



מכון טכנולוגי חולון
Holon Institute of Technology

מוסמך במדעים (M.Sc.) במדעי הנתונים

סילבוסים

תואר שני ללא תזה

מאי 2022

תוכן עניינים

3.....	אבטחה והגנה על מאגרי נתוני עתק
4.....	אתיקה במדעי הנתונים
6.....	יישומים של מדעי הנתונים ברובוטיקה וב-IOT
7.....	יישומים של מדעי הנתונים באפידמיולוגיה ובריאות הציבור
8.....	כריית טקסטים
9.....	כריית נתונים מתקדמת
10.....	למידה לא מונחית מתקדמת
12.....	למידה עם חיזוק
13.....	למידה עמוקה
14.....	למידה עמוקה ליישומי ראייה ממוחשבת
15.....	למידת מכונה
16.....	למידת מכונה למערכות סייבר
17.....	מסחר אלקטרוני
18.....	מערכות תומכות החלטה
19.....	מתמטיקה למדעי הנתונים
20.....	ניהול והנדסת תשתיות נתוני עתק בענן
22.....	ניתוח רשתות חברתיות
23.....	סדרות עתיות ושיטות חיזוי
25.....	סמינריון
26.....	עיבוד אותות
27.....	עיבוד אותות ביו-רפואיות
28.....	עיבוד אותות על גרפים
29.....	עיבוד שפה טבעית
31.....	עיבוד שפה טבעית מתקדמת
33.....	עיבוד תמונות
34.....	עיבוד תמונות ביו-רפואיות
35.....	עקרונות למדעי הנתונים ובסיסי הנתונים
36.....	סילבוס מורחב לביצוע עבודת גמר לתואר שני במדעי הנתונים
37.....	כללי:

37	הנחות יסוד:
38	שלבי הפרויקט ומטלות (סיום כל מטלה מותנה באישור המנחה):
39	אבני דרך (עדכונים ללוחות הזמנים באתר הקורס):
41	טיפים לעבודה נכונה:
42	נספח ג': כתיבת מסמכים בתוכנית לתואר שני במדעי הנתונים:
42	נספח ד': מבנה הצעת המחקר - מטלה 2
45	נספח ה': הנחיות לסקר ספרות מטלה 3
47	נספח ו' - מבנה ותבנית לדו"ח עבודת גמר
49	נספח ז': דף כותרת להגשת עבודת גמר
50	נספח ח': רשימת תיוג
52	נספח ט' - בחינת הגמר
53	רכב אוטונומי
55	שיטות בייסיאניות בלמידה
56	שיטות הסתברותיות וסטטיסטיות מתקדמות למדעי הנתונים
58	שיטות מתקדמות בלמידת מכונה

אבטחה והגנה על מאגרי נתוני עתק Protection Security of Big Data

קוד הקורס: 90212

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם: אין דרישת קדם

סוג הקורס: שיעור, מעבדה, ולמידה מבוססת פרויקט PBL

מטרות הקורס:

הקורס מיועד להקנות לסטודנטים היכרות עם עולם האבטחה וההגנה על מאגרים של נתוני עתק, והבנה של הכלים הקיימים המאפשרים פעילויות מניעתיות, תגובתיות ושיקומיות.

נושאי הקורס:

במהלך הקורס הסטודנטים ילמדו מצד אחד, להבדיל בין הגנת המידע לבין אבטחת מידע, ומצד שני, שיטות המאפשרות למנוע ולזהות אירועים חריגים היכולים לגרום נזק למערכות מידע ובמיוחד למאגרי נתוני עתק. במהלך הקורס יודגשו האתגרים של ההתקדמויות הנוכחיות והעתידיות בשימוש במאגרים של נתוני עתק ושירותי ענן בהיבטים של אתיקה, אבטחת מידע ופרטיות. הסטודנטים יתנסו בבעיות מחקריות ויישומים אמיתיים תוך דגש על פרויקט מסכם של הקורס (בשיטת למידה מבוססת פרויקט) ולפי מסלול ההתמחות

ביבליוגרפיה:

1. ISO/IEC (218), Information technology — Security techniques — Information security management systems — Overview and vocabulary, ISO/IEC 27000:2018(E), Fifth edition, 2018-02, https://standards.iso.org/ittf/PubliclyAvailableStandards/c073906_ISO_IEC_27000_2018_E.zip
2. Dunning T., Friedman E., (2016) Sharing Big Data Safely: Managing Data Security, O'Reilly Media

אתיקה במדעי הנתונים Ethics and data Science

קוד הקורס: 90206

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם: אין דרישת קדם

סוג הקורס: שיעור

מטרות הקורס:

לאתיקה שורשים עתיקים המשקפים את תפקידה המרכזי בחיי האדם. בינה מלאכותית (AI) התפתחה באופן משמעותי רק במחצית המאה האחרונה. צמיחה מתפרצת, הכוח והפוטנציאל של AI בעשור האחרון מעלה את החשיבות של שקיפות, מוסריות ואתיקה של מערכות מבוססות AI. תוך שנים מעטות מגוון רב של מערכות שמשפיעות על חיינו החל משווקים פיננסיים, תחבורה, בריאות ועד מערכות צבאיות ישתמשו בנתוני עתק ובבינה מלאכותית לקבלת החלטות. לצורך אימון ובניית מודלים, מערכות אלו ישתמשו בנתוני עתק. חלקם הגדול יהיו גם נתונים פרסונליים. מי יבקר וישלוט בנתונים הרגישים? מכאן עולה הצורך והחשיבות של הקניית ידע ומודעות להיבטים של אתיקה למדעני עתיד בתחום זה.

