תרגיל בית 3 – MDP ומבוא ללמידה

**עברו על כלל ההנחיות לפני תחילת התרגיל.**

# **הנחיות כלליות:**

* תאריך ההגשה: לחלק א' של התרגיל (MDP) – עד ליום האחרון של הסמסטר - 08/04/2024ב23:59  
   לחלק ב' של התרגיל (מבוא ללמידה) – עד לסוף מועדי א' - 17/05/2024ב23:59
* את המטלה יש להגיש **בזוגות בלבד.**
* יש להגיש מטלות מוקלדות בלבד. פתרונות בכתב יד לא ייבדקו.
* ניתן לשלוח שאלות בנוגע לתרגיל בפיאצה בלבד.
* המתרגל האחראי על תרגיל זה:  **דניאל אלגריסי.**
* בקשות דחיה מוצדקות (מילואים, אשפוז וכו') יש לשלוח למתרגל האחראי (**ספיר טובול**) בלבד.
* במהלך התרגיל ייתכן שנעלה עדכונים, למסמך הנ"ל – תפורסם הודעה בהתאם.
* העדכונים הינם מחייבים, ועליכם להתעדכן עד מועד הגשת התרגיל.
* שימו לב, העתקות תטופלנה בחומרה.
* התשובות לסעיפים בהם מופיע הסימון [Writing Hand on Apple ](https://emojipedia.org/apple/ios-14.6/writing-hand/) צריכים להופיע בדוח.
* לחלק הרטוב מסופק שלד של הקוד.
* אנחנו קשובים לפניות שלכם במהלך התרגיל ומעדכנים את המסמך הזה בהתאם. גרסאות עדכניות של המסמך יועלו לאתר. הבהרות ועדכונים שנוספים אחרי הפרסום הראשוני יסומנו כאן בצהוב. ייתכן שתפורסמנה גרסאות רבות – אל תיבהלו מכך. השינויים בכל גרסה יכולים להיות קטנים.

**לצורך הנוחות:**

**הבהרות ועדכונים גרסה ראשונה סומנו ככה.**

**הבהרות ועדכונים גרסה שניה סומנו ככה.**

**שימו לב שאתם משתמשים רק בספריות הפייתון המאושרות בתרגיל (מצוינות בתחילת כל חלק רטוב)  
לא יתקבל קוד עם ספריות נוספות**

**מומלץ לחזור על שקפי ההרצאות והתרגולים הרלוונטיים לפני תחילת העבודה על התרגיל.**

**חלק א׳ – MDP (44 נק׳)**

**רקע**

בחלק זה נעסוק בתהליכי החלטה מרקובים, נתעניין בתהליך עם **אופק אינסופי** (מדיניות סטציונרית).

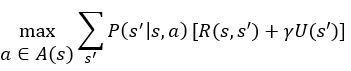
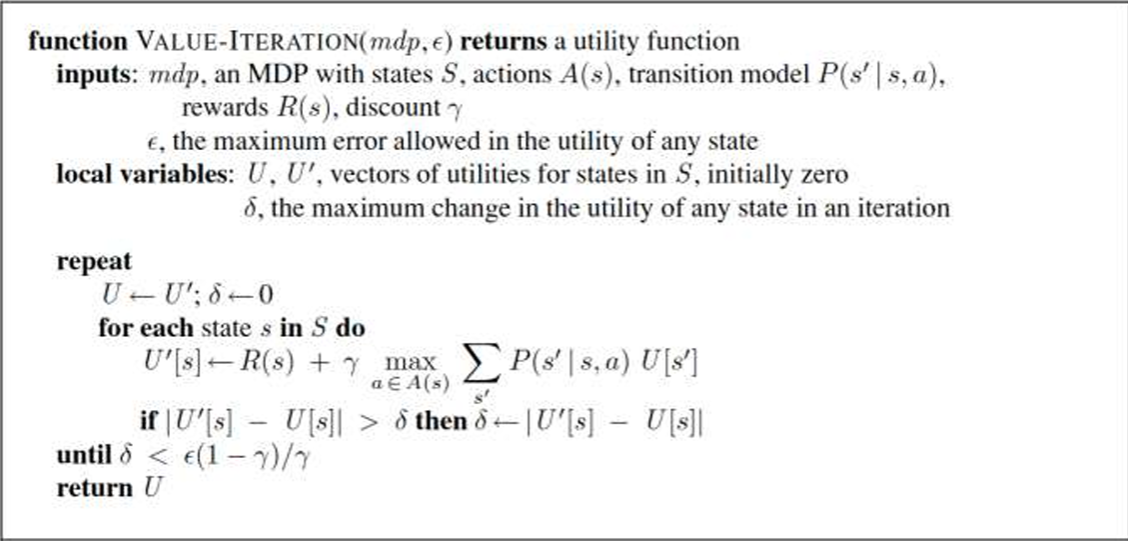
**חלק א׳ - חלק היבש** [Writing Hand on Apple ](https://emojipedia.org/apple/ios-14.6/writing-hand/)

