

## 53243 כריית נתונים Data Mining

נקודות זכות: 3.0, שעות: הרצאה 3.  
דרישות קדם: אין

### תאור הקורס

הקורס עוסק בעקרונות ושימושים מתקדמים של כריית נתונים. במהלך הקורס נלמדים נושאים הקשורים לסיווג ושיבוץ נתונים בהדמיה סטטיסטית וממוחשבת. כמו כן, נבחנות השיטות ללמידה מבוקרת ולא מבוקרת ושיטות להפחתת המימד. מוצגים שימושים בתחומי כריית טקסט ומודלי שוק ההון. הקורס מלווה בתרגול במערכת MATLAB.

### נושאי הקורס

1. מבוא.
2. למידה מבוקרת ולא מבוקרת.
3. שיטות גרסיה: גרסיה מירבית, לינארית ולא-לינארית.
4. שיטת (SVM) Support Vector Machine – ללמידה מבוקרת.
5. שימושי SVM לבעיות שוק ההון וניהול.
6. שיטות הפחתת המימד (PCA & Projection Pursuit) ושימושיהן.
7. שיטות למידה לא מבוקרת, קלסטריציה הירארכית.
8. מודל - Gaussian Mixture Model (GMM)
9. האלגוריתמי שיבוץ: EM, PAM, K-MEANS.
10. שימושים לניתוח סדרות עתיות של שוק ההון.
11. כריית הטקסט וweb mining.
12. בניית association rules.

### הרכב הציון

- פרויקט גמר (40%)
- עבודות בית (30%)
- מצגת בקורס (30%)

### ספרות

- Bishop, C.M., *Pattern Recognition and Machine Learning (Information Science and Statistics)*, Springer, 2006.
- Witten, I.H. and Frank, E., *Data Mining: Practical Machine Learning Tools and Techniques*, 2<sup>nd</sup> ed., Morgan Kaufmann, 2005.
- Larose, D.T., *Data Mining Methods and Models*, Wiley-IEEE Press, 2006.
- Stork, D.G. and Yom-Tov, E., *Computer Manual in MATLAB to Accompany Pattern Classification*, 2<sup>nd</sup> ed., Wiley-Interscience, 2004.
- Tan, P.N. and Steinbach, M. and Kumar, V., *Introduction to Data Mining*, Addison Wesley, 2005.