נושאי הקורס:

- שינויים סוציו אקונומיים, תרבותיים, טכנולוגיים ואתיקה
- הגדרת נתוני עתק, נתונים אישיים ונתונים אנונימיים
- נתונים אישיים, חוקיות, יושרה, סודיות, הוגנות, שקיפות ו- General Data Protection Regulation (GDPR).
- חיי נתוני עתק: נתונים שונים ומגוונים כגון: דמוגרפיים, סוציו- אקונומיים, השכלה, תעסוקה, בריאות, חיי פנאי, צריכה של שירותים מקוונים, בידור וכו' נאספים לאורך חיינו משלב העובר ועד מותנו.
- סוגיות אתיות: מודעות, שליטה ובקרה, אמון, בעלות, ניטור, מעקב והבטחה, זהות דיגיטלית, אנונימיות, פרטיות.
- דוגמאות לשימוש בנתוני עתק בעולמות הרפואה, חינוך, ערים חכמות, IOT, כלכלה ומסחר.
- מה אנו צריכים לדעת על אתיקה?
- סוגיית אתיקה בבינה מלאכותית, מדוע עכשיו?
- קודי אתיקה מקצועית, אכיפה וסמכות, ערכים מקצועיים.
- השפעת בינה מלאכותית על אתיקה מקצועית.
- סוגיית הטייה (bias) של אלגוריתמים.
- כיצד להיות מדען נתונים תוך שמירת אתיקה.
- הצגת פרויקט הקורס ע"י סטודנטים, דיון קבוצתי.

• הצגת פרויקט הקורס ע"י סטודנטים, דיון קבוצתי.

ביבליוגרפיה:

1. Paula Boddington ,Towards a Code of Ethics for Artificial Intelligence , Springer International Publishing AG, 2017
2. Mark Coeckelbergh ,AI Ethics, MIT Press, 2020
3. Mike Loukides ,Hilary Mason and DJ Patil, Ethics and Data Science, O'Reilly Media, Inc, 2018

יישומים של מדעי הנתונים ברובוטיקה וב-IOT Data Science applications in robotics and the Internet of Things

קוד הקורס: 90209

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם: למידה עם חיזוק

סוג הקורס: שיעור ולמידה מבוססת פרויקט PBL

מטרות הקורס:

הקורס מיועד להקנות לסטודנטים היכרות עם שימוש במדעי הנתונים בעולם הרובוטיקה ואינטרנט של הדברים.

נושאי הקורס:

במהלך הקורס הסטודנטים ילמדו מגוון רחב של יישומי מדעי הנתונים ברובוטיקה ובמרחבת הדברים (Internet of things). הנושאים המרכזיים הם: תכנון מערכת לאיסוף נתונים (תוך שימוש בחיישנים, מפעילים וזרוע רובוטית), בניית מערכת, איסוף נתונים, עיבודיהם, ניתוחיהם וייצוגיהם במסגרת תהליך קבלת החלטות. הסטודנטים יתנסו בזרוע רובוטית ומערכות IoT תוך דגש על פרויקט מסכם של הקורס (בשיטת למידה מבוססת פרויקטים) ולפי מסלול ההתמחות.

ביבליוגרפיה:

1. Dobot Magician User Manual v1.7.0 (Up-to-date version available at:
https://www.dobot.cc/downloadcenter.html?sub_cat=73#sub-download)
2. Hock, O., Sedo, J. (2017). Forward and inverse kinematics using pseudoinverse and transposition method for robotic arm dobot. In: Kinematics. IntechOpen
3. Holler, J., Tsiatsis, V., Mulligan, C., Avesand, S., Karnouskos, S., Boyle, D. (2014). From Machine-to-machine to the Internet of Things: Introduction to a New Age of Intelligence. Academic Press.
4. Ferrarini, P., & Luccio, F. (2018). Computational Thinking: First Algorithms, Then Code. Springer.

5. McKenzie, R. M., Barraclough, T. W., & Stokes, A. A. (2017). Integrating soft robotics with the robot Operating system: a hybrid Pick and Place arm. *Frontiers in Robotics and AI*, 4, 39.

יישומים של מדעי הנתונים באפידמיולוגיה ובריאות הציבור Applications of Data Science in Epidemiology and Public Health

קוד הקורס: 90208

היקף הקורס: 3 ש"ס, 3 נ"ד

דרישות קדם: כריית נתונים מתקדמת או כריית טקסטים או למידה לא מונחית מתקדמת

סוג הקורס: שיעור ולמידה מבוססת פרויקט PBL

מטרות הקורס:

הקורס מיועד להקנות לסטודנטים היכרות עם שימוש במדעי הנתונים בעולם בריאות הציבור ובדגש על תחום האפידמיולוגיה.

נושאי הקורס:

במהלך הקורס הסטודנטים ילמדו מגוון רחב של יישומים של מדעי הנתונים בבריאות הציבור ובאפידמיולוגיה. הנושאים המרכזיים הם: מקורות הנתונים הבריאותיים והרפואיים השונים, ארגונייהם, עיבודיהם, ניתוחיהם וייצוגיהם במסגרת תהליך קבלת החלטות. הסטודנטים יתנסו בבעיות מחקריות ויישומים אמיתיים תוך דגש על פרויקט מסכם של הקורס (בשיטת למידה מבוססת פרויקטים) ולפי מסלול ההתמחות.