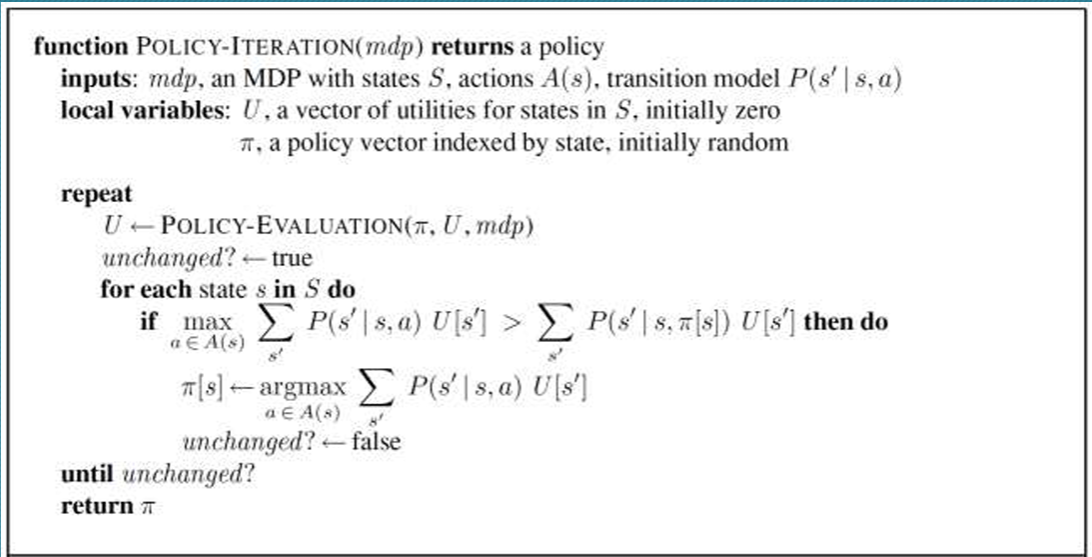
1. בתרגול ראינו את משוואת בלמן כאשר התגמול ניתן עבור המצב הנוכחי בלבד, כלומר , למתן תגמול זה נקרא "תגמול על הצמתים" מכיוון שהוא תלוי בצומת שהסוכן נמצא בו.   
   בהתאם להגדרה זו הצגנו בתרגול את האלגוריתמים Value iteration ו-Policy Iteration למציאת המדיניות האופטימלית.

כעת, נרחיב את ההגדרה הזו, לתגמול המקבל את המצב הנוכחי והמצב אליו הגיע הסוכן, כלומר: , למתן תגמול זה נקרא "תגמול תוצאתי". לצורך שלמות ההגדרה, נגדיר שאם לכל   
 מתקיים - אז .

1. (1 נק') התאימו את הנוסחה של התוחלת של התועלת מהתרגול, עבור התוחלת של התועלת המתקבלת במקרה של ״ תגמול תוצאתי ״, אין צורך לנמק.
2. (1 נק') כתבו מחדש את נוסחת משוואת בלמן עבור המקרה של ״ תגמול תוצאתי ״, אין צורך לנמק.

בסעיפים הבאים התייחסו גם למקרה בו , והסבירו מה לדעתכם התנאים שצריכים להתקיים על הסביבה\ על מנת שתמיד נצליח למצוא את המדיניות האופטימלית.

1. (2 נק') נסחו את אלגוריתם Value Iteration עבור המקרה של ״ תגמול תוצאתי ״.
2. (2 נק') נסחו את אלגוריתם Policy Iteration עבור המקרה של ״ תגמול תוצאתי ״.



In policy iteration we use the expected utility function from **a** to calculate POLICY-EVALUATION, if then the Utility won’t necessarily converge in both Policy and Value iteration, meaning a finite horizon is required in order to find an optimal solution for this MDP.

נתון הגרף הבא:



נתונים:

* (Discount factor) .
* אופק אינסופי.
* – קבוצת המצבים – מתארים את מיקום הסוכן בגרף.
* – קבוצת המצבים הסופיים.
* קבוצת הפעולות לכל מצב (על פי הגרף), לדוגמא: .
* תגמולים ("תגמול תוצאתי"):
* מודל המעבר הוא דטרמיניסטי, כלומר כל פעולה מצליחה בהסתברות אחת.

