מכון טכנולוגי חולון
Holon Institute of Technology

סמסטר קיץ – תשע"ז, מועד ג'
תאריך: 10.01.18
המרצה: דמיטרי פטשוב
קבוצות הרצאה: 0-21207-62/63

מבחן בקורס Big Data, למידת מכונה ומולטימדיה בסביבת Python

הוראות:

1. המבחן נעשה במחשב.
2. אין שימוש בחומר עזר.
3. מותר להשתמש במחשבון, דף ריק מתוכן וכלי כתיבה.
4. משך הבחינה - 3 שעות
5. במבחן שלוש חלקים:
   - חלק ראשון חובה - יש לענות על כל השאלות.
   - חלק שני בחירה - יש לענות על שתי שאלות מתוך שלוש.
   - חלק שלישי בונוס - לא חובה לענות על חלק זה.
6. אין לעזוב את המבחן לפני שהוגש כנדרש (אחרת הציון 0 אוטומטית)
7. שימו לב! לא ליצור את הפרויקט בקונן C (אם קיים קונן נוסף) אחרת יש סיכוי לעובדן מידע.
8. שימו לב! תוודאו שאתם עובדים בגרסת פייתון נכונה.
9. יש ליצור קובץ פונקציות אחד בלבד ובו לממש את כל הפונקציות הנדרשות.
10. כל פונקציה שאתם מתבקשים לכתוב חייבת להיות בעלת שם זהה לזה שבשאלה! כולל התייחסות לאותיות קטנות/גדולות. פונקציה שהשם שלה לא יהיה רשום במדויק, לא תקבל ניקוד.
11. אין להגיש שום סקריפט שלא התבקשתם לכתוב וזה כולל סקריפט הבדיקות שלכם.
12. כל קטע קוד שלא קשור לאחת הפונקציות שהתבקשתם לכתוב יגרור הורדה בציון.
13. שימו לב! יש להתחשב במקרי הקצה שעלולים להיות. לדוגמה קלט לא תקין או כתובת שלא קיימת.
14. אתם יכולים להשתמש בכל פונקציה שתרצו לפתרון המבחן.
15. במהלך המבחן אין גישה לאינטרנט.
16. מי שמסיים את המבחן מחכה במקומו עד להגעת המרצה בכדי להגיש את הקוד.
17. יש לרשום את מספרי ת.ז. על גבי טופס זה ולהגיש אותו בסוף המבחן.

מכון טכנולוגי חולון
Holon Institute of Technology

חלק א' (חובה): 50%

שאלה 1 (7 נק.)

1. הגדירו פונקציה בשם : **myIdentityMatrix**
הפונקציה צריכה לקבל מספר טבעי n.
הפונקציה צריכה להחזיר מטריצת יחידה ממימד n.

דוגמה :

n = 3
myMat = **myIdentityMatrix**(n)

תוצר הפונקציה יהיה :

$$\text{myMat} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

שאלה 2 (10 נק.)

2. הגדירו פונקציה בשם : **myArrayMean**
הפונקציה צריכה לקבל רשימה ולחשב ממוצע של כלל המספרים שישנם ברשימה.
שימו לב מספרים יכולים להיות בעלי ייצוגים שונים (int, float32, float64, etc...).
עליכם להכליל בממוצע את כלל המספרים בעלי ייצוג של מספר. כמו כן, אין להכליל בממוצע את המספרים בעלי ייצוג שאינו מספר (string, tuple, list, etc...).
הפונקציה צריכה להחזיר את הממוצע של כלל המספרים ברשימה (לא כולל תתי רשימות).

לדוגמה :

myList = [1, [3.1], 2.0, 'number', (13, 3), ['hello', 8], numpy.uint8(12), '20']
myMean = **myArrayMean(myList)**

תוצר הפונקציה יהיה :

$$\text{myMean} = \frac{1+2+12}{3} = 5$$

פונקציות עזר :

numbers.Number

שאלה 3 (10 נק.)