ביבליוגרפיה:

1. Shortliffe E. H., Cimino J. J. (2014), *Biomedical Informatics: Computer Applications in Health Care and Biomedicine*, Springer-Verlag
2. Venot A., Burgun A., Quantin C. (2014), *Medical Informatics, e-Health: Fundamentals and Applications*, Springer-Verlag

3. Marconi, K. and Lehmann, H. (2014), Big Data and Health Analytics, Taylor & Francis
4. Kudyba S.P. (2016), Healthcare Informatics: Improving Efficiency through Technology, Analytics, and Management, CRC Press

כריית טקסטים Text Mining

קוד הקורס: 90102

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם: למידה עמוקה

סוג הקורס: שיעור

מטרות הקורס:

מטרת הקורס היא להכיר ללומד את הכלים והשיטות לניתוח של טקסטים בסוגים ובאופנים שונים וחילוץ מידע וידע מתוך טקסטים אלו בשיטות חישוביות. הקורס יסקור את השיטות השונות ליצוג טקסטים (Embedding Word, BOW), שיטות לסיווג טקסטים, חילוץ מידע וישויות, שיטות לא מונחות לניתוח טקסטים, תרגום ועוד. הקורס ישלב שיטות מתחום ה NLP (עיבוד שפה טבעית) ניתוחים סטטיסטיים ושיטות עדכניות של למידה עמוקה.

נושאי הקורס:

- עקרונות בניתוח טקסט
- שיטות ליצוג טקסט (Embedding Word, BOW)
- ניתוח תחבירי וסמנטי של טקסט
- שיטות מתקדמות לסיווג טקסטים (RNN, LSTM, ATTENTION, BERT)
- שיטות מתקדמות חילוץ תבניות וישויות
- אישכול טקסטים MODELLING TOPIC
- תרגום טקסטים

ביבליוגרפיה:

1. Neural Network Methods in Natural Language Processing (Synthesis Lectures on Human Language Technologies): Yoav Goldberg)2017, Graeme Hirst: 9781627052986

2. Speech and Language Processing 3rd edition. Dan Jurafsky and James H. Martin. 2017.

3 . מאמרים עדכניים נוספים כפי שיתפרסמו מעת לעת

כריית נתונים מתקדמת Advanced Text Mining

קוד הקורס: 90105

היקף הקורס: 3 ש"ס 3 נ"ז

דרישות קדם: למידה עמוקה

סוג הקורס: שיעור

מטרות הקורס:

מטרת הקורס היא להכיר ללומד נושאים מתקדמים מתחום כריית הנתונים. הקורס יתמקד בנושאים נבחרים בעלי חשיבות בכריית נתונים ולמידת מכונה במגוון תחומים. השילוב בין כריית נתונים ולמידת מכונה הוא נושא חם שזוכה לאחרונה לפופולריות רבה בתחומי המחקר והבינה המלאכותית. הקורס יעסוק בנושאי כריית נתונים מתוך רשתות, קהילות, נתונים מבוססי זמן, מקום והתאמה אישית.

נושאי הקורס:

- כרייה של נתוני רשת וגרפים
- ניתוח קהילות
- כריית נתונים מבוססי מיקום
- כריית נתונים עיתיים (לפי ציר זמן)
- זיהוי אנומליות
- מערכות המלצה

ביבליוגרפיה:

1. Reza Zafarani, Mohammad Ali Abbasi, Huan Liu. Social Media Mining: An

Introduction. Cambridge University Press. (2014)

2. Pang-Ning Tan, Michael Steinbach, Anuj Karpatne, and Vipin Kumar. 2018.

Introduction to Data Mining (2nd Edition) (2nd ed.). Pearson.

3. מאמרים עדכניים נוספים כפי שיתפרסמו מעת לעת.

למידה לא מונחית מתקדמת Advanced Topics in Supervised Machine Learning

קוד הקורס: 90107

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם: למידת מכונה

סוג הקורס: שיעור

מטרות הקורס:

באופן כללי, התחום Learning Machine מתייחס לזיהוי אוטומטי של דפוסי נתונים. כבזה הוא מהווה קרקע פורייה להתפתחויות סטטיסטיות ואלגוריתמיות חדשות. מטרת הקורס היא לספק קורס מתקדם עם אופי מתמטי קפדני להתפתחויות אלה עם דגש על שיטות וניתוח שלהם.

נושאי הקורס:

התיאוריה הסטטיסטית של למידת מכונה:

- סיווג, רגרסיה
- Empirical Risk Minimization, Regularization
- Suprema of Empirical Processes

אלגוריתמיקה:

- Boosting
- Kernel Methods
- Convex Optimization

:Online Learning

- Online Convex Optimization
- Partial Information: Bandit Problems
- Blackwell's Approachability

ביבליוגרפיה:

1. I. Goodfellow, Y. Bengio and A. Courville. Deep Learning. MIT Press 2016.
2. G. Giraud, Introduction to High-Dimensional Statistics. Chapman and Hall / CRC, 2014.
3. S. Shalev-Shwartz, B.-D. Shai, Understanding Machine Learning: From Theory to Algorithms. Cambridge University Press, 2014.
4. S. Bubeck, N. Cesa-Bianchi. Regret Analysis of Stochastic and Nonstochastic Multi-armed Bandit Problems. Now Publishers Incorporate, 2012.