1. (יבש 2 נק') הרץ את האלגוריתם Value iteration שכתבת על הגרף הנתון. ומלא את הערכים בטבלה הבאה, כאשר . (ייתכן שלא צריך למלא את כולה).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  | 5 | 5 | 5 | 5 | 0 |  |
|  |  |  |  | 7 | 7 | 7 | 7 | 0 |  |
|  |  |  |  | 1.5 | 1.5 | 1.5 | -1 | 0 |  |
|  |  |  |  | 2.5 | 2.5 | 2.5 | -1 | 0 |  |
|  |  |  |  | 2.5 | 2.5 | 2.5 | -1 | 0 |  |
|  |  |  |  | 0.25 | 0.25 | -1.5 | -1 | 0 |  |
|  |  |  |  | 0.25 | 0.25 | -1.5 | -1 | 0 |  |

1. (יבש 2 נק') הרץ את האלגוריתם Policy iteration שכתבת על הגרף הנתון. ומלא את הערכים בטבלה הבאה, כאשר המדיניות ההתחלתית מופיעה בעמודה הראשונה בטבלה. (ייתכן שלא צריך למלא את כולה).

הניחו שבמידה ולא קיים שיפור, האלגוריתם יבחר תמיד להשאיר את הפעולה הקודמת.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | Top-Left(S4) |  |  |  |  |

1. (יבש 2 נק') חיזרי על הסעיף הקודם. הפעם עם **אופק סופי כאשר** (שימי לב, המדיניות לא חייבת להסתיים במצב מסיים, ישנם מצבים שלא יכולים להגיע למצב מסיים עם אופק זה. ישנם צמתים עם מספר תשובות נכונות, נקבל את כולם).
2. (1 נק') ללא תלות בשינוי של הסעיף הקודם. אם , מה מספר המדיניות האופטימליות הקיימות? נמקו.
3. (1 נק') ללא תלות בשנוי של הסעיף הקודם, הסבירי מה היה קורה אם

בתשובתך, התייחסי גם לערכי התועלות של כל צומת וגם לשינוי במדיניות, אין צורך לחשב.

Here the discount factor is 1, in addition to a positive reward on the path from s­1 to s2 and vice versa, we get the same results in the first iteration but after that s1 will want to go to s2 given that his utility plus the reward (according to our equation) with probability and discount factor 1 will give him more than what he already has (5+2=7), which will lead to the utility of ­s2 changing in the next iteration accordingly and for the same reasons (2+7=9), and so on…

The utility wont converge as explained above, but the policy will converge to the same as before with the adjustment of s1 and s2 pointing at each other, never including the goal state.

**חלק ב׳ - היכרות עם הקוד**

חלק זה הוא רק עבור היכרות הקוד, עבורו עליו במלואו ווודאו כי הינכם מבינים את הקוד.

mdp.py – **אתם לא צריכים לערוך כלל את הקובץ הזה**.   
בקובץ זה ממומשת הסביבה של ה-mdp בתוך מחלקת MDP. הבנאי מקבל:

* board - המגדיר את המצבים האפשריים במרחב ואת התגמול לכל מצב, תגמול על הצמתים בלבד.
* terminal\_states – קבוצה של המצבים הסופיים (בהכרח יש לפחות מצב אחד סופי).
* transition\_function – מודל המעבר בהינתן פעולה, מה ההסתברות לכל אחת מארבע הפעולות האחרות. ההסתברויות מסודרות לפי סדר הפעולות.
* gamma – discount factor המקבל ערכים - .

בתרגיל זה לא נבדוק את המקרה בו .

הערה: קבוצת הפעולות מוגדרת בבנאי והיא קבועה לכל לוח שיבחר.

למחלקת MDP יש מספר פונקציות שעשויות לשמש אתכם בתרגיל.

* print\_rewards() – מדפיסה את הלוח עם ערך התגמול בכל מצב.
* print\_utility(U) – מדפיסה את הלוח עם ערך התועלת U לכל מצב.
* print\_policy(policy) – מדפיסה את הלוח עם הפעולה שהמדיניות policy נתנה לכל מצב שהוא לא מצב סופי.
* step(state, action) – בהינתן מצב נוכחי state ופעולה action מחזיר את המצב הבא באופן דטרמיניסטי. עבור הליכה לכיוון קיר או יציאה מהלוח הפונקציה תחזיר את המצב הנוכחי state.

