מכון טכנולוגי חולון
Holon Institute of Technology

סמסטר א – תשע"ט, מועד ב'
תאריך: 24.03.2019
המרצה: דמיטרי פטשוב
קבוצות הרצאה: 0-21210-1/2

אלגוריתמים במולטימדיה ולמידת מכונה בסביבת פייתון

מס' זהות:

הוראות:

1. המבחן נעשה במחשב.
2. אין שימוש בחומר עזר.
3. במהלך המבחן אין גישה לאינטרנט.
4. מותר להשתמש במחשבון וכלי כתיבה.
5. משך הבחינה - 3 שעות
6. יש לרשום את המספר ת.ז. על גבי טופס זה ולהגיש אותו בסוף המבחן.
7. 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.
8. מי שמסיים את המבחן מחכה במקומו עד להגעת אוסף הבחינות בכדי להגיש את הקוד (במידה ועזבתם לפני, הציון שתקבלו יהיה 0).
9. הימנעו מליצור את הפרויקט בכונן C (במידה ומתאפשר), אחרת יש סיכוי לעובדן מידע.
10. שימו לב! תוודאו שאתם עובדים בגרסת פייתון נכונה.
11. יש ליצור קובץ פונקציות אחד בלבד ובו לממש את כל הפונקציות הנדרשות.
12. לקובץ הפונקציות שהנכם מגישים יש לקרוא בשם המתחיל באות "s" קטנה וללא רווח מס' הת.ז. שלכם. לדוגמה: s123456789
13. כל פונקציה שאתם מתבקשים לכתוב חייבת להיות בעלת שם זהה לזה שבשאלה! כולל התייחסות לאותיות קטנות/גדולות. פונקציה שהשם שלה לא יהיה רשום במדויק כמו בטופס הבחינה, לא תקבל ניקוד.
14. שימו לב! כל פונקציה שקורסת לא תקבל ניקוד.
15. אין להגיש שום סקריפט שלא התבקשתם לכתוב וזה כולל סקריפט הבדיקות.
16. כל קטע קוד שלא קשור לאחת הפונקציות שהתבקשתם לכתוב יגרור הורדה בציון וזה כולל פלטים שלא התבקשתם לעשות.
17. שימו לב! יש להתחשב במקרי הקצה שעלולים להיות. לדוגמה קלט לא תקין או כתובת שלא קיימת. במקרה של שגיאה יש להחזיר None
18. אתם יכולים להשתמש בכל פונקציה שתרצו לפתרון המבחן.
19. במידה וישנה פונקציה שאינכם יודעים מה יעודה, נסאו פקודה:
    print FunName.__doc__

מכון טכנולוגי חולון
Holon Institute of Technology

פונקציות עזר אפשריות

Basic functions:
isinstance, id, hex, ord, str, eval, enumerate, list, range, raw_input, tuple, type, zip, chr, cmp, in, not, or, and; **String operations:** count, endswith, find, isdecimal, isdigit, isnumeric, join, lower, upper, replace, split; **List operations:** append, extend, insert, remove, pop, count; **Dictionary operations:** has_key, items, keys, values; **Set operations:** issubset, union, intersection, difference, symmetric_difference, copy, add, remove, discard, pop

copy:
copy

os:
getcwd, listdir, path.isfile, path.exists, path.join, path.isdir

numbers:
Number

numpy:
zeros, ones, eye, tri, mean, median, sum, prod, dot, transpose, trace, ceil, floor, round, max, min, argmax, argmin, sort, argsort, reshape, concatenate, all, any, pi, asarray, copy, logical_or, logical_and, logical_not, logical_xor, array_equal, isnan, linspace, arrange, mod, exp, log, sin, cos, tan, sinc, sinh, cosh, tanh, arcsin, arccos, arctan, arctan2, arcsinh, arccosh, arctanh, intp, int8, int16, int32, int64, uint8, uint16, uint32, uint64, float16, float32, float64, complex64, complex128, random.rand, random.randn, random.randint, random.random, random.shuffle, random.normal, random.uniform, ndarray.astype, linalg.eig, linalg.eigh, linalg.eigvals, linalg.norm, linalg.det, linalg.matrix_rank

cv2:
copyMakeBorder, BORDER_REPLICATE, BORDER_CONSTANT, BORDER_REFLECT, filter2D, split, merge, cvtColor, COLOR_BGR2GRAY, COLOR_GRAY2RGB, imread

matplotlib.pyplot:
figure, subplot, subplot2grid, plot, stem, scatter, show, bar, hist, hist2d, pie, psd, imshow, axis, grid close, colorbar, draw, gca, gcf, legend, loglog, semilogx, semiology, pause, title, xlabel, ylabel, xlim, ylim, xticks, yticks

מכון טכנולוגי חולון
Holon Institute of Technology

חלק א' (חובה): 50%

שאלה 1 (5 נק.)

