

אוניברסיטת בן-גוריון
המחלקה להנדסת תעשייה וניהול

למידת מכונה
Machine Learning

364-1-1811

מרצה: פרופ' בעז לרנר

boaz@bgu.ac.il, שעות קבלה: בתאום מראש

בודק תרגילים ואחראי מעבדות: רועי וולף wolfr@post.bgu.ac.il

שעות קבלה: בתאום מראש במייל

מבנה הקורס: הרצאה: כן תרגול: לא מעבדה: כן נקודות: 3

תיאור הקורס: הקורס מציג ומלמד עקרונות, שיטות ואלגוריתמים של למידת מכונה (machine learning) עם דגש על ייצוג ידע, חיזוי והיסק הסתברותי. ללא הנחת הנחות מיוחדות על עולם התוכן, למידת מכונה עושה שימוש בנתונים כדי להשיג שתי מטרות מרכזיות: (1) בניית מודלים המתארים קשרים בין משתנים, ועי"כ מייצגים ידע ומאפשרים בין היתר להבין השפעות בעולם התוכן, למשל, בין סיבה ("עישון") לתוצאה ("סרטן ריאות"); ו-2) בניית מודלים מנבאים המאפשרים חישוב ערך לא ידוע למשתנה מבוקש ("חומרת תאונה") בהינתן ערכים ידועים למשתנים אחרים (למשל, משתני דמוגרפיה והשכלה של הנהג יחד עם משתני דרך ורכב). בשנים האחרונות, עם הקפיצה הטכנולוגית בכל תחומי החיים, המבוססת בבעיות גדולות המלוות לעיתים בנתוני עתק (big data), תחום למידת המכונה הפך לנפוץ ביותר במדע, הנדסה וטכנולוגיה, ובקורס נתרכז בהיבטים מרכזיים שלו: למידה, סיווג, שיערוך, אישכול, קבלת החלטות והיסק סטטיסטי. נלמד את התיאוריה ועקרונות התחום ולצדם את האלגוריתמים והשיטות המקובלות, הן תיאורטית והן מעשית.

מטרות הקורס: הכרת עקרונות למידת מכונה, היסק הסתברותי, למידה אנושית ולמידה במערכות, כריית נתונים (data mining); הכרת שיטות ואלגוריתמים של מערכות לומדות לשם אישכול, סיווג, שיערוך, היסק וקבלת החלטות באופן הסתברותי; הכרת יישומי למידת מכונה בתעשייה; התנסות מודרכת בתוכנת R בעבודה על נתונים ושימוש במערכות לומדות ובניתוח ביצועיהן ע"י ביצוע תרגולים מעשיים.

דרישות הקורס:

1. פיתוח תוכנה מונחה עצמים (364-1-1421)
2. מודלים ברגרסיה לינארית (364-1-1061)

מבנה ציון הקורס:

מרכיב	משקל
1. מבחן	60%
2. תרגילים (בסביבת Python/R) (חלק א' 15% וחלק ב' 25%)	40%
3. הצגת קונספט (רשות/בונוס)	3%

סה"כ: 100% + 3% (בונוס/רשות)

חובת קבלת ציון עובר (56 ומעלה) בבחינה הסופית: כן
אם ציון הבחינה יהיה נמוך מ-56, ציון הקורס יהיה כציון הבחינה.

פרוט המפגשים

מפגש	נושאי המפגש	ספרות
1	הקדמה ללמידת מכונה ולקורס	1,2,5
2	מבנה מערכת לומדת לזיהוי תבניות (כולל בחירת ומיצוי מאפיינים)	1,2,6
3	למידה במערכות ביולוגיות ומלאכותיות ולימוד מונחה	1,2
4	אי-ודאות, היסק הסתברותי, משפט בייס ושימושיו	1-5
מעבדה 1	היכרות ותרגול ב-Python/R, טעינת נתונים, סטטיסטיקה תיאורית	
5	תורת ההחלטה הבייסיאנית	1-5
6	עצי החלטה	3
מעבדה 2	היכרות ותרגול ב-Python/R למערכות לומדות מונחות	
7	רשתות נוירונאליות	1
8	למידה לא מונחת ואישיכול	2,4
מעבדה 3	המשך תרגול עם מערכות לומדות מונחות ועם לא מונחות ב-Python/R. היכרות עם Kaggle.	
9	רשתות בייסיאניות	6
10	שיערוך צפיפות פונקציית הסתברות	1,2,4,5
11	יישומי למידת מכונה	
12	כריית נתונים + הדגמות למערכות של למידת מכונה	
13	השלמות, חזרה למבחן והצגת קונספט (סטודנטים)	

רשימת ספרות:

ספרי הקורס:

1. Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*, Clarendon Press. Oxford.
2. Duda, R. O., Hart, P. E. and Stork, D. G. (2001). *Pattern Classification* (2ed ed.), Wiley.
3. Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill.
4. Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*, Springer Press. Oxford.
5. Hastie, T., Tibshirani, R. & Friedman, J. (2009). *The Elements of Statistical Learning*, 2nd ed. Springer.
6. Koller, D. and Friedman, N. (2009). *Probabilistic graphical models*, The MIT Press.