
האוניברסיטה העברית בירושלים

סילבוס

נושאים בביג דאטה בכלכלה - 57493

תאריך עדכון אחרון 05-03-2023

נקודות זכות באוניברסיטה העברית: 4

תואר: בוגר

היחידה האקדמית שאחראית על הקורס: כלכלה

השנה הראשונה בתואר בה ניתן ללמוד את הקורס: 0

סמסטר: סמסטר א' או / ו ב'

שפת ההוראה: עברית

קמפוס: הר הצופים

מורה אחראי על הקורס (רכז): שרית אגמי

דוא"ל של המורה האחראי על הקורס: sarit.agami@mail.huji.ac.il

שעות קבלה של רכז הקורס:

מורי הקורס:

תאור כללי של הקורס:

הקורס דן בשיטות לניתוח נתונים גדולים, כלומר נתונים בהם מספר התצפיות הוא גדול ו/או מספר המשתנים הוא גדול. מטרת הקורס הוא להכיר את השיטות וליישמן על אוספי נתונים אמיתיים. היישום יעשה באמצעות התוכנה R.

מטרות הקורס:

תוצרי למידה

בסיומו של קורס זה, סטודנטים יהיו מסוגלים:
לנתח נתונים גדולים: תיאור והסקת מסקנות מהן.

דרישות נוכחות (%) :
90%

שיטת ההוראה בקורס:

רשימת נושאים / תכנית הלימודים בקורס:

1. *Principal Component Analysis (PCA)*
 - *Population principal components, sample principal components, visualizing principal components.*
 - *The Kaiser-Meyer-Olkin Measure of Sampling Adequacy, Bartlett's test of sphericity*
 - *Principal components regression (PCR)*
 - *Robust PCA*
 - *PCA on discrete data*
2. *Canonical Correlation Analysis (CCA)*
 - *Population canonical correlations, sample canonical correlations*
 - *Canonical correlations and transformed data*
 - *Canonical correlations and regression*
3. *Discriminant Analysis*
 - *Linear discriminant rules*
 - *Discrimination under Gaussian assumptions*
 - *PCA, discrimination and regression*
4. *Cluster Analysis*
 - *Hierarchical Agglomerative Clustering*
 - *k-Means Clustering*
 - *Principal Components and Cluster Analysis*

-
- 5. Factor Analysis
 - Population k-Factor Model, Sample k-Factor Model
 - Factor Loadings
 - Factor Scores and Regression
 - 6. Multidimensional Scaling
 - 7. Graphical models for high-dimensional data
 - 8. Ridge regression
 - 9. Lasso

חומר חובה לקריאה:
אין

חומר לקריאה נוספת:
Multivariate analysis
Mardia, Kantilal Varichand; London:Academic Press;1979

An introduction to applied multivariate analysis with R
Everitt, Brian;New York: Springer; 2011

הערכת הקורס - הרכב הציון הסופי :
מבחן מסכם בכתב/בחינה בעל פה 0 %
הרצאה 10 %
השתתפות 10 %
גשת עבודה 0 %
גשת תרגילים 5 %
גשת דו"חות 0 %
פרויקט מחקר 75 %
בחנים 0 %
אחר 0 %

מידע נוסף / הערות:

במהלך הקורס יינתנו תרגילי בית ליישום הנלמד. יינתן ציון על כל תרגיל, כאשר המשקל של התרגילים יהיה 5% מהציון הסופי. בנוסף, הסטודנטים יידרשו להעביר הצגה קצרה בכיתה של מאמר המיישם שיטות בביג דאטה, משקל ההצגה הינו 10% מהציון הסופי. במהלך הסמסטר כל סטודנט יבחר נושא לכתיבת עבודה סמינריונית לפי המעניין אותו או מתוך אוסף נושאים שיוצעו לבחירה, ויתקיימו מפגשים פרטניים מול כל סטודנט כדי לעבוד על הנושא שבחר.
* השיעורים יוקלטו, והקלטות השיעורים תהיה זמינות לכל הסטודנטים בכל עת