3. הגדירו פונקציה בשם :**myKNearest**
הפונקציה מקבלת מספר טבעי k, מספר ממשי כלשהו ומערך חד-מימדי מסוג ndarray המכיל מספרים בעלי יצוג float64.
הפונקציה צריכה להחזיר מערך חד-מימדי מסוג ndarray המכיל k איברים הקרובים ביותר בגודלם למספר הממשי שנשלח לפונקציה.

לדוגמה :

k = 3
x = -1.2
myArray = numpy.array([-3.2, 8, 1, 0, -1.1, 0.1, 12])
myKN = **myKNearest(k, x, myArray)**

תוצר הפונקציה יהיה :

myKN = [-1.1, 0, 0.1]

מכון טכנולוגי חולון
Holon Institute of Technology

שאלה 4 (10 נק.)

4. הגדירו פונקציה בשם : **myDirLongName**

הפונקציה צריכה להחזיר את שם הקובץ (string) הארוך ביותר בתיקיית העבודה (תיקיה בה יצרתם את הפרויקט).

לדוגמה :

myFileName = **myDirLongName**()

נניח שבתיקיית העבודה שלי ישנם שלוש קבצים : functions.py, main.py, exam.txt

תוצר הפונקציה יהיה :

myFileName = ' functions.py '

פונקציות עזר :

os.getcwd(), os.listdir(myFolder), os.path.isfile(myFullPath)

שאלה 5 (13 נק.)

5. הגדירו פונקציה בשם : **myFastToneReplacement**

להלן הפונקציה :

```
def myFastToneReplacement(img, fromA, toB, equalsC):
    myImg = img.copy()
    # ----------------------------------------------------------------
    D = myImg.shape
    if len(D) == 2:
        for i in range(D[0]):
            for j in range(D[1]):
                if myImg[i, j] >= fromA and myImg[i, j] <= toB:
                    myImg[i, j] = equalsC
    elif len(D) == 3:
        for m in range(D[0]):
            for n in range(D[1]):
                for k in range(D[2]):
                    if myImg[m, n, k] >= fromA and myImg[m, n, k] <= toB:
                        myImg[m, n, k] = equalsC
    else:
        return None
    # ----------------------------------------------------------------
    return myImg
```

תחליפו את קטע הקוד שבין הקווים השבורים בשורת קוד אחת בלבד.

פונקציות עזר :

numpy.logical_and(boolMat1, boolMat2)

מכון טכנולוגי חולון
Holon Institute of Technology

חלק ב' (בחירה): 50%

שאלה 6 (25 נק.)

6. הגדירו פונקציה בשם : **myMedianFilt**

הפונקציה צריכה לקבל תמונה בגוווני אפור וסדר הפילטר. תמונה חייבת להיות מסוג ndarray. סדר הפילטר חייב להיות מספר טבעי.
הפונקציה צריכה להחזיר את התמונה המפולטרת על ידי מסנן החציון.
גודל התמונה המתקבלת חייב להיות זהה לזה של התמונה המקורית. הקצבות של התמונה צריכים להיות מוגדרים נכון.
תזכורת :

דוגמה :

medImg = **myMedianFilt(img, filtOrder)**

פונקציות עזר :

np.reshape(myMat, myShapeTuple) np.sort(myArr)
np.median(myMat) np.zeros(myShapeTuple) cv2.imread(myStr, 0)
cv2.copyMakeBorder(myMat, top, bottom, left, right, cv2.BORDER_REPLICATE)

שאלה 7 (25 נק.)

7. הגדירו פונקציה בשם : **myPhaseRound**

הפונקציה צריכה לקבל מטריצה (מערך דו מימדי) מסוג ndarray בעלת ערכים בין 0 ל 180.
הפונקציה צריכה להחזיר מטריצה בה כל איבר ממטריצה שנקלטה מעוגל לערך הקרוב ביותר מבין : 0, 45, 90, 135, 180. כל ערך שמעוגל ל 180 צריך להחליף ב 0.
לפי כך, הפונקציה צריכה להחזיר מטריצה בעלת ערכים {0, 45, 90, 135} בלבד.