**חלק ג׳ – רטוב**  
כל הקוד צריך להיכתב בקובץ mdp\_implementation.py

מותר להשתמש בספריות:

All the built-in packages in python, numpy, matplotlib, argparse, os, copy, typing, termcolor, random

עליכם לממש את הפונקציות הבאות:

* (רטוב 6 נק'): value\_iteration(mdp, U\_init, epsilon) – בהינתן ה-mdp, ערך התועלת ההתחלתי U\_init, וחסם העליון לשגיאה מהתוחלת של התועלת האופטמילי epsilon מריץ את האלגוריתם value iteration ומחזיר את U המתקבל בסוף ריצת האלגוריתם. TODO
* (רטוב 4 נק'): get\_policy(mdp, U) – בהינתן ה-mdp וערך התועלת U (המקיים את משוואת בלמן) מחזיר את המדיניות (במידה וקיימת יותר מאחת, מחזיר אחת מהן). TODO
* (רטוב 4 נק'): policy\_evaluation(mdp, policy) – בהינתן ה-mdp, ומדיניות policy מחזיר את ערכי התועלת לכל מצב. TODO
* (רטוב 6 נק'): policy\_iteration(mdp, policy\_init) - בהינתן ה-mdp, ומדיניות התחלתית policy\_init, מריץ את האלגוריתם policy iteration ומחזיר מדיניות אופטימלית. TODO

למימוש הפונקציות הבאות ניתן להשתמש באיזה ספריות שתרצו.

בסעיפים הבאים ניתן להוסיף ארגומנטים לפונקציות אך יש לתעד אותם היטב

* Shape

  Description automatically generated(רטוב 5 נק'): get\_all\_policies(mdp, U, …) – בהינתן ה-mdp, וערך התועלת U (המקיים את משוואת בלמן) מדפיס\מציג את כל המדיניות המקיימות ערך זה בלוח בודד (יש לבצע ויזואליזציה להצגת כל המדיניות), לדוגמא:

הפונקציה מחזירה את מספר המדיניות (policies) השונות הקיימות המקיימות את U. TODO

↓↑→←

אלו הסימונים שניתן להעביר לפונקציית ההדפסה (למעשה ניתן להעביר מה שתרצו בכל תא במטריצה) בכל תא, שימי את כל החיצים המתאימים למדיניות כלשהי.

למשל בדוגמה, עבור התא (1,1) נשים את כל החיצים

* (רטוב 5 נק'): get\_policy\_for\_different\_rewards(mdp, …) – בהינתן ה-mdp מדפיס\מציג את המדיניות האופטימלית כתלות ב-~~r~~ R (ערכי התגמול לכל מצב שאינו סופי). TODO   
  יש להחזיר רשימה של ערכי R שבהם יש שינוי במדיניות מהקטן לגדול  
  שימו לב – הפונקציה לא מחזירה ערך, רק מדפיסה.

שימו לב 2 – ניתן להניח שלא יהיו שינויים במדיניות עבור ערכי קטנים מ- וגדולים מ- . בנוסף, דיוק של 2 ספרות אחרי הנקודה הינו מספק.

דוגמא חלקית של פתרון אפשרי:

A picture containing text, crossword puzzle

Description automatically generated

[Writing Hand on Apple ](https://emojipedia.org/apple/ios-14.6/writing-hand/) בנוסף לקוד עליכם לצרף להגשה היבשה את התצוגות של הפונקציות על הסביבה שניתנה בתרגיל. רק על שני הפונקציות האחרונות: get\_all\_policies, get\_policy\_for\_different\_rewards

**עבור מצבים סופיים וקירות (WALL), הערך שצריך לחזור בתאים אלו עבור טבלאות המדיניות הוא None. כל ערך אחר לא יתקבל כתשובה.**

main.py – דוגמת הרצה לשימוש בכל הפונקציות.

בתחילת הקובץ אנו טוענים את הסביבה משלושה קבצים:

board, terminal\_states, transition\_function

ויוצרים מופע של הסביבה (mdp).

* שימו לב, שכרגע הקוד ב-main לא יכול לרוץ מכיוון שאתם צריכים להשלים את הפונקציות הרלוונטיות ב- mdp\_implementation.py.
* בנוסף, על מנת לראות את הלוח עם הצבעים עליכם להריץ את הקוד בIDE לדוגמה PyCharm.

