

תרגיל בית 10

Numpy and Image Processing

הנחיות כלליות:

- קראו **בעיון** את השאלות והקפידו שהתכניות שלכם פועלות בהתאם לנדרש.
- את התרגיל יש לפתור לבד!
- הקפידו על כללי ההגשה המפורסמים באתר. בפרט, יש להגיש את כל הפתרונות לשאלות יחד בקובץ `ex10_012345678.py` המצורף לתרגיל, לאחר החלפת הספרות 012345678 שבשם הקובץ במספר תעודת הזהות שלכם, כל 9 הספרות כולל ספרת ביקורת.
- אופן ביצוע התרגיל: שימו לב, בתרגיל זה עליכם להשלים את הקוד בקובץ המצורף.
- **אין לשנות את שמות המחלקות, הפונקציות, המתודות והמשתנים שכבר מופיעים בקובץ השלד של התרגיל.**
- **אין למחוק את ההערות שמופיעות בקובץ השלד.**
- היות ובדיקת התרגילים עשויה להיות אוטומטית, יש להקפיד על פלטים מדויקים על פי הדוגמאות (עד לרמת הרווח).
- בדיקה עצמית: כדי לוודא את נכונותן ואת עמידותן של התוכניות לקלטים שגויים, בכל שאלה הריצו את תוכניתכם עם מגוון קלטים שונים, אלה שהופיעו כדוגמאות בתרגיל וקלטים נוספים עליהם חשבתם (וודאו כי הפלט נכון).
- **ניתן להניח שהקלט תקין בהתאם להערות המפורטות בהוראות כל שאלה, אלא אם מצויין אחרת.**
- מועד אחרון להגשה: כמפורסם באתר.

שאלה 1

נתונים קבצי מידע המתארים את כמות התבואה מסוגים שונים שגידלו מספר חקלאים בשנה מסוימת. כל שורה בטבלה מייצגת סוג תבואה (Crop) וכל עמודה מייצגת חקלאי (Farmer) כלשהו. השורה הראשונה מייצגת את שמות החקלאים. לדוגמא, אלו הם נתוני התבואה עבור שנת 2018 השמורים בקובץ "Crops2018.csv":

	Joe	Nataly	Martha	Dan	David	Bob
Wheat	10	20	30	40	50	60
Corn	12	22	32	42	52	62
Carrot	14	24	34	44	54	64
Banana	16	26	36	46	56	66

א. ממשו את הפונקציה `load_crops(filename)` המקבלת את שם הקובץ בתור מחרוזת, קוראת אותו, ומחזירה 3 מערכי `numpy`. הראשון הוא מערך חד מימדי (וקטור) אשר מכיל את שמות החקלאים (`string`). השני גם הוא מערך חד מימדי, ומכיל את שמות (סוגי התבואה) (העמודה השמאלית בטבלה). השלישי הוא מערך דו מימדי (מטריצה) ומכיל את הנתונים המספריים בטבלה (כל מספר מייצג משקל תבואה מסוג מסוים שקנה חקלאי מסוים ביחידות של טון). איברי מערך זה יהיו משתנים מטיפוס `.int`.

ניתן להניח שמבנה הקובץ תקין, אך יש לתפוס שגיאות `IO`. במידה וקרתה, יש להדפיס למסך הודעה `"IO Error!"`.

שימו לב שהטבלה הנ"ל והקובץ המצורף הם רק דוגמא ועל הפונקציה לתמוך בכל גודל של מטריצה.

בסעיפים הבאים אין להשתמש בלולאות.

בדוגמאות ההרצה, כאשר מוחזר מערך, חשוב שהערכים בו יהיו זהים. פרמטר `dtype` אינו משנה.

בכולם, נתייחס אל `weights` בתור מטריצת המשקלים שנטענה מ `load_crops` שבסעיף א', אל `crops` בתור וקטור התבואות, ו-`farmers` בתור וקטור שמות החקלאים. כולם מטיפוס `numpy`.

ב. ממשו את הפונקציה `section1B(farmers, crops, weights)`. הפונקציה מחזירה את ממוצע כלל המשקלים של החקלאים במקומות הזוגיים. במקרה שלנו, יודפס ממוצע המשקלים עבור `Nataly,Dan,Bob` כך:

```
>>> section1B(farmers, crops, weights)
```

```
43.0
```

ג. ממשו את הפונקציה `section1C(farmers, crops, weights)`. הפונקציה מחזירה מערך `numpy` של סוגי התבואה בעלי ממוצע משקלים מקסימלי.