1. הגדירו פונקציה בשם : **mySquaredMatrix**
הפונקציה צריכה לקבל מספר טבעי n.
הפונקציה צריכה להחזיר מטריצה ריבועית (מערך דו-מימדי) ממימד 2n בה יש אחדים ברביעה שמאלי תחתון וברביעה ימיני עליון.

דוגמה :

n = 2
myMat = **mySquaredMatrix** (n)

תוצר הפונקציה יהיה :

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

שאלה 2 (10 נק.)

2. הגדירו פונקציה בשם : **countMyStrings**
הפונקציה צריכה לקבל מחרוזת.
הפונקציה צריכה להחזיר מילון המכיל את כל המילים הקיימות במחרוזת ללא חזרות וללא חשיבות לאותיות גדולות/קטנות בתור המפתחות של המילון וערך עבורן יהיה מספר הופעות של כל אחת מהמלים במחרוזת שהתקבלה, כולל מקרים בהם מילה היא חלק ממילה אחרת.

לדוגמה :

myStr = 'Hello Hello We Do Hi no No no we Donot Hey HE he heY'
myDict = **countMyStrings(myStr)**

תוצר הפונקציה יהיה :

myDict :

hello : 2,	we : 2,	do : 2,	hi : 1
no : 4,	donot : 1,	hey : 2,	he : 6

שאלה 3 (10 נק.)

3. הגדירו פונקציה בשם: **myListMedian**
הפונקציה מקבלת רשימה המכילה מספרים ורשימות נוספות במבנה דומה לרשימה עצמה.
הפונקציה צריכה להחזיר חציון של כל האיברים הנמצאים ברשימה שהתקבלה כולל כל תתי הרשימות.

לדוגמה :

myList = [1, [2], 3, [11, -1, [2, 3, [1.0]], 2], -3, 1.0, [[0]]]
myMedian = **myListMedian(myList)**

תוצר הפונקציה יהיה :

myMedian = 1.5

מכון טכנולוגי חולון
Holon Institute of Technology

שאלה 4 (10 נק.)

4. הגדירו פונקציה בשם: **UpFolderContain**
הפונקציה לא מקבלת שום קלט.
הפונקציה צריכה להחזיר רשימה של שמות כל הקבצים והתיקיות הנמצאים בתיקיה בה נמצאת תיקיית העבודה של הפרויקט בו אתם עובדים.
לדוגמה:

myList = **UpFolderContain()**

שאלה 5 (15 נק.)

5. הגדירו פונקציה בשם: **myColorShift**

```
def myColorShift(img):
    if img.__class__ != np.ndarray:
        return None

    im = img.copy()
    D = im.shape
    if len(D) != 3:
        return None

    # ----------------------------------------------------------
    for m in range(D[0]):
        for n in range(D[1]):
            if im[m,n,0] > 200:
                im[m,n,1] = im[m,n,0]
    # ----------------------------------------------------------

    return im
```

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

מכון טכנולוגי חולון
Holon Institute of Technology

חלק ב' (בחירה): 50%

שאלה 6 (25 נק.)

6. הגדירו פונקציה בשם : **myLRcycles**

הפונקציה צריכה לקבל מטריצת המידע (ndarray), וקטור עמודה של תיוגים (ndarray), מטריצת משקלים W ההתחלתית (ndarray), פרמטר גמא (מספר חיובי), ומספר איטרציות (מספר טבעי).

הפונקציה צריכה להחזיר את מטריצת ה- W לאחר התהליך של רגרסיה לוגיסטית בעלת מספר חזרות מוגדר מראש.

תזכורת :

עדכון מטריצת W :

$$\boldsymbol{w_{k+1} = w_k + \gamma_k \frac{\partial l(w)}{\partial w_k}}$$

קירוב לנגזרת החלקית :

$$\boldsymbol{\frac{\partial l(w)}{\partial w} = \sum_{n=1}^{N} \left[x_n \cdot \left(y_n - g(x_n w) \right) \right]}$$

הגדרת הסיגמואיד :

$$\boldsymbol{g(z) = sig(z) = \frac{1}{1 + e^{-z}}}$$

W = **myLRcycles(**TrainData, TrainLabel, W, gamma, iterNum**)**

שאלה 7 (25 נק.)

7. הגדירו פונקציה בשם : **myGradDescent**

הפונקציה צריכה לקבל אות חד מימדי וזמן אשר מוגדרים על ידי מערכים של ndarray, נקודת התחלה שהיא ערך של זמן, מספר איטרציות וגודל צעד.

הפונקציה צריכה להחזיר את ערך הזמן אליו הגיע האלגוריתם לאחר כמות האיטרציות שניתנה עם גודל צעד הנתון.

יש להשתמש בגרדיאנט רובסטי. משמע, נגזרת מסדר ראשון אשר פחות רגישה לרעשים.