**חלק ב׳ - מבוא ללמידה (56 נק׳)**

[Writing Hand on Apple ](https://emojipedia.org/apple/ios-14.6/writing-hand/) **חלק א׳ – חלק היבש (28 נק')**

**k****NN – נעים להכיר**

בחלק זה תכירו אלגוריתם למידה בשם kNN, או בשמו המלא k-Nearest Neighbors, כאשר ה־k הוא למעשה פרמטר!  
יהי סט אימון עם ־דוגמות, , כאשר .  
כלומר הדוגמות הינן וקטורים ־ממדיים והתגיות הינן מדומיין כלשהו, הבעיה היא בעיית קלסיפיקציה (סיווג).  
אם לא נאמר אחרת, הקלסיפיקציה תהיה בינארית, כלומר .  
עבור כל דוגמה בסט האימון, ניתן להסתכל על הכניסה ה־ בווקטור כעל הfeature ה־ של הדוגמה, קרי כל דוגמה מיוצגת על ידי ־ערכים: .  
תהליך ה"אימון" של האלגוריתם הוא טריוויאלי – פשוט שומרים את סט האימון במלואו.  
תהליך הסיווג הוא גם פשוט למדי – כאשר רוצים לסווג דוגמה מסט המבחן מסתכלים על השכנים הקרובים ביותר שלה במישור ה־ממדי מבין הדוגמות בסט האימון, ומסווגים את הדוגמה על פי הסיווג הנפוץ ביותר בקרב השכנים.  
על מנת להימנע משוויון בין הסיווגים, נניח בדרך כלל כי ־אי זוגי, או שנגדיר היטב שובר שוויון.  
אם לא נאמר אחרת, במקרה של שוויון בקלסיפיקציה בינארית, נסווג את הדוגמה כחיובית .

שאלות הבנה

1. (3 נק') כאמור, בתהליך הסיווג אנו בוחרים עבור הדוגמה את הסיווג הנפוץ ביותר של השכנים הקרובים ביותר, אולם עלינו להגדיר את פונקציית המרחק עבור קביעת סט שכנים זה.  
   שתי פונקציות מרחק נפוצות הינן מרחק אוקלידי ומרחק מנהטן.
2. עבור איזה ערכים של נקבל שאין תלות בבחירת פונקציית המרחק? (נמקי)
3. עבור בעיית קלסיפיקציה בינארית תנו דוגמה פשוטה לערכי , סט אימון ודוגמת מבחן בה השימוש בכל אחת מפונקציות המרחק הנ"ל משנה את סיווג דוגמה המבחן.

מעתה, אלא אם כן צוין אחרת, נשתמש במרחק אוקלידי.

נתונה קבוצת האימון הבאה, כאשר :



1. (1 נק') איזה ערך של עלינו לבחור על מנת לקבל את הדיוק המרבי **על קבוצת האימון**? מה יהיה ערך זה? (הדוגמא לא יכולה להיות שכנה של עצמה)
2. (1 נק') עבור איזה ערך של נקבל מסווג של קבוצת האימון? קרי כל דוגמת מבחן תקבל את הסיווג הנפוץ של כלל קבוצת האימון?
3. (2 נק') נמקו מדוע שימוש בערכי גדולים או קטנים מדי יכול להיות גרוע עבור קבוצת הדגימות הנ״ל.
4. (2 נק') שרטט את גבול ההחלטה של 1-nearest neighbor עבור הגרף.

**השוואה בין מודלי למידה:**

1. (3 נק') הציגו מסווג מטרה וקבוצת אימון בעלת לכל היותר 10 דוגמאות כך שלמידת עץ ID3 תניב מסווג אשר עונה נכון עבור כל דוגמת מבחן אפשרית (כלומר יתקבל מסווג המטרה), אך למידת KNN תניב מסווג שעבורו קיימת לפחות דוגמת מבחן אחת עליה הוא יטעה, לכל ערך K שייבחר.
2. (3 נק') הציגו מסווג מטרה וקבוצת אימון בעלת לכל היותר 10 דוגמאות כך שלמידת מסווג KNN עבור ערך K מסוים תניב מסווג אשר עונה נכון עבור כל דוגמת מבחן אפשרית (כלומר יתקבל מסווג המטרה), אך למידת עץ ID3 תניב מסווג אשר עבורו קיימת לפחות דוגמת מבחן אפשרית אחת עליה הוא יטעה.
3. (3 נק') הציגו מסווג מטרה וקבוצת אימון בעלת לכל היותר 10 דוגמאות כך שלמידת מסווג KNN עבור ערך K מסוים תניב מסווג אשר עבורו קיימת לפחות דוגמת מבחן אפשרית אחת עליה הוא יטעה, וגם למידת עץ ID3 תניב מסווג אשר עבורו קיימת לפחות דוגמת מבחן אחת אפשרית עליה הוא יטעה.
4. (3 נק') הציגו מסווג מטרה וקבוצת אימון בעלת לכל היותר 10 דוגמאות כך שלמידת מסווג KNN עבור ערך K מסוים תניב מסווג אשר עונה נכון עבור כל דוגמת מבחן אפשרית (כלומר יתקבל מסווג המטרה), וגם למידת עץ ID3 תניב מסווג עונה נכון עבור כל דוגמת מבחן אפשרית (כלומר יתקבל מסווג המטרה).

**מתפצלים ונהנים**

(7 נק') כידוע, בעת סיווג של דוגמת מבחן על ידי עץ החלטה, בכל צומת בעץ אנו מחליטים לאיזה צומת בן להעביר את דוגמת המבחן על ידי ערך סף שמושווה לfeature של הדוגמה. לפעמים ערך הסף קרוב מאוד לערך הfeature של דוגמת המבחן. היינו רוצים להתחשב בערכים "קרובים" לערך הסף בעת סיווג דוגמת מבחן, ולא לחרוץ את גורלה של הדוגמה לתת־עץ אחד בלבד; לצורך כך נציג את האלגוריתם הבא:  
  
יהיו עץ החלטה , דוגמת מבחן , ווקטור המקיים .  
כלל אפסילון־החלטה שונה מכלל ההחלטה הרגיל שנלמד בכיתה באופן הבא:   
נניח שמגיעים לצומת בעץ המפצל לפי ערכי התכונה , עם ערך הסף .  
אם מתקיים  אזי ממשיכים **בשני** המסלולים היוצאים מצומת זה, ואחרת ממשיכי לבן המתאים בדומה לכלל ההחלטה הרגיל. לבסוף, מסווגים את הדוגמה *בהתאם לסיווג הנפוץ ביותר של הדוגמאות הנמצאות בכל העלים אליהם הגענו במהלך הסיור על העץ (במקרה של שוויון – הסיווג ייקבע להיות ).*

יהא עץ החלטה לא גזום, ויהא העץ המתקבל מ־ באמצעות גיזום מאוחר שבו הוסרה הרמה התחתונה של (כלומר כל הדוגמות השייכות לזוג עלים אחים הועברו לצומת האב שלהם).  
הוכיחו\הפריכו: **בהכרח** קיים ווקטור כך שהעץ עם כלל אפסילון־החלטה והעץ עם כלל ההחלטה הרגיל יסווגו כל דוגמת מבחן ב בצורה זהה.

**חלק ב׳ - היכרות עם הקוד  
רקע**

חלק זה הוא רק עבור היכרות הקוד, עבורו עליו במלואו ווודאו כי הינכם מבינים את הקוד.  
בחלק של הלמידה, נעזר ב 𝑑𝑎𝑡𝑎𝑠𝑒𝑡, הדאטה חולק עבורכם לשתי קבוצות: קבוצת אימון train.csv וקבוצת מבחן test.csv.   
ככלל, קבוצת האימון תשמש אותנו לבניית המסווגים, וקבוצת המבחן תשמש להערכת ביצועיהם.

בקובץ utils.py תוכלו למצוא את הפונקציות הבאות לשימושכם:  
 load\_data\_set, create\_train\_validation\_split, get\_dataset\_split   
אשר טוענות/מחלקת את הדאטה בקבצי ה־csv למערכי np.array (קראו את תיעוד הפונקציות).

הדאטה של ID3 עבור התרגיל מכיל מדדים שנאספו מצילומים שנועדו להבחין בין גידול שפיר לגידול ממאיר. כל דוגמה מכילה 30 מדדים כאלה, ותווית בינארית **diagnosis** הקובעת את סוג הגידול (0=שפיר, 1=ממאיר). כל התכונות (מדדים( רציפות. העמודה הראשונה מציינת האם האדם חולה (M) או בריא (B). שאר העמודות מציינות כל תכונות רפואיות שונות של אותו אדם (התכונות מורכבות ואינכם צריכים להתייחס למשמעות שלהן כלל).

תיקיית :

* תיקיה זו אלו מכילה את קבצי הנתונים עבור .

קובץ :

* קובץ זה מכיל פונקציות עזר שימושיות לאורך התרגיל, כמו טעינה של וחישוב הדיוק.
* בחלק הבא יהיה עליכם לממש את הפונקציה . קראו את תיעוד הפונקציות ואת ההערות הנמצאות תחת התיאור TODO.

קובץ :

* קובץ בדיקה בסיסי שיכול לעזור לכם לבדוק את המימוש.

קובץ :

* קובץ זה מכיל 3 מחלקות שימושית לבניית עץ שלנו.
  + המחלקה : מחלקה זו מממשת הסתעפות של צומת בעץ. היא שומרת את התכונה ואת הערך שלפיהם מפצלים את הדאטה שלנו.
  + המחלקה : מחלקה זו מממשת צומת בעץ ההחלטה.   
    הצומת מכיל שאלה ואת שני הבנים כאשר הוא הענף בחלק של הדאטה שעונה על שאלת הצומת   
    (הפונקציה של ה מחזירה ).  
    ו־ הוא הענף בחלק של הדאטה שעונה על שאלת הצומת   
    (הפונקציה של ה מחזירה ).
  + המחלקה : מחלקה זו מממשת צומת שהוא עלה בעץ ההחלטה. העלה מכיל לכל אחד מהמחלקות בדאטה את מספר הדוגמאות בעלה עבור כל מחלקה (למשל: ).