לדוגמה במקרה שלנו:

```
>>> section1C(farmers, crops, weights)
```

```
array(['Banana'], dtype='<S6')
```

ד. ממשו את הפונקציה section1D(farmers, crops, weights). הפונקציה מחזירה מערך numpy של שמות החקלאים, אשר גידלו יותר מ-50 טון עבור לפחות 3 סוגי תבואה שונים (בכל אחד מ-3 מהסוגים הנ"ל).

לדוגמה במקרה שלנו:

```
>>> section1D(farmers, crops, weights)
```

```
array(['David', 'Bob'], dtype='<S6')
```

ה. נגדיר חקלאי "מאוזן" בתור חקלאי שהפרש בין משקל הגידול המקסימלי שלו למינימלי שלו קטן מ-8 טון. ממשו את הפונקציה section1E(farmers, crops, weights) אשר מחזירה למסך את מספר החקלאים המאוזנים.

לדוגמה במקרה שלנו:

```
>>> section1E(farmers, crops, weights)
```

6

שאלה 2

מרכז הספורט באוניברסיטה כולל חדר כושר ובו מספר עמדות אימון. במטרה לאמוד את זמינות המתקנים בעמדות האימון השונות, ערך מנהל חדר הכושר רישום של מספר המשתמשים בכל עמדה על פני חלונות של שעתיים, החל משעת הפתיחה (6:00) ועד לשעת הסגירה (00:00) ביום מסוים.

נתון קובץ בשם station_usage.csv, אשר מכיל טבלה בה כל ערך מספרי מציין את מספר המשתמשים בעמדת אימון

א. נתון קובץ בשם station_usage.csv, אשר מכיל טבלה בה כל ערך מספרי מציין את מספר המשתמשים בעמדת אימון מסוימת בחדר כושר, בחלון זמן מסוים. עמודות הטבלה מייצגות את עמדות האימון (stations), ושורות הטבלה מייצגות את חלונות הזמן (time slots).

מצורפת דוגמה לקובץ המתואר:

Time Slot	Free weights	Pull-up bar	Bench press	Ab crunch	Cross trainer
T6-8	0	0	0	0	0
T8-10	4	3	5	8	14
T10-12	3	1	2	2	12
T12-14	1	2	4	1	8
T14-16	0	0	0	0	0

T16-18	1	2	4	5	16
T18-20	3	4	5	8	10
T20-22	4	5	2	6	20
T22-24	0	0	0	0	0

ממשו את הפונקציה `load_usage(filename)` אשר מקבלת מחרוזת של שם הקובץ ומחזירה 3 מערכי `numpy`. הראשון חד מימדי, ומכיל את שמות המתקנים השונים בחדר כושר. השני חד מימדי, ומכיל את חלונות הזמן (בתור מחרוזות). השלישי דו מימדי (מטריצה) המכילה את נתוני המשתתפים (מסוג `int`).

שימו לב: על הפונקציה לזהות שורות שבהם סך המשתמשים הם 0 ולהתעלם מהם. כלומר, שורה אשר סך המשתמשים בה הוא 0, לא תיכנס למטריצה, וה-`time_slot` שלה לא ייכלל בוקטור המתאים.

הערה: ניתן להניח שהקובץ תקין ואין צורך לטפל בשום שגיאה.

בסעיפים הבאים אין להשתמש בלולאות.

בדוגמאות ההרצה, כאשר מוחזר מערך, חשוב שהערכים בו יהיו זהים. פרמטר ה-`dtype` אינו משנה.

בכולם, נתייחס אל `usage` בתור מטריצת המשתמשים שנטענה מ `load_usage` שבסעיף א', אל `time_slots` בתור וקטור חלונות הזמן, ואל `station_names` בתור וקטור שמות התחנות. כולם מטיפוס `numpy` של מערך.

ב. ממשו את הפונקציה `section2B(station_names, time_slots, usage)`. הפונקציה מחזירה את מספר המשתמשים המקסימלי בחלון זמן כלשהו (כלומר, סך המשתמשים על כל המכשירים).

לדוגמה במקרה שלנו:

```
>>> section2B(station_names, time_slots, usage)
```

37

ג. ממשו את הפונקציה `section2C(station_names, time_slots, usage)`. הפונקציה מחזירה את מספר התאים בטבלה עם פחות מ-2 משתמשים.

לדוגמה במקרה שלנו:

```
>>> section2C(station_names, time_slots, usage)
```

4

ד. ממשו את הפונקציה `section2D(station_names, time_slots, usage)`. הפונקציה מחזירה את עמדת אימון בעלת מספר משתמשים יומי מקסימלי. במידה ויש כמה כאלו, תוחזר אחת מהן.

