



החוג לניהול / Dept. of Management

## 7024-1221 - כריית מידע עסקי

דרישות קדם: מבוא לאנליטיקה עסקית  
טיפול יישומי בנתונים (למעט סטודנטים הפטורים מקורס זה)

### פרטי הקבוצות בקורס

#### סמסטר א' – תשפ"א

קבוצה	יום בשבוע	שעה	תאריך בחינה מועד א+ב	מרצה	דואר אלקטרוני
01	ה	15:00-17:30	ללא בחינה- בקורס פרוייקט מסכם	ד"ר רונן מאירי	Ronen.meiri@outlook.com

עוזר הוראה (תרגול פרונטלי) - TBD

שעת קבלה – בתיאום מראש.

שימו לב שחלק ממפגשי ההרצאה יהיו משולבים עם תרגול העוסק בלימוד תוכנה

### היקף הלימודים

היקף 3 ש"ס לקורס

ECTS = 4 = 2 ש"ס – European Credit Transfer and Accumulation System (ECTS), ערך הניקוד של הקורס במוסדות להשכלה גבוהה בעולם שהינם חלק מ"תהליך בולוניה".

### תיאור הקורס

עולם ה Machine Learning צובר היום תאוצה בעולם העסקי והופך לכלי חשוב בהרבה ארגונים להשגת יתרון תחרותי אסטרטגי. המושג Machine learning מאפשר למכונות "ללמוד" או ליצר תהליכים קבלת החלטות מבוססי נתונים מבלי הצורך המפורש לקודד אוסף של חוקים או תהליכים. השימוש ב ML כיום מתרחב מעבר לשימושים העסקיים הקלסיים (של שיווק ממוקד, וניהול סיכונים אשראי) לתחומים חדשים לרבות: פענוח בדיקות ברפואה, פרמצטיקה, מסחר, מכונות אוטונומיות, Cyber Security ושימושים נוספים אחרים.

הקורס הינו קורס מתקדם המיועד להרחיב את הידע בכלים של Machine Learning לסטודנטים שלמדו את הקורס מבוא לאנליטיקה עסקית. הקורס יהווה נקודת כניסה מעשית לעולם ה Machine Learning ויתמקד בסקירה תיאורתית בסיסית והתנסות מעשית באלגוריתמים ומתודולוגיות שונות, תוך יישום הטכנולוגיות על בעיות עסקיות ממשיות. במהלך הקורס ילמדו של מודלים שונים מעולמות של Supervised Learning & Predictive analytics ושיטות של Unsupervised Learnign כמו Cluster Analysis. במהלך הקורס ילמדו שיטות שונות לאמידת איכות המודל, הערך הכלכלי של מודל, Hyperparametrization, Ensampling, Boosting, רשתות נוירונים ועוד. זהו קורס חשוב לכל מי שרצה להעמיק בעולמות של Machine Learning באופן מעשי הן באקדמיה והן בתעשייה.

הלימוד בקורס מורכב מהרצאות בכיתה, תרגילים קצרים על הנושאים שנלמדים בקורס, שאלות רשות לאלו שמעוניינים להעמיק את הידע שלהם ב machine learning ופרויקט/עבודת גמר מסכמת על החומר שנלמד בקורס

בחלקו השני של הקורס, במהלך מפגש או שניים, אנו נתמקד בהכנה נדרשת לקורס ה"סמינר באנליטיקה עסקית" שיעבר בסמסטר ב.

### הערות:

חובת הגשה של 80% מהתרגילים בקורס.  
 בנוס על פתרון כל שאלות הבנוס.  
 עבודת הגמר בזוגות.

לביצוע הפרוייקט המסכם בקורס יש צורך להשתמש בכלי תכנות. הקורס כולל מעבר מהיר על קוד/תוכנה/ות כריית מידע ומחייב לימוד ותרגול עצמי בהיקף משמעותי. הקורס מניח שלסטודנטים הבנה בסיסית לפחות בעקרונות תכנות וניסיון קודם (בסיסי לפחות) בתכנות בשפת python.