דוגמה :

PhaseMat = **myPhaseRound(myMat)**

פונקציות עזר :

np.logical_and(BoolMat1, BoolMat2), np.logical_or(BoolMat1, BoolMat2)
np.random.uniform(lowVal, highVal, ShapeTuple)

שאלה 8 (25 נק.)

8. הגדירו פונקציה בשם : **myHistEq**

הפונקציה צריכה לקבל תמונה צבעונית. תמונה חייבת להיות מסוג ndarray. התמונה חייבת להיות ברזולוציה של 8 ביט.

הפונקציה צריכה להחזיר תמונה שעברה השוואת היסטוגרמה (Histogram equalization).
תזכורת :

שלבים של האלגוריתם :

1. חישוב היסטוגרמה
2. חישוב פונקצית צפיפות ההסתברות
3. חישוב פונקצית הסתברות מצטברת
4. שימוש בפונקצית הסתברות מצטברת כטבלת חיפוש (Lookup table)

$$\boldsymbol{PDF} = \frac{\boldsymbol{Hist}}{\sum \boldsymbol{Hist}}$$

$$\boldsymbol{CDF(k)} = \sum_{\boldsymbol{n=0}}^{\boldsymbol{k}} \boldsymbol{PDF(n)}$$

$$\boldsymbol{newImg[i,j,k]} = \boldsymbol{CDF(Img[i,j,k])} * \mathbf{255}$$

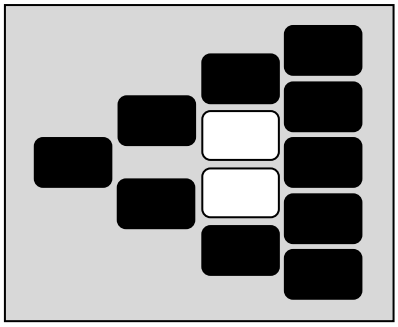
מכון טכנולוגי חולון
Holon Institute of Technology

חלק ג' (בונוס): 15%

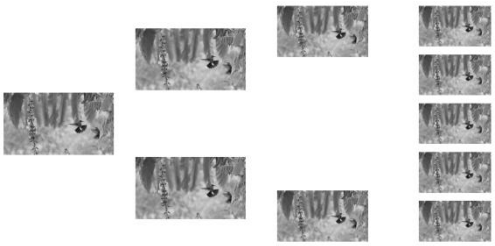
שאלה 9 (15 נק.)

9. הגדירו פונקציה בשם : **myPlotShape**
הפונקציה צריכה לקבל תמונה בגווני אפור. תמונה חייבת להיות מסוג ndarray.
הפונקציה צריכה לפלוט גרף של ספריית matplotlib שנראה כך :

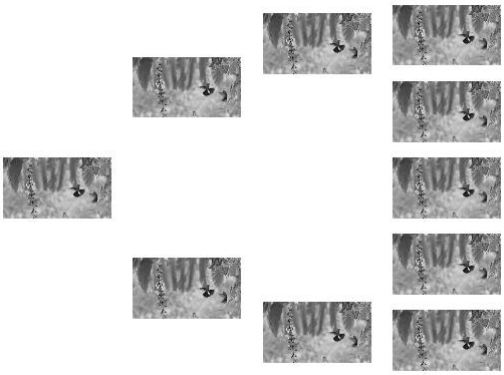
סקיצה :



דוגמה א :



דוגמה ב :



שימו לב ששתי התמונות לדוגמה בעצם אותה התמונה. יש הבדלים בתצוגה בגלל גודל החלון ולכן תצוגה יכולה להיות שונה במסכים או מחשבים שונים.
סקיצה מתארת לכם איך התמונות צריכות להיות ממוקמות בתוך החלון, כאשר שחור זו תמונה ולבן זה מקום ריק מתמונה.

בהצלחה!