לדוגמה במקרה שלנו:

```
>>> section2D(station_names, time_slots, usage)
```

Cross trainer

ה. לכל time slot יש את סך המשתמשים שלו (כלומר, סך המשתמשים על כל המכשירים). נתבונן על הממוצע של ערך זה על פני כל ה time_slots. ממשו את הפונקציה section2E(station_names, time_slots, usage) הפונקציה מחזירה את ה time_slots שסך המשתמשים בו הוא הרחוק ביותר (בערך מוחלט) מהממוצע שצוין. ניתן להניח שיש אחד כזה.

לדוגמה במקרה שלנו:

```
>>> section2E(station_names, time_slots, usage)
```

T12-14

שאלה 3

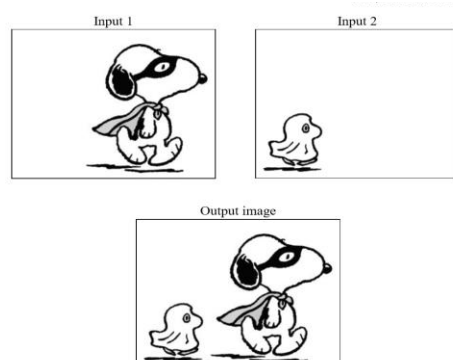
שימו לב: בשאלה זו אין להשתמש בלולאות.

ממשו את הפונקציה `mergelimages(im1,im2)` המקבלת שני מערכים מטיפוס `ndarray` המייצגים תמונה, ומחזירה מערך חדש המייצג תמונה הממזגת את 2 תמונות הקלט.

במערך הפלט, כל איבר יכיל את הערך הנמוך יותר (=כנה יותר) מבין שני ערכי הפיקסלים במיקום התואם בשני מערכי הקלט. למשל - אם במיקום (i,j) במערך `im1` נמצא הערך x ובמיקום זה במערך `im2` נמצא הערך y , ומתקיים $x \leq y$ אז במיקום (i,j) במערך הפלט יוצב x , אחרת יוצב y .

- מערכי הקלט והפלט מכילים מספרים שלמים בטווח 0-255 המייצגים פיקסלים בגווני אפור, כאשר 0 מייצג את הצבע השחור, ו-255 את הצבע הלבן.
- ניתן להניח ששתי תמונות הקלט הן באותו גודל בדיוק, וגודל מערך הפלט יהיה זהה לגודל זה.
- כדי לקבל את גודלו של מערך מסוג `ndarray` ניתן לגשת לשדה `shape` המחזיר `tuple` בפורמט הבא: `(number of rows, number of columns)`.

דוגמא לפעולת הפונקציה:



שאלה 4

דור הביא את התמונה בשחור-לבן של דמות הדיסני האהובה עליו לשיבת צוות של קורס פייתון.

לרוע מזלו, נשפך עליה הרבה מלח ופלפל, וכעת 90% ממנה מלאה רעש salt&pepper כפי שלמדנו בתרגול!

כלומר, ב-90% מהפיקסלים בתמונה, הערך נמחק ובמקומו נמצא אחד מ-2 הערכים: 0,255.

עזרו לדור וממשו את הפונקציה `denoise_disney_character(disney_file, save_file)` אשר מקבלת את כתובת הקובץ לתמונה `disney_file`, ושומרת את התמונה המשוחררת הכי טובה שתוכלו בכתובת `save_file`. כלומר, תמונה בה הדמות תהיה כמה שיותר בולטת, והרקע כמה שפחות בולט.

- מערכי הקלט והפלט מכילים מספרים שלמים בטווח 0-255 המייצגים פיקסלים בגווי אפור, כאשר 0 מייצג את הצבע השחור, ו-255 את הצבע הלבן.
- ב-90% מהפיקסלים בתמונה, הערך נמחק ובמקומו נמצא אחד מ-2 הערכים: 0,255.
- (מומלץ אך לא חובה): להשתמש בקטעי הקוד מהתרגול המבצעים `denoising`. בקודים הנ"ל, ראינו שניתן להתאים:
 - גודל שכונה (פיקסלים שכנים להיעזר בהם).
 - סוג פונקציה (הדרך שבה נעזרים בפיקסלים השכנים).
 - מספר איטרציות (כמות הפעמים שמתקנים פיקסל "מלוכלך" ע"י פיקסלים שכנים "נקיים").
- **בניגוד לתרגילים ושאלות קודמות, כאן איננו מעוניינים במקרה הכללי עבור תמונה כלשהי, אלא בפתרון עבור התמונה הספציפית שסופקה לכם בתרגיל הנ"ל.**

ניקוד: כל תמונה בה ניתן לראות בבירור את הדמות תיחשב כמוצלחת, אם כי עבור תמונות מוצלחות במיוחד יינתן בonus!

בהצלחה!