
האוניברסיטה העברית בירושלים

סילבוס

סמינר בהסקה סיבתית לניתוח מאגרי נתונים - 67709

תאריך עדכון אחרון 10-08-2023

נקודות זכות באוניברסיטה העברית: 2

תואר: מוסמך

היחידה האקדמית שאחראית על הקורס: מדעי המחשב

השנה הראשונה בתואר בה ניתן ללמוד את הקורס: 0

סמסטר: סמסטר ב'

שפת ההוראה: אנגלית

קמפוס: קרית א"י ספרא

מורה אחראי על הקורס (רכז): ד"ר אמיר גלעד

דוא"ל של המורה האחראי על הקורס: amirg@cs.huji.ac.il

שעות קבלה של רכז הקורס: ייקבע במייל

מורי הקורס:

תאור כללי של הקורס:

נהוג לומר ש"מתאם אינו מעיד על סיבתיות" (Causation imply not does Correlation). אולם, היקף ההשפעה של הסקה סיבתית (Analysis Causal) נרחב הרבה יותר מגילוי קורלציות, אסוציאציות, וניתוח תחזיות של מודלים לומדים. הסקה סיבתית חשובה במיוחד לתחומים כמו מדעי החברה, בריאות, וכו'. לדוגמה, שיטה זו תשמש חוקר ביו-רפואי המעוניין לגלות את ההשפעה של תרופה חדשה על סוג ספציפי של סרטן, או כלכלן שמעוניין לגלות האם תכנית התמחות מסוימת משפרת את סיכויי התעסוקה. ניתוח סיבתי מהווה את הבסיס לקביעת מדיניות נכונה באמצעות שיטות להערכת ההשפעות של "התערבות" מסוימת על העולם. הדרך הטובה ביותר לבצע ניתוח סיבתי היא על ידי ניסויים מבוקרים עם הקצאות אקראיות, אך ביצוע ניסויים כאלה לא תמיד אפשרי בשל סיבות אתיות, שיקולי עלות, או בשל אילוצי המציאות. על כן, ניסויים תצפיתיים או דאטה תצפיתי משמשים במקרים רבים לביצוע ניתוח סיבתי. אולם, הדאטה יכול לגרום לנו להגיע מסקנות שונות מאוד כתלות באופן שבו מנתחים אותו. בסמינר זה, ינתן רקע קצר לניתוח סיבתי פורמלי המתבסס על דאטה תצפיתי. בהמשך, הסטודנטים יציגו שיטות מתקדמות לניתוח סיבתי ושימושים עדכניים לניתוח סיבתי עבור ניתוח דאטה (Analysis Data), הגינות (Fairness), ואף להסברים לאינטליגנציה מלאכותית (XAI).

מטרות הקורס:

הסטודנטים יכירו מאמרים וכלים מרכזיים בתחום ההסקה הסיבתית וכיצד ניתן להשתמש בהם לצורך ניתוח מאגרי נתונים, מתן הסברים לתהליכים במאגרי נתונים ובמודלים של אינטליגנציה מלאכותית. לומדים מודלים של (Fairness) להגנות קריטריונים ויצירת (XAI).

תוצרי למידה

בסימו של קורס זה, סטודנטים יהיו מסוגלים:

בסיום הסמינר, הסטודנטים ידעו להבין ולהציג מאמרים בתחום ההסקה הסיבתית ויכירו את יישומים מרכזיים של תחום זה לצרכי ניתוח מאגרי נתונים ולמידת מכונה.

דרישות נוכחות (%) :

80

שיטת ההוראה בקורס: בתחילת הקורס תינתן רשימת מאמרים לבחירת הסטודנטים. המאמרים יוצגו ע"י הסטודנטים בקורס באנגלית. משך הרצאה בודדת הינו שעה אקדמית. כל הרצאה תתחיל עם הידע המקדים הנדרש להבנת המאמר או המאמרים שיוצגו. לאחר מכן, ההרצאה תמשיך לפרט על המאמר או המאמרים שנבחרו בהתייעצות עם המורה. טיטה של כל מצגת תועבר למורה לפחות שבוע לפני מועד ההרצאה להערות ואישור.

רשימת נושאים / תכנית הלימודים בקורס:

רקע על הסקה סיבתית הכולל את המוטיבציה, מודלים סיבתיים, הנחות יסוד, קריטריונים להסקה

סיבתית נכונה ממסדי נתונים קיימים, ושימושים בהסקה סיבתית במסדי נתונים ובאלגוריתמים לומדים.

חומר חובה לקריאה:

Judea Pearl, Madelyn Glymour, Nicholas P. Jewell, *Causal Inference in Statistics: A Primer*

חומר לקריאה נוספת:

1. Judea Pearl - *Causality - Models, Reasoning, and Inference*, 2nd Edition, 2009
2. Judea Pearl and Dana Mackenzie: *The Book of Why: The New Science of Cause and Effect*

מרכיבי הציון הסופי:

מצגת / הצגת פוסטר / הרצאה / סמינר / פרוסמינר / הצעת מחקר 70 %
נוכחות / השתתפות בסיום 30 %

מידע נוסף / הערות: