

אמינות במערכות לומדות

מרצה: יניב רומנו

מספר קורס בפקולטה למדעי המחשב: 236601

מספר קורס בפקולטה להנדסת חשמל: 048100

קישור לאתר הקורס: ייתן בהמשך קישור למודל

הודעות:

מעוניינים/ות להירשם לקורס?

- אנא מלאו את הטופס בקישור הבא: <https://forms.gle/TG2G7T9gK2gT6bb3A>
- חשוב לציין שרישום בקישור הנ"ל אינו מהווה אישור הרשמה לקורס. בשל מגבלות מקום ייתכן שרק חלק מהסטודנטים שמילאו את פרטיהם יוכלו להשתתף בקורס.
- ניתן גם לשלוח מייל הקורס בדוא"ל yromano@technion.ac.il

מקצועות קדם:

1. 094422, 044202, 094412 הסתברות (או קורסים דומים)
2. 236756, 046195 מערכות לומדות

מקצועות קדם מומלצים:

1. 094423 מבוא סטטיסטיקה (או קורסים דומים, כגון 046202 עיבוד וניתוח מידע)
2. 236781, 046211 למידה עמוקה

סילבוס:

בקורס זה נלמד על שיטות להבטחת אמינות, עמידות, הוגנות, והישנות של מסקנות המתקבלות ממערכות למידה מורכבות, כגון למידה עמוקה. נחשף לבעיות של אלגוריתמי למידה מודרניים ונכיר שיטות מעשיות הנסמכות על תיאוריה עשירה להתמודד עם אתגרים אלו. הקורס ייתן במתכונת סמינר ויכסה מאמרים בחזית התחום של מדעי הנתונים. נתרכז בשיערוך אי ודאות בחיזוי, כמו כן זיהוי ומניעת הטייה מגדרית של מערכות לומדות. בנוסף, נסקור שיטות מתקדמות לבדיקות השערות מרובות ככלי לזיהוי תגליות מדעיות ממאגרי מידע ולאינטרפרטציה של אלגוריתמי למידה.

מקורות:

מאמרים מעיתונים ומכנסים מובילים בתחום של למידה חישובית וסטטיסטיקה.

דוגמאות למאמרים שיכוסו בקורס:

A tutorial on conformal prediction. Journal of Machine Learning Research, G. Shafer, and V. Vovk, 2008.

Predictive inference with the jackknife+. R.F. Barber, E. J. Candes, A. Ramdas, and R. J. Tibshirani, arXiv:1905.02928, 2019.

Equality of opportunity in supervised learning. M. Hardt, E. Price, and N. Srebro. NeurIPS, 2016.

Panning for gold: Model-X knockoffs for high dimensional controlled variable selection. E. J. Candes, Y. Fan, L. Janson, L. Lv. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 2018.

תוצאות למידה:

1. חשיפה לאתגרים בשימוש במערכות לומדות ביישומים שבהם האמינות היא דרישה הכרחית.
2. למידת מתודולוגיות להבטחת אמינות המסקנות המתקבלות ממערכות לומדות.

מפגשים:

- יתקיימו מדי שבוע בימי רביעי בשעות 10:30-12:30.
- מיקום: TBD.
- סה"כ יתקיימו 13 מפגשים, מתוכם:
 - 10-11 המפגשים הראשונים יינתנו על ידי מרצה הקורס. ההרצאות יינתנו בעברית וחומר הלימוד באנגלית.
 - המפגשים האחרונים יינתנו על ידי משתתפי הקורס (ראו הערה על עבודה בזוגות תחת **הרכב הציון**). על כל סטודנט/ית יהיה להציג מאמר וכיווני המשך אפשריים. המאמר שיתואר בכיתה וכיווני המשך שהוצעו יהיו את הבסיס לפרויקט.
- יתקיימו דיונים במהלך כל המפגשים ולכן הנוכחות בהרצאות היא חובה.

הרכב הציון (בקירוב):

- השתתפות פעילה וביצוע משימות בנוגע לחומר הלימוד: ~10%
- הצגה של מאמר מתוך רשימה (או מאמר אחר באישור המרצה): 20%-25%
- פרויקט המתבסס על המאמר שנבחר להצגה בכיתה: 65%-70%
 - מטרת הפרויקט היא לבדוק ביסודיות מאמר מדעי ולהציע הרחבה מסוימת.
 - דוגמאות להרחבות אפשריות: שיפור ביצועי האלגוריתם המוצע במאמר, הפעלה של האלגוריתם על מידע או domain אחר, וכד'.
 - מועד להגשת דוח המסכם את תוצאות הפרויקט יפורסם בהמשך. ההגשה של הפרויקט תהיה בסיום הסמסטר ולאחר תקופת המבחנים, אם כי מומלץ לעבוד על הפרויקט בזמן הסמסטר.
- כתלות במספר הנרשמים נמליץ על עבודה בזוגות הן לצורך הצגת מאמר והן לשם ביצוע הפרויקט.

שעות קבלה:

ניתן לתאם עם המרצה בכתובת דוא"ל yromano@technion.ac.il.