למידה עם חיזוק Reinforcement Learning

קוד הקורס: 90108

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם: למידת מכונה

סוג הקורס: שיעור

מטרות הקורס:

כדי להגשים את החלומות וההשפעה של AI דורשים מערכות אוטונומיות שלומדות לקבל החלטות טובות. למידת חיזוק היא פרדיגמה חזקה אחת לשם כך והיא רלוונטית למגוון עצום של משימות, כולל רובוטיקה, משחקים, מידול צרכנים ותחום הבריאות. קורס זה יקנה את היסודות לתחום למידה עם חיזוק והסטודנטים ילמדו על אתגרי הגישה והגישות, כולל הכללה וחקירה. באמצעות שילוב של הרצאות ומשימות כתובות וקידוד, התלמידים יתוודעו היטב לרעיונות וטכניקות מפתח עבור RL. המטלות יכללו את היסודות של למידת חיזוק וכן למידת חיזוק עמוקה - תחום חדש ומבטיח ביותר המשלב טכניקות למידה עמוקה עם למידת חיזוק.

נושאי הקורס:

- תהליכי מרקוב
- חיפוש והערכת פוליסה
- Q Learning
- Deep Q learning
- Fast reinforcement learning (Exploration/Exploitation)
- Batch Reinforcement Learning
- Monte Carlo Tree Search

ביבליוגרפיה:

1. Reinforcement Learning: An Introduction, Sutton and Barto, 2nd Edition, MIT Press, Cambridge, MA, 2018.
2. Reinforcement Learning: State-of-the-Art, Marco Wiering and Martijn van Otterlo, Eds.

למידה עמוקה Deep Learning

קוד הקורס: 90006

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם : למידת מכונה

סוג הקורס: שיעור

מטרות הקורס:

למידה עמוקה היא אחת המיומנויות המבוקשות ביותר בתחום מדע הנתונים והבינה המלאכותית. בקורס זה ילמד הסטודנט את היסודות של הלמידה העמוקה, יבין כיצד לבנות רשתות עצביות וילמד כיצד להוביל פרויקטים מצליחים של למידה עמוקה. בנוסף יילמדו נושאים מתקדמים כגון רשתות מעורבות, LSTM, RNNs, GAN, Adam, Dropout, ועוד.

נושאי הקורס:

- מבני רשתות עצביות: פרסטרון רב שכבתי, רשתות קונבולוציה, רשתות חוזרות, רשתות עם זכרון, רשתות שארית.
- אלגוריתם ה-backpropagation
- אלגוריתם stochastic gradient descent ותכונותיו.
- פונקציות מחיר ושיטות אופטימיזציה מסדר ראשון (גודל צעד משתנה, שיטת מומנטום, שיטת נסטרוב, אלגוריתם אדם) ואחרים.
- למידה עמוקה עבור בעיות סיווג ורגרסיה.
- מעבר מסיווג לזיהוי עצמים: שיטת (RNCC) region CNN והרחבותיה השונות, YOLO SSD, ושיטות מתקדמות אחרות.
- מודלים יוצרים, autoencoder, autoencoder variational – רשתות יריביות יוצרות GAN, wasserstien GAN, (GAN) מעגלי.

ביבליוגרפיה:

1. Deep Learning By Ian Goodfellow, Yoshua Bengio, Aaron Courville Online book, 2017
2. מאמרים עדכניים נוספים כפי שיתפרסמו מעת לעת

למידה עמוקה ליישומי ראייה ממוחשבת Deep Learning for Computer vision application

קוד הקורס: 90215

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם: עיבוד תמונות, מבוא לתכנות Python (קורס השלמה), למידת מכונה

סוג הקורס: שיעור

מטרות הקורס:

בשנים האחרונות מתחוללת מהפכת הבינה המלאכותית (Artificial Intelligence), המשפיעה על כל תחום בחיינו. אחד התחומים המרתקים הוא ראייה ממוחשבת, ובפרט זיהוי חזותי. המרכיב המרכזי המשמש כיום בבינה מלאכותית הוא אלגוריתמי לימוד מכונה (Machine Learning) המבוססים על מתודולוגיית הלמידה העמוקה (Deep learning). מטרת הקורס היא להציג באופן נרחב ומעמיק את מתודולוגיית הלמידה העמוקה ולאפשר לסטודנטים לרכוש ניסיון מעשי במימוש המתודולוגיה לשימושי ראייה ממוחשבת וזיהוי חזותי, תוך שימוש ב-TensorFlow ו-PyTorch כפלטפורמות למידה עמוקה מרכזית.

נושאי הקורס:

- סקירה כללית של תחום הראייה הממוחשבת
- מבוא לסיווג תמונה- גישה מבוססת נתונים, KNN, מסווג לינארי.
- אופטימיזציה- עקרונות בסיסיים, פונקציות מחיר ושיטות אופטימיזציה.
- רשתות נוירונים
- רשתות נוירונים קונבולוציוניות (CNN)
- אימון רשתות נוירונים
- הכרת סביבות תוכנה
- ארכיטקטורות CNN
- רשתות Recurrent Neural Networks
- שכבות Attention

ביבליוגרפיה:

1. Y. Bengio, I. Goodfellow, A. Courville, Deep Learning, MIT Press, 2016, available online at <http://www.deeplearningbook.org/>
2. Y. Dong, Li Deng, Automatic Speech Recognition A Deep Learning Approach, Springer-Verlag London, 2015.
3. Y. Goldberg, A Primer on Neural Network Models for Natural Language Processing, 2015, available online at <http://u.cs.biu.ac.il/~yogo/nnlp.pdf>.