תזכורת :

$$\boldsymbol{t_{k+1} = t_k - \gamma_k \frac{\partial f(t)}{\partial t_k}}$$

דוגמה :

$t_n =$ **myGradDescent(**$signal, time, t_0, iterNum, stepSize$ **)**

מכון טכנולוגי חולון
Holon Institute of Technology

שאלה 8 (25 נק.)

8. הגדירו פונקציה בשם : **myDiagEdgeDetect**

הפונקציה צריכה לקבל תמונה 8-ביט ריבועית ופרמטר מחצית עובי הקו rad (מספר חיובי).
הפונקציה צריכה להחזיר תמונה ריבועית, 8-ביט שהתקבלה עם קו אלכסוני בו חושבה
עוצמת הגרדיאנט (Intensity Gradient). זאת אומרת שהקו האלכסוני ותווך של rad
פיקסלים ימינה ממנו ו- rad פיקסלים למתה ממנו יכיל ערכים של עוצמת הגדיאנט וכל שאר הפיקסלים בתמונה יהיו ללא שינוי.
הישתמשו במסכות סובל לחישוב הגרדיאנט.
יש לעגל את כל הערכים של עוצמת הגרדיאנט. כל הערכים המתקבלים אשר גדולים מ 255 יש לשנות ל 255.

תזכורת :

$$\boldsymbol{SobelYmask} = \begin{bmatrix} \mathbf{1} & \mathbf{2} & \mathbf{1} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{-1} & \mathbf{-2} & \mathbf{-1} \end{bmatrix}, \ \boldsymbol{SobelXmask} = (\boldsymbol{SobelYmask})^{\boldsymbol{T}}$$

דוגמה :

Img = **myDiagEdgeDetect**(image, rad)

מכון טכנולוגי חולון
Holon Institute of Technology

חלק ג' (בונוס): 15%

שאלה 9 (15 נק.)

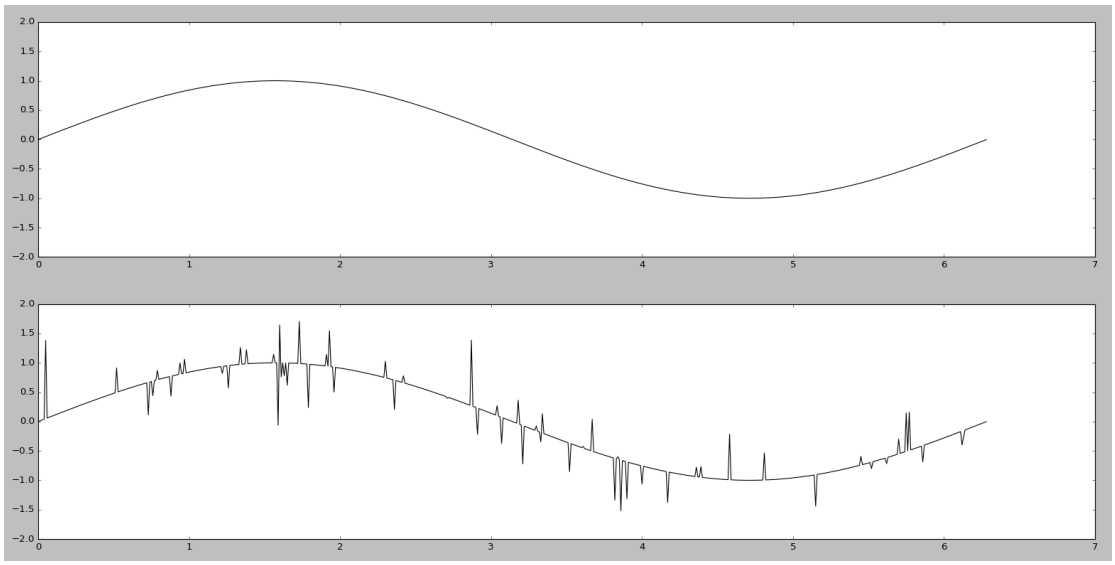
9. הגדירו פונקציה בשם : **SaltAndPepperSignalNoise**

הפונקציה צריכה לקבל אות חד מימדי (ndarray) ואחוז הדגימות הרועשות (מספר בין 0 ל- 1).

הפונקציה צריכה להחזיר את האות שהיא קיבלה עם תוספת רעש בדגימות אקראיות לפי התפלגות אחידה (מיקום הדגימות הרועשות מתפלג אחיד). כמות הדגימות היא בהתאם לאחוז הדגימות הרועשות מכמות הכוללת של הדגימות באות (במידה והערך לא שלם, יש לבצע עיגול). בכל אחד ממקומות בהם מתווסף הרעש, יש להוסיף רעש אקראי בהתפלגות נורמלית (גאוסיאנית) עם תוחלת 0 ושונות 0.5.

דוגמה :

NoisySignal = **SaltAndPepperSignalNoise**(Signal, NoisePr)



בהצלחה!