפשוטה ביותר לסיווג היא ע"פ הדוגמה (או הדוגמאות) הדומות, מה שניקרא – Nearest Neighbour Classifier

בשאלה זו נבדוק את הביצועים של מסווג KNN כאשר K הוא מספר השכנים. א. חשבו את ביצועי מסווג KNN עבור $k=1...10$ ציירו את התוצאות על גרף. מה התוצאה האופטימלית? איך מתבצעת ההחלטה כאשר k זוגי?

שאלה 2 - PCA

בשאלה זו נבדוק את השימוש של PCA לצורך סיווג תמונות של מספרים כתובים בכתב יד מתוך MNIST.

- א. חשבו את טרנספורמציה PCA על סט האימון. ציירו את הסיפרה הממוצעת ואת 6 ה principle components - הראשונים.
- ב. ציירו גרף של הווריאנס הכולל ש"מוסבר" ע"י ה הרכיבים הראשיים.
- ג. כמה בסיסים צריך בשביל להגיע ל- 95% ווריאנס? כמה ל- 80%?
- ד. הטילו את הספרות למימד 2 וצייר את הווקטורים המתקבלים (נקודות על המישור, scatter plot) כל סיפרה בצבע אחר.
- ה. חזרו על שאלה 1 כאשר כל סיפרה מיוצגת ע"י ההטלה שלה למימד 2, 10 ו- 20.
- ו. עבוד סיפרה כלשהי, הטילו את הסיפרה למימד k' ושחזרו אותה בחזרה. ציירו את השיחזור עבור 150, 100, 50, 10, 5, 2 $k=$ כתבו את הנוסחה שבה השתמשתם להטלה ולשיחזור.

- ז. הבעיה העיקרית בשימוש ב PCA - היא הקושי בתיאור של dataset מורכב, כמו במקרה של שלנו. הפתרון הוא לחשב טרנספורמציה PCA לכל מחלקה/סיפרה בנפרד בעזרת סט האימון. ולמדוד את ההתאמה של סיפרה חדשה לכל אחד מהמודלים.
1. חשבו מודל PCA לכל סיפרה בנפרד. הציגו את 6 ה principle components - של כל מודל.
 2. מה ההבדל יחסית למודל של כל הספרות יחד?
 3. חשבו את ההטלה של התמונות מה test set - ל כל אחד מהמודלים (סה"כ 10 הטלות לכל תמונה).
 4. שחזרו את התמונות מכל אחת מההטלות (סה"כ 10 תמונות משוחזרות מכל תמונה). ציירו את כל השיחזורים של תמונה אחת לדוגמה.
 5. חשבו את המרחק בין השיחזורים לתמונה המקורית.
 6. המודל ששיחזר "הכי טוב" ייבחר כמודל הנכון.
- מה הביצועים של מסווג כזה?

חלק 2 – סיווג ע"י BOW

- בחלק זה של העבודה נממש מסווג מסוג SVM שיסווג תמונות לשתי מחלקות, כאשר התמונות מיוצגות ע"י (BOW) Bag Of Words
- בשאלה זו נשתמש בתמונות מתוך ה dataset -הזה:
- http://people.csail.mit.edu/torralba/code/spatialenvelope/spatial_envelope_256x256_static_8outdoorcategories.zip
- המכיל תמונות מ- 8 קטגוריות שונות כגון - הרים, חופים, יער. השייכות של כל תמונה לקטגוריה המתאימה נתונה בשם הקובץ. יש לבחור שתי קטגוריות מתוך השמונה. המטרה היא לבנות מסווג שיידע לסווג תמונה לקטגוריה הנכונה, כשהתמונה מיוצגת ע"י BOW.
- חלקו את הנתונים לנתוני אימון, TRAIN-SET ונתוני בדיקה TEST-SET ביחס של 4:1, כלומר 80% מהתמונות ישמשו לאימון ו- 20% לבדיקה.
- שלב האימון – יתבצע על ה TRAIN-SET -בלבד:
1. חישוב מאפייני dense-SIFT מכל התמונות. (אפשר להשתמש בפונקציה מתוך OpenCV לחישוב SIFT בנקודות מוגדרות מראש).
 2. Vector Quantization – חישוב K-means על סט המאפיינים. מספר המרכזים צריך להיות גדול מספיק (לפחות 100) ונתון לשיקולכם.
 3. לכל תמונה חישוב היסטוגרמה של מספר המאפיינים השייכים ל-cluster ה-k.

4. אימון מסווג מסוג Support Vector Machine (SVM) לינארי כאשר כל תמונה מיוצגת ע"י היסטוגרמה שחושבה ב-(3) ולכל תמונה מוצמד label שמשייך אותה למחלקה המתאימה.

כתוצאה משלב האימון, המודל מורכב מהמרכזים שחושבו בשלב ה-VQ ע"י K-means והמסווג שאומן בשלב (4).

שלב הבדיקה – יתבצע על test-set בלבד לכל תמונה –

1. חישוב מאפייני SIFT.
 2. Vector Quantization – לכל מאפיין חישוב המרכז הקרוב ביותר.
 3. חישוב היסטוגרמה של מספר המאפיינים בכל מרכז.
 4. סיווג ע"י המסווג שאומן בשלב האימון.
- מה הפרמטרים האופטימליים של המסווג? יש להציג את הביצועים עבור מספר ערכים של הקבוע C ע"י עקומות ROC וחישוב ה-AUC. יש להציג את הביצועים של המסווג ע"י ROC ו-AUC. השיבו: איך ישתנו הביצועים עבור חלוקות שונות לtrain ולtest? מה התוצאה הממוצעת?