למידת מכונה Machine Learning

קוד הקורס: 90004

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם: עקרונות למדעי הנתונים ובסיסי נתונים (במקביל)

סוג הקורס: שיעור

מטרות הקורס:

תלמידי הקורס יכירו את היסודות של למידת מכונה ויקבלו סקירה של מושגים, טכניקות ואלגוריתמים רבים בלימוד מכונה, החל מנושאים כמו סיווג ורגרסיה לינארית ומסתיים בנושאים עדכניים יותר כמו שיפור SVM, HMM, ורשתות בייסיאניות. הקורס ייתן לתלמיד את הרעיונות הבסיסיים והאינטואיציה העומדת מאחורי שיטות לימוד מכונות מודרניות, כמו גם הבנה קצת יותר פורמלית של איך, מדוע ומתי הם עובדים.

נושאי הקורס:

- רגרסיה (פשוטה ורב משתנית)
- רגרסיה לוגיסטית
- KNN
- עצי החלטה
- רשתות בייסיאניות
- SVM
- רשתות עצביות
- למידה לא מונחית
- צמצום ממדים

ביבליוגרפיה:

1. Giraud, Christophe. Introduction to High-Dimensional Statistics. Chapman and Hall / CRC, 2014. ISBN: 9781482237948.
2. Bubeck, Sebastien, and Nicolo Cesa-Bianchi. Regret Analysis of Stochastic and Nonstochastic Multi-armed Bandit Problems. Now Publishers Incorporate, 2012. ISBN: 9781601986269.
3. Shalev-Shwartz, Shai, and Shai Ben-David. Understanding Machine Learning: From Theory to Algorithms. Cambridge University Press, 2014. ISBN: 9781107057135

למידת מכונה למערכות סייבר Machine Learning for Cyber Systems

מספר קורס: 70214

היקף הקורס: 3 ש"ס, 3 נ"ד

דרישות קדם: כריית נתונים מתקדמת או כריית טקסטים

סוג הקורס: שיעור ולמידה מבוססת פרויקט PBL

נושאי הקורס:

במהלך הקורס הסטודנטים ירחיבו את הידע בלמידת מכונה דרך יישומים בבעיות ואתגרים ספציפיים למערכות סייבר (Systems Physical-Cyber) הסטודנטים ילמדו במהלך הקורס יישומים, שיטות ואלגוריתמים של למידת מכונה לתכנון מנגנונים חכמים למערכות סייבר (למשל: זיהוי אנומליות ו/או סביבות לא יציבות, שיפור יעילות של תהליכים, התאמה אישית של מערכת למשתמש). במהלך הקורס יודגשו האתגרים של ההתקדמויות הנוכחיות והעתידיות של יישומי בינה מלאכותית במערכות סייבר (למשל, רכבים אוטונומיים, עוזרים ווירטואליים, ומכשור רפואי חכם) בהיבטים של אתיקה, אבטחת מידע ופרטיות. הסטודנטים יתנסו בבעיות מחקריות ויישומים אמיתיים תוך דגש על פרויקט מסכם של הקורס (בשיטת למידה מבוססת פרויקט) ולפי מסלול ההתמחות.

מטרות הקורס:

הקורס מיועד להקנות לסטודנטים היכרות עם יישומים שונים של תחום הלמידות מכונה ועם השיטות והאלגוריתמים הנפוצים לעיבוד וניתוח נתונים במרחב הסייבר.

ביבליוגרפיה:

1. Russell S., Norvig P. (2018), Artificial Intelligence: A Modern Approach, Global Edition, Pearson Education Limited
2. Rawat D., Rodrigues K., Stojmenovic I. (2015), Cyber-Physical Systems: From Theory to Practice, 1st Edition, CRC Press

מסחר אלקטרוני E-Commerce

קוד הקורס: 90216

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם: אין דרישת קדם

סוג הקורס: שיעור, מעבדה, ולמידה מבוססת פרויקט PBL

מטרות הקורס:

מטרת הקורס הינה הקנייה של הבנה לגבי מושגים בתחום המסחר האלקטרוני ברשת האינטרנט. הצגה של פוטנציאל העסקי של שווקים אלקטרוניים. אופן המימוש של מסחר אלקטרוני. כלים ומתודולוגיות לביצוע מסחר. אספקטים אסטרטגיים של מוצרי מידע. השפעת האינטרנט על שרשרת הערך. יישומים טיפוסיים. מודלים מבוססים מסחר אלקטרוני כגון (B2C , B2B) טכניקות שיווק וסוגיות של אבטחת מידע.

נושאי הקורס:

- מבוא למסחר אלקטרוני , מושגים ומודלים עסקיים
- שיווק ופרסום באינטרנט
- קמעונאות ברשת
- תעשיית שירותים מקוונים
- ממשל זמין, למידה אלקטרונית, מסחר בין צרכנים
- מכירות פומביות, פורטלים וקהילות
- ספקי תוכן מקוונים
- מערכות תשלומים וסליקה

- ניהול קשרי לקוחות
- כדאיות השקעה במסחר אלקטרוני
- סוגיות אתיות במסחר אלקטרוני

ביבליוגרפיה:

Laudon, Kenneth C. and Carol Guercio Traver, E-Commerce: Business, Technology, Society 2014 (11th ed). New Jersey: Pearson Education (Prentice-Hall), ISBN-10: 0133507165; ISBN-13: 978-0133507164

מערכות תומכות החלטה Decision Support Systems

קוד הקורס: 90207

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם: כריית נתונים מתקדמת או כריית טקסטים או למידה לא מונחית מתקדמת

סוג הקורס: שיעור ולמידה מבוססת פרויקט PBL

מטרות הקורס:

הקורס מיועד להקנות לסטודנטים היכרות עם פיתוח ושימוש במערכות תומכות החלטה בתחומים שונים ודרך למידה מבוססת פרויקט בנושאים הספציפיים למסלול ההתמחות.

נושאי הקורס:

הקורס מספק ידע תיאורטי ומעשי על צרכי המידע בתהליכי קבלת החלטות של יחידים וארגונים. אפיון של סיטואציות החלטה מבחינת: רמת המבניות, אופן ההצגה, המידע הנדרש, ומודלים שיסייעו בקבלת החלטות. חברת עקרונות הסיווג והצגת העקרונות של מערכות מידע המסייעות בתהליכי קבלת החלטות. הקניית יכולת מתודולוגית להיעזר במערכות מידע ממוחשבות בסיטואציות החלטה. גישות ושיטות לעיצוב מערכות תומכות החלטה. הקורס דן ביחסי הגומלין בין סיטואציות החלטה והמידע הנדרש לבין קבלת החלטות. בנוסף, נלמדות גישות ומתודולוגיות לעיצוב מערכות מידע לתמיכה בקבלת החלטות. הסטודנטים יתנסו בבעיות מחקריות ויישומים אמיתיים תוך דגש על פרויקט מסכם של הקורס (בשיטת למידה מבוססת פרויקטים) ולפי מסלול ההתמחות.

Turban, E., Sharda, R., and Delen, D., Decision Support Systems and Intelligence Systems, Prentice Hall, 9th edition, 2010

מתמטיקה למדעי הנתונים Mathematics for Data Science

קוד הקורס: 90002

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם: אין דרישת קדם

סוג הקורס: שיעור, מעבדה, ולמידה מבוססת פרויקט PBL

מטרות הקורס:

החלק הראשון של הקורס מיועד להקנות ידע מוצק בנושאים מרכזיים של אלגברה לינארית בהם ישנו שימוש רחב במדע הנתונים. החלק השני יעסוק בתחומים של אנליזה נומרית לסטטיסטיקה.

נושאי הקורס:

- פרקים באלגברה לינארית
- טרנספורמציות לינאריות
- דמיון מטריצות, ערכים עצמיים ווקטורים עצמיים
- לבסון
- מרחב מכפלה פנימית
- היטלים אורתוגונליים
- פירוק SVD
- פירוק QR
- 1. טורי פורייה
- 2. התמרת פורייה
- 3. DFT

ביבליוגרפיה:

1. F.Su, **Mastering Linear Algebra**, THE GREAT COURSES, 2019.

2. M. Deisenroth, A. Faisal, C. Ong, **Mathematics for Machine Learning**, To be published by Cambridge University Press.
3. P.N. Klein, **Coding the Matrix**, Newtonian Press, 2013, <https://codingthetmatrix.com/>
- 4.G. Strang, **Linear Algebra and Learning from Data**, Wellesley-Cambridge Press, 2019

ניהול והנדסת תשתיות נתוני עתק בענן Management and Engineering of cloud data infrastructure

קוד הקורס: 90005

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם: אין דרישות קדם לקורס אבל נדרש ידע במערכות הפעלה ותכנות.

סוג הקורס: שיעור, מעבדה, ולמידה מבוססת פרויקט PBL

מטרות הקורס :

הקורס מיועד להקנות לסטודנטים היכרות עם התשתיות נתוני עתק בענן, ויכולות להגדיר צרכים ניהוליים והנדסיים של פרויקט מבוסס טכנולוגיות נתוני עתק בענן.

נושאי הקורס:

במהלך הקורס הסטודנטים ילמדו מהם העקרונות הניהוליים וההנדסיים של תשתיות נתוני עתק בענן. לאורך הקורס הסטודנטים ילמדו מה הן השבבות השונות של תשתיות נתוני עתק בענן וכיצד לבנות ארכיטקטורה של פרויקט שישתמש בטכנולוגיות הללו. במהלך הקורס יודגשו האתגרים של ההתקדמויות הנוכחיות והעתידיות של שימוש במאגרי נתוני עתק בענן בהיבטים של אתיקה, אבטחת מידע, פרטיות וכלכלה. הסטודנטים יתנסו בבעיות מחקריות ויישומים אמתיים תוך דגש על פרויקט מסכם של הקורס (בשיטת למידה מבוססת פרויקט) ולפי מסלול ההתמחות. במהלך הקורס יתבצע תרגול ושימוש בתשתיות ענן של AWS

הנושאים שיילמדו בקורס כוללים:

- הכרה עם אוסף המוצרים המסופקים על ידי AWS.
- שימוש בשירותי וירטואליזציה. הגדרה של שרתי ענן במערכות הפעלה שונות.
- שימוש בשירותי אחסון נתונים כדוגמת S3.
- העברת נתונים אוטומטית.
- אחסון של בסיסי נתונים רציונליים בענן.
- שימוש במערכת עיבוד מבוזר. הגדרת קלסטרים.

- שימוש ב-DOCKERS
- אבטחת מידע בענן.
- שימוש בכלים לכריית נתונים מבוססי ענן.

