

בית הספר למוסמכים במינהל עסקים ע"ש ליאון רקנאטי

1242.3253.01 – נושאים מתקדמים למדעי הנתונים למנהל עסקים (לשעבר: נושאים מתקדמים בכריית מידע וגילוי ידע)

(דרישות קדם: מבוא ליישומי דאטה במנהל עסקים או מבוא למדע הנתונים למנהל עסקים)

סמטר א' – תשפ"א

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

שעת קבלה – בתיאום מראש

הערות:

לביצוע הפרוייקט המסכם בקורס מומלץ להשתמש בכלי תכנות כגון R. הקורס כולל מעבר מהיר ביותר על יסודות שפת R לצורך משימות כריית מידע ומחייב לימוד עצמי משמעותי. הקורס מניח שלסטודנטים הבנה בסיסית לפחות בעקרונות תכנות וניסיון קודם (בסיסי לפחות) בתכנות. לכן: **לסטודנטים ללא ניסיון/לימוד קודם של תכנות (ברמה בסיסית לפחות) מומלץ ללמוד לפני או במקביל את הקורס "טיפול יישומי בנתוני אנליטיקה עסקית" (או לחילופין את "הקורס כריית מידע עם יישומים ב R").**

היקף הלימודים

היקף ה"ס לקורס: 1

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

תיאור הקורס

הקורס יתמקד בהצגת פתרונות כריית מידע לבעיות מבוססות נתונים בתחום העסקי. הפתרונות מתבססים על יישום תאוריות וכלים מתחום ה Data Science וכריית המידע (Data Mining). הקורס מאפשר התנסות מעשית בשימוש בכלי כריית מידע וקבלת בסיס מעשי ותיאורטי בתחום. הקורס הינו קורס מתקדם המיועד להרחיב את הידע בכלי כריית מידע ולסטודנטים שלמדו את הקורס מבוא ליישומי דאטה במנהל עסקים. 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 ותרגול שימוש בתוכנת כריית מידע - R). כמו כן, יכלול הקורס פרוייקט מסכם בו יישמו התלמידים את הנלמד בקורס באמצעות תוכנת כריית מידע.

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. יכיר תוכנת כריית מידע
4. יקבל 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 להם ינתן פתרון לדוגמא וכן תרגיל המהווה תחרות בין קבוצות.
- יתכן בונוס אישי של עד 10 נק' על הצטיינות בתחרות ועל השתתפות ותרומה אישית משמעותית לדיון בהרצאות.
- 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.

<https://coller.tau.ac.il/MBA-students/programs/2019-20/MBA/regulations/exams>

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
Text Mining – LDA (optional)	6
Ensemble Methods: Random Forests, Bagging, Gradient Boosting, Adaboost	7
R Data Mining Workshop	8
Introduction to Neural Networks/Deep Learning and 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

הערות:

- קיימת אפשרות שחלק מהרצאות 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