

בית הספר למוסמכים במינהל עסקים ע"ש ליאון רקנאטי

1242.3253.02 – נושאים מתקדמים למדעי הנתונים למנהל עסקים (לשעבר: נושאים מתקדמים בכריית מידע וגילוי ידע)

דרישות קדם: מבוא למדע הנתונים למנהל עסקים או מבוא ליישומי דאטה במנהל עסקים
דרישה מקבילה: טיפול יישומי בנתוני אנליטיקה עסקית

סמסטר א' – תשפ"ג

קבוצה	יום בשבוע	שעה	תאריך בחינה	מרצה	דואר אלקטרוני	טלפון
02	ה'	15:45-18:30 (מחצית שנייה)	אין בחינה	ד"ר תומר גבע ד"ר רן תל-ניר		

שעת קבלה – בתיאום מראש

הערות לגבי לימוד עצמי ורקע קודם נדרש:

לביצוע הפרוייקט המסכם בקורס מומלץ להשתמש בחבילות כריית מידע וכתובת קוד בשפות Python או R. הקורס כולל מעבר מהיר ביותר על חבילת כריית מידע בפייתון וכולל חומר ללימוד עצמי בשפת R. לכן גם סטודנטים שלמדו R או פייתון בקורסי קדם או קורסים מקבילים יהיו צריכים במהלך הקורס לבצע לימוד עצמי משמעותי נוסף. הקורס מניח שלסטודנטים הבנה בסיסית לפחות בעקרונות תכנות וניסיון קודם (בסיסי לפחות) בתכנות.

היקף הלימודים

היקף הי"ס לקורס: 1

ECTS = 4 ECTS 1 י"ס – (European Credit Transfer and Accumulation System) ECTS, ערך הניקוד של הקורס במוסדות להשכלה גבוהה בעולם שהינם חלק מ"תהליך בולוניה".

תיאור הקורס

הקורס יתמקד בהצגת פתרונות Data Science ו Machine Learning לבעיות מבוססות נתונים בתחום העסקי. הקורס מאפשר התנסות מעשית בשימוש בכלים ואלגוריתמים מתקדמים וקבלת בסיס מעשי ותיאורטי בתחום. הקורס הינו קורס מתקדם המיועד להרחיב את הידע לסטודנטים שלמדו את קורס הקדם. בהרצאות הקורס יוצגו בעיות נבחרות וילמדו שיטות ואלגוריתמים של כריית מידע בדגש על Supervised Learning ו Predictive Modeling. הלימוד בקורס מורכב מהרצאות בכיתה, מעבודות תלמידי הקורס, וע"י לימוד ותרגול עצמי (קריאת חומר, מעבר אופציונלי על הרצאות מוקלטות, תרגול ולימוד שימוש בכלי machine learning ב R או בפייתון). כמו כן, יכלול הקורס פרוייקט מסכם בהיקף נרחב בו יישמו התלמידים את הנלמד בקורס.

תפוקות למידה

עם סיום הקורס בהצלחה הסטודנט:

1. יקבל רקע רחב לגבי בעיות, כלים ושיטות עבודה בתחום Business Analytics ו Data Science
2. יקבל בסיס מקצועי לעבוד באופן עצמאי עם כלי כריית מידע Data Science
3. יכירו חבילת כריית מידע בפיייתון או R
4. יקבלו בסיס להמשיך להתפתח בתחום באופן עצמאי

הערכת הסטודנט בקורס והרכב הציון

אחוז	מטלה	תאריך	גודל קבוצה/ הערות
100%	פרוייקט מסכם	הגשה עד שבוע מיום הלימודים האחרון בסמסטר	הפרוייקט הינו במתכונת של בחינת בית. ההגשה בזוגות בלבד!

- בקורס ינתנו מספר תרגילי בית רשות כולל תרגיל המהווה תחרות בין קבוצות.
- יתכן בונוס אישי של עד 15 נק' על הצטיינות בתחרות, פתרון תרגילי בית רשות אתגריים (יצוינו באופן ספציפי), השתתפות ותרומה משמעותית לדיון בהרצאות, ומענה על שאלונים וסקרים בקורס.
- הציונים בפרוייקט יחושבו בעיקר לפי ביצועים כמותיים
- בהרצאת אורח ובהרצאות דיון יש חובת נוכחות. מידע לגבי מועדי הרצאות אלו ינתן במהלך הסמסטר.
- למרצה זכות לדחות את מועד הגשת הפרוייקט

מדיניות שמירה על טווח ציונים

החל משנה"ל תשס"ט מונהגת בפקולטה מדיניות שמירה על טווח ציונים בקורסי התואר השני. עקרונות השיטה חלים על כל קורסי התואר השני, ומדיניות השמירה על טווח הציונים תיושם לגבי הציון הסופי בקורס זה. מידע נוסף בנושא זה מתפרסם בהרחבה באתר הפקולטה. [לתקנוני מדיניות שמירת טווח ציונים](#)

הערכת הקורס ע"י הסטודנטים

בסיומו של הקורס הסטודנטים ישתתפו בסקר הוראה על מנת להסיק מסקנות לטובת צרכי הסטודנטים והאוניברסיטה.

אתר הקורס

אתר הקורס יהווה המקום המרכזי בו ימסרו הודעות לסטודנטים, לפיכך מומלץ להתעדכן בו מדי שבוע, לפני השיעור, ובכלל – גם בתום הסמסטר. (לצורך תיאום עינייני הבחינה למשל). שקפי הקורס יהיו באתר הקורס באתר. לתשומת לבכם - בכיתה ידונו גם נושאים (ובפרט דוגמאות) שאינם מופיעים בשקפים או מופיעים בכותרת בלבד. כל אלו הינם חלק בלתי נפרד מחומר הקורס.

נושאים	נושא
Introduction	1
Basic Algorithms - Search methods (e.g, Gradient Descent) for regression and classification	2
Advanced Model Evaluation Consideration and Practical Implementation Concerns	3
Feature Selection and Regularization (Ridge regression and LASSO)	4
Basic Text Mining - Document representation , document categorization, Naïve Bayes algorithm filtering application	5
Expected Value Framework	6
Ensemble Methods: Random Forests, Bagging, Gradient Boosting, Adaboost	7
Introduction to scikit-learn (machine learning in Python)**	8
Introduction to Neural Networks, Deep Learning and the Backpropagation algorithm.	9
Clustering	10
Network-based Mining	11
Advanced Implementation Tips	12

* התכנית הינה בסיס לשינויים. סדר ההרצאות לא בהכרח תואם לרשימת הנושאים. חלק מההרצאות ימשכו יותר משיעור אחד
 ** אופציונלי: לסטודנטים המעוניינים לעבוד עם R ינתנו חומרי הדרכה ללימוד עצמי והרצאות מוקלטות בנושא.

הערות:

- קיימת אפשרות שחלק מההרצאות הקורס יוחלפו בהרצאות אורח /הרצאות על נושאים אקטואליים . בהתאם לנסיבות.
- במידת הצורך, בקורס צפויות עד שתי הרצאות השלמה בימי השישי בבוקר במהלך הסמסטר. עדכון לגבי מועדי ההרצאות (במידה ויועברו) ינתן במהלך הסמסטר.
- במידת הצורך, יתכן והרצאה אחת או שתיים יועברו באופן מתקשב

ספרי לימוד

"An Introduction to Statistical Learning: With Applications in R." (Springer 2014). James, Gareth, Daniela Witten, and Trevor Hastie.

"The Elements of Statistical Learning", Second edition, Trevor Hastie, Robert Tibshirani, Jerome H. Friedman (2009)

חומר קריאה רלוונטי נוסף ימסר במהלך הסמסטר