ביבליוגרפיה:

Maxim B., Heisel M., Ali N., Bahsoon R., Mistrik I., (2017) Software Architecture for Big Data and the Cloud, Morgan Kaufmann

ניתוח רשתות חברתיות Social Networks Analysis

קוד הקורס: 90213

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם: כריית נתונים מתקדמת או למידה לא מונחית מתקדמת

סוג הקורס: שיעור ולמידה מבוססת פרויקט PBL

מטרת הקורס:

הקורס מיועד להקנות לסטודנטים היכרות עם שימוש במדעי הנתונים בעולם רשתות החברתיות ולהבין את התופעות האופייניות לסביבות המחוברות.

נושאי הקורס:

במהלך הקורס הסטודנטים ילמדו את תורת הרשת המודרנית ואת יישומיה הרבים, החל מרשתות תקשורת ורשתות חברתיות במסגרת של אוטומציה של תהליכים, פעילות של מנועי חיפוש ואירגוניים שיווקיים, גילוי הריגות והונאה, אפידמיולוגיה ובאופן כללי תהליכים דינמיים ברשתות כמו התפשטויות מידע ודעות, התנהגויות ומחלות. הסטודנטים יתנסו בבעיות מחקריות ויישומים אמיתיים תוך דגש על פרויקט מסכם של הקורס (בשיטת למידה מבוססת פרויקטים) ולפי מסלול ההתמחות.

ביבליוגרפיה:

1. Luke, Douglas. 2015. A User's Guide to Network Analysis in R. Use R! Cham: Springer International Publishing.
2. Yang, Song, Franziska B. Keller, and Lu Zheng. 2016. Social Network Analysis: Methods and Examples. SAGE Publications.
3. M., Krishna Raj P., Ankith Mohan, and K. G. Srinivasa. 2018. Practical Social Network Analysis with Python. Springer.

סדרות עתיות ושיטות חיזוי

Time series and prediction methods

קוד הקורס: 90106

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם: שיטות הסתברותיות וסטטיסטיקות מתקדמות

סוג הקורס: שיעור

מטרות הקורס:

להקנות ידע בסיסי ומתקדם בסדרות עתיות. לפתח מיומנויות של ביצוע ניתוח סדרות עתיות בתחומים שונים של כלכלה וניהול ובביצוע מחקר, לפתח מיומנויות בשימושי מחשב בשפת תכנות פייתון.

נושאי הקורס:

- דוגמאות מתחומים שונים, גישות כלליות לניתוח נתוני עיתים, טיפול בטרנד ועונתיות
- תהליך סטוכסטי סטציונרי, פונקציית אוטוקורלציה, אמידה של פונקציית תוחלת
- תכונות של אוטוקורלציה, תהליכים ליניאריים
- הערכת של פונקציית אוטוקורלציה, נוסחת ברלט
- מודלים ARMA, פונקציית אוטוקורלציה חלקית
- מבוא לפילטרים ליניאריים ותורת ספקטראלית. צפיפות ספקטראלית של תהליך ARMA
- הערכת פרמטרים של תהליך ARMA
- תהליכים לא סטציונריים: SARIMA, ARIMA
- סדרות עתיות רב מימדיות, קרוס-ספקטרום
- ניתוח השפעות שינוי מדעניות בעזרת סדרות עתיות - Interrupted time series analysis

ביבליוגרפיה:

1. Peter Brockwell & Richard Davis (2002), "Introduction to Time Series and Forecasting", 2nd edition, Springer.
2. Robert H. Shumway & David S. Stoffer (2011), "Time Series Analysis and its Applications", 3rd edition, Springer.
3. Robert H. Shumway & David S. Stoffer (2015), "Time Series Analysis and its Applications - EZ", Free Texts in Statistics.

4. David McDowall (2019), "Interrupted time series analysis". Oxford University press.
5. Chris Chatfield (2004), "The Analysis of Time Series, an Introduction", 6th edition, Chapman & Hall/CRC.
6. Philip Hans Franses (2004), "Time Series Models for Business and Economic Forecasting", Cambridge University Press.
7. Aileen Nielsen (2019), Practical Time Series Analysis: Prediction with Statistics and Machine Learning, O'Reilly.

סמינריון

היקף הקורס : 2 ש"ס, 2 נ"ז

דרישות קדם: אין דרישת קדם

נושאי הקורס:

נושאי הסמינריון יקבעו בהתאם לעבודות הגמר שיוצגו על ידי הסטודנטים ויהיו משתנים משנה לשנה.

עיבוד אותות Signal Processing

קוד הקורס: 90104

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם: מתמטיקה למדעי הנתונים

סוג הקורס: שיעור

מטרות הקורס:

זהו מבוא מהיר לתחום עיבוד אותות ספרתי מוכוון למידה מבוססת פרויקט. הקורס מקנה ידע מעשי בעיבוד אותות ספרתי תוך כדי שימוש הספריות של MATLAB.

נושאי הקורס:

- אותות בזמן בדיד ובזמן רציף
- דגימה, משפט הדגימה, אליאסינג ותדר נייקוויסט
- מושג המערכת
- ניתוח במישור הזמן
- קונבולוציה
- מושג המערכת, התמרת, Z פונקציית תמסורת
- IDEF - IDEF
- ניתוח במישור התדר
- מושג המסנן
- שיטות סינון שונות

ביבליוגרפיה:

1. B.A. Shenoi, **Introduction to Digital Signal Processing and Filter Design**, John Wiley & Sons, 2006
2. G. Blanchet, M.Charbit, **Digital Signal and Image Processing using MATLAB**, ISTE 2006
3. E.Tarr, **Hack Audio: An Introduction to Computer Programming and Digital Signal Processing in MATLAB**, Taylor & Francis, 2019.