## תפוקות למידה

עם סיום הקורס בהצלחה הסטודנט:

1. יקבל רקע רחב לגבי עם בעיות, כלים ושיטות עבודה בתחום
2. יקבל בסיס מקצועי לעבוד באופן עצמאי עם כלי Machine Learning
3. יכיר תוכנת כריית מידע
4. יקבל בסיס להמשיך להתפתח בתחום באופן עצמאי
5. יקבל הכנה לסמינר באנליטיקה עסקית

## הערכת הסטודנט בקורס והרכב הציון

אחוז	מטלה	תאריך	גודל קבוצה/ הערות
5%	הגשת 80% מהתרגילים שה"כ יהיו כ 10 תרגילים בקורס	במהלך הסמסטר	פרטני
85%	פרוייקט מסכם	הגשה עד תחילת המפגש האחרון של הקורס	בקבוצות של שני סטודנטים
10%	השתתפות פעילה כולל במפגשי ההכנה לקורס הסמינר		הגשת הצעת הכנה לקורס הסמינר תהיה ברביעיות רביעיות בלבד (ינתן אישור חריג לקבוצה אחת קטנה יותר במקרה שמספר הסטודנטים בקורס יחייב זאת). הקבוצות ימשיכו בקורס הסמינר באנליטיקה עסקית

\* בקורס חובת נוכחות בכל ההרצאות והמפגשים. הנוכחות תבדק מעת לעת. סטודנטים שלא יהיו נוכחים בזמן בדיקת נוכחות בנסיבות שאינן מוצדקות ע"פ תקנון האוניברסיטה (כגון מילואים או מחלה) – יקבלו ציון לא השלים בקורס.

\* בקורס ינתנו מספר תרגילי רשות להם ינתן בנוס אישי נוסף של עד 5 נק' על הגשת כל תרגילי הרשות.

\* הפרוייקט המסכם נועד לבדוק את הידע האישי של הסטודנטים. הפרוייקט מבוסס על ידע מהקורס הנוכחי ומקורסי הקדם. עיקר הציון בפרוייקט יהיה על סמך איכות הביצועים בפתרון בעיית כריית מידע

\* תתכן בהתאם לשיקול דעת המרצה, בחינת בקיאות קצרה על תוכן הפרוייקט (לכל או חלק מהסטודנטים). מועד הבחינה יקבע בהמשך ע"י המרצה והמזכירות. ציון בחינה זו ישוקלל בציון הפרוייקט

## מדיניות שמירה על טווח ציונים

בחוג לניהול מונהגת מדיניות שמירה על טווח ציונים. מדיניות זו מתייחסת לממוצע הציונים הסופיים בקורס.

מידע בנושא זה מתפרסם בהרחבה באתר החוג לניהול, בסעיף ציונים בתקנון.

## 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 ענייני 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.

## תכנית הקורס \*

נושא	נושאים
1	Introduction
2	Basic Algorithms <ul style="list-style-type: none"><li>- Search methods for Regression algorithms</li><li>- The loss function and optimization using Gradient Decent algorithm</li></ul>
3	Performance measurements for model Evaluation and Practical Implementation Concerns
4	Overfitting and Evaluation
5	Feature Selection and Regularization
6	Ensamble models and Random Forest
7	Boosting Methods
9	Cluster Analysis
9	Basic Text Mining <ul style="list-style-type: none"><li>- Document representation , document categorization, Naïve Bayes algorithm filtering application</li></ul>
10	Introduction to Neural Networks/Deep Learning
11	Data Mining Software Tutorial

\*הרשימה הינה רשימת נושאים ולא סדר הרצאות. חלק מהנושאים יועברו במשך מספר הרצאות. ההתכנית הינה בסיס לשינויים.

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

לעיון נוסף וההעמקה

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

חומר קריאה רלוונטי נוסף ימסר במהלך הסמסטר.