קובץ :

* קובץ זה מכיל את המחלקה של שתצטרכו לממש חלקים ממנה, עיינו בהערות ותיעוד המתודות.

קובץ :

* קובץ הרצת הניסויים של ID3, הקובץ מכיל את הניסויים הבאים, שיוסברו בהמשך:

**חלק ג׳ – חלק רטוב ID3 (28 נק')**

עבור חלק זה מותר לכם להשתמש בספריות הבאות:

All the built in packages in python, sklearn, pandas ,numpy, random, matplotlib, argparse, abc, typing.

**אך כמובן שאין להשתמש באלגוריתמי הלמידה, או בכל אלגוריתם או מבנה נתונים אחר המהווה חלק מאלגוריתם למידה אותו תתבקשו לממש.**

1. (3 נק') השלימו את הקובץ utils.py ע"י מימוש הפונקציה .   
   קראו את תיעוד הפונקציה ואת ההערות הנמצאות תחת התיאור TODO.  
   (הריצו את הטסטים המתאימים בקובץ *לוודא שהמימוש שלכם נכון).*  
   שימו לב! בתיעוד ישנן הגבלות על הקוד עצמו, אי־עמידה בהגבלות אלו תגרור הורדת נקודות.  
   בנוסף, שנו את ערך ה בתחילת הקובץ מ־ למספר תעודת הזהות של אחד מהמגישים.
2. (10 נק') **אלגוריתם ID3:**
   1. השלימו את הקובץ ID3.py ובכך ממשו את אלגוריתם כפי שנלמד בהרצאה. TODO  
      שימו לב שכל התכונות רציפות. אתם מתבקשים להשתמש בשיטה של חלוקה דינמית המתוארת בהרצאה. כאשר בוחנים ערך סף לפיצול של תכונה רציפה, דוגמאות עם ערך השווה לערך הסף משתייכות לקבוצה עם הערכים הגדולים מערך הסף. במקרה שיש כמה תכונות אופטימליות בצומת מסוים בחרו את התכונה בעלת האינדקס המקסימלי.   
      כלל המימוש הנ"ל צריך להופיע בקובץ בשם , באזורים המוקצים לכך.  
      (השלימו את הקוד החסר אחרי שעיינתם והפנמתם את הקובץ ואת המחלקות שהוא מכיל).
   2. ממשו את שנמצאת ב TODO

[Writing Hand on Apple ](https://emojipedia.org/apple/ios-14.6/writing-hand/)*והריצו את החלק המתאים ב* ציינו בדו"ח את הדיוק שקיבלתם.

1. **גיזום מוקדם.**

פיצול צומת מתקיים כל עוד יש בו יותר דוגמאות מחסם המינימום ,𝑚כלומר בתהליך בניית העץ מבוצע "גיזום מוקדם" כפי שלמדתם בהרצאות. שימו לב כי פירוש הדבר הינו שהעצים הנלמדים אינם בהכרח עקביים עם הדוגמאות. לאחר סיום הלמידה (של עץ יחיד), הסיווג של אובייקט חדש באמצעות העץ שנלמד מתבצע לפי רוב הדוגמאות בעלה המתאים.

* 1. (2 נק') [Writing Hand on Apple ](https://emojipedia.org/apple/ios-14.6/writing-hand/)הסבירו מה החשיבות של הגיזום באופן כללי ואיזה תופעה הוא מנסה למנוע?
  2. (3 נק') **עדכנו** את המימוש בקובץ כך שיבצע גיזום מוקדם כפי שהוגדר בהרצאה.   
     הפרמטר מציין את המספר המינימלי בעלה לקבלת החלטה, קרי יבוצע גיזום מוקדם אם ורק אם מספר הדוגמות בצומת קטן שווה לפרמטר הנ"ל. TODO
  3. (8 נק') **שימו לב, זהו סעיף יבש ואין צורך להגיש את הקוד שכתבתם עבורו.**בצעו כיוונון לפרמטר על קבוצת האימון:

1. בחרו לפחות חמישה ערכים שונים לפרמטר .