עיבוד אותות ביו-רפואיות Bio-Medical Signal Processing

קוד הקורס: 90211

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם: עיבוד תמונות

סוג הקורס: שיעור ולמידה מבוססת פרויקט PBL

מטרות הקורס :

הקורס מיועד להקנות לסטודנטים היכרות עם היסודות של עולם האותות הביו-רפואיות ועם השיטות והאלגוריתמים הנפוצים לעיבודיהם וניתוחיהם.

נושאי הקורס:

במהלך הקורס הסטודנטים ירחיבו את הידע בעיבוד אותות ביישומים ביו-רפואיים במגוון רחב של שיטות לעיבוד וניתוח של נתונים רפואיים למכשור שונה (בין היתר: מד טמפרטורה, מד לחץ, מד תאוצה, EEG, ECG, MEG, MCG, מד שמע, US, fNIRS) הסטודנטים ילמדו את השיטות העיקריות לעיבוד וניתוח אותות ביו-רפואיים. הסטודנטים יתנסו בבעיות מחקריות ויישומים אמיתיים תוך דגש על פרויקט מסכם של הקורס (בשיטת למידה מבוססת פרויקטים) ולפי מסלול ההתמחות.

ביבליוגרפיה:

Najarian K., Splinter R., (2012) Biomedical Signal and Image Processing, 2nd Edition,

CRC Press

2. Shortliffe E. H., Cimino J. J. (2014), Biomedical Informatics: Computer Applications in

Health Care and Biomedicine, Springer-Verlag

3. Venot A., Burgun A., Quantin C. (2014), Medical Informatics, e-Health: Fundamentals

and Applications, Springer-Verlag

עיבוד אותות על גרפים Graph Signal Processing

קוד הקורס: 90202

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם: עיבוד אותות

סוג הקורס: שיעור ולמידה מבוססת פרויקט PBL

מטרות הקורס:

הקורס הינו קורס חדשני בתחום המתפתח בקצב אדיר בשנים האחרונות. עיבוד אותות על גרפיים משלב טכניקות ורעיונות מתחומים שונים: תורת הגרפים, עיבוד אותות קלאסי, התורה הספקטרלית. היישומים הם בניתוח רשתות למיניהן, מדעי הנתונים, מדעי המוח ועוד. במהלך הקורס כל סטודנט יבצע פרויקט מעשי.

נושאי הקורס:

- גרפים: גרף מכוון/לא מכוון, ממושקל/לא ממושקל.
- מטריצת שכבות
- מטריצת לפלסיאן
- הספקטרום של לפלסיאן
- אותות על גרפים
- התמרת פורייה של אותות על גרפים מבוססת לפלסיאן
- אנליזה ספקטרלית מבוססת לפלסיאן
- שיטות סיבון
- אותות על גרפים תלויים בזמן
- יישומים של עיבוד אותות על גרפים

ביבליוגרפיה:

1. **L. Stankovic, M. Dakovic, E. Seidic**, Introduction to Graph Signal Processing, Springer, 2019.
2. **Ortega, P. Frossard, J. Kovacevic, J. M. F. Moura**, Graph Signal Processing: Overview, Challenges, and Applications, [Proceedings of the IEEE , Volume: 106, Issue: 5](#), 2018

עיבוד שפה טבעית Natural Language Processing

קוד הקורס: 90101

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם : למידת מכונה, למידה עמוקה (במקביל)

סוג הקורס: שיעור

מטרות הקורס:

תלמידי הקורס יכירו את המורכבות של המידול בשפות אנושיות, בתוספת ידע מעשי על פתרון משימות מסווג מסמכים לתרגום מכונה. בקורס זה נדון באלגוריתמים שונים של NLP נסביר ויישם אלגוריתמים בסיסיים וטכניקות בתחום. לדוגמה התלמידים ילמדו להבין גישות תחביר וסמנטיקה, הבנת גישות לשיח, יצירת דיאלוג וניהול במכונה, ולסיכום והבנת שיטות קיימות לגישות סטטיסטיות לתרגום מכונה. הבנת טכניקות הלמידה של מכונה בשימוש ב-NLP, כולל מודלי Markov מוסתרים ודקדוקים הסתברותיים ללא הקשר, אשכולות ושיטות ללא השגחה, מודלים ליניאריים ועוד. הסטודנט יידרש לרקע מתמטי מספיק כדי להיות מסוגל לקרוא מאמרים בלמידת מכונה אשר כוללים טיעונים לגבי סטטיסטיקה והסתברות בנוסף מיומנויות תכנות טובות כדי לפתח מודלים עבור הפרויקט הסופי בשפות המתאימות לכך, כגון R, Python וכו'

נושאי הקורס:

- מבוא ומושגי יסוד – סקירה
- טכניקות הערכה, ומידול שפה
- כריית טקסט – מדדים הסתברותיים להערכת דמיון ואשכולות
- תיוג, ניתוח ותחביר של טקסט
- תיוג סטוכסטי ומודלים לוג-ליניאריים
- מודלי חיזוי בעיבוד שפה טבעית
- אלגוריתם NER בעיבוד שפה טבעית
- עיבוד דיבור ומערכות דיאלוג
- עיבוד שיח: פילוח, החלטה אנאפורה
- תרגום מכונה
- יצירת שפה טבעית
- שיטות "לא מפוקחות" בעיבוד שפה טבעית

1. H. Lane , H. Hapke , C. Howard. Natural Language Processing in Action: Understanding, analyzing, and generating text with Python, 1st Edition, Manning Publications, 2019.