2. עבור כל ערך, חשבו את הדיוק של האלגוריתם על ידי  על קבוצת האימון בלבד.   
כדי לבצע את חלוקת קבוצת האימון ל- קבוצות יש להשתמש בפונקציה [sklearn.model\_selection.KFold](https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.KFold.html" \l "sklearn.model_selection.KFold" \o "sklearn.model_selection)עם הפרמטרים ,   
ו־ אשר שווה למספר תעודת הזהות של אחד מהשותפים.

* + 1. [Writing Hand on Apple ](https://emojipedia.org/apple/ios-14.6/writing-hand/)השתמשו בתוצאות שקיבלתם כדי ליצור גרף המציג את השפעת הפרמטר על הדיוק.   
       צרפו את הגרף בדו״ח. (לשימושכם הפונקציה בתוך הקובץ ).
    2. [Writing Hand on Apple ](https://emojipedia.org/apple/ios-14.6/writing-hand/) הסבירו את הגרף שקיבלתם. לאיזה גיזום קיבלתם התוצאה הטובה ביותר ומהי תוצאה זו?
  1. [Writing Hand on Apple ](https://emojipedia.org/apple/ios-14.6/writing-hand/)(2 נק׳) השתמשו באלגוריתם ID3 עם הגיזום המוקדם כדי ללמוד מסווג מתוך **כל** קבוצת האימון ולבצע חיזוי על קבוצת המבחן.   
     השתמשו בערך ה־ האופטימלי שמצאתם בסעיף c. (ממשו שנמצאת ב  *והריצו את החלק המתאים ב ).* ציינו בדו"ח את הדיוק שקיבלתם. האם הגיזום שיפר את הביצועים ביחס להרצה ללא גיזום?

**הוראות הגשה**

* הגשת התרגיל תתבצע אלקטרונית בזוגות בלבד.
* הקוד שלכם ייבדק (גם) באופן אוטומטי ולכן יש להקפיד על הפורמט המבוקש. הגשה שלא עומדת בפורמט לא תיבדק (ציון 0).
* המצאת נתונים לצורך בניית הגרפים אסורה ומהווה עבירת משמעת.
* הקפידו על קוד קריא ומתועד. התשובות בדוח צריכות להופיע לפי הסדר.
* יש להגיש קובץ zip יחיד בשם AI3\_<id1>\_<id2>.zip (ללא סוגריים משולשים) שמכיל:
* קובץ בשם AI\_HW3.PDF המכיל את תשובותיכם לשאלות היבשות.
* קבצי הקוד שנדרשתם לממש בתרגיל **ואף קובץ אחר:**
* קובץ – utils.py
* בחלק של עצי החלטה – ID3.py, ID3\_experiments.py
* בחלק של mdp - mdp\_implementation.py

**אין להכיל תיקיות בקובץ ההגשה, הגשה שלא עומדת בפורמט לא תיבדק.**

נספח MDP:

דוגמת הרצה (שימו לב שהרצה זו השתמשה במודל הסתברותי שונה משלכם)

יצירת הסביבה:

mdp = MDP(board=board\_env,  
 terminal\_states=terminal\_states\_env,  
 transition\_function=transition\_function\_env,  
 gamma=1.0)

הדפסת הלוח עם התגמולים לכל מצב:

print('@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@')  
print("@@@@@@ The board and rewards @@@@@@")  
print('@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@')  
mdp.print\_rewards()

פלט:

תמונה שמכילה טקסט, צילום מסך, גופן

התיאור נוצר באופן אוטומטי

Value iteration:

print('@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@')  
print("@@@@@@@@@ Value iteration @@@@@@@@@")  
print('@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@')  
  
U = [[0, 0, 0, 0],  
 [0, 0, 0, 0],  
 [0, 0, 0, 0]]  
print("\nInitial utility:")  
mdp.print\_utility(U)  
print("\nFinal utility:")  
U\_new = value\_iteration(mdp, U)  
mdp.print\_utility(U\_new)  
print("\nFinal policy:")  
policy = get\_policy(mdp, U\_new)  
mdp.print\_policy(policy)

פלט:



Policy iteration:

print('@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@')  
print("@@@@@@@@@ Policy iteration @@@@@@@@")  
print('@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@')  
  
print("\nPolicy evaluation:")  
U\_eval = policy\_evaluation(mdp, policy)  
mdp.print\_utility(U\_eval)  
  
policy = [['UP', 'UP', 'UP', 0],  
 ['UP', 'WALL', 'UP', 0],  
 ['UP', 'UP', 'UP', 'UP']]  
print("\nInitial policy:")  
mdp.print\_policy(policy)  
print("\nFinal policy:")  
policy\_new = policy\_iteration(mdp, policy)  
mdp.print\_policy(policy\_new)  
  
print("Done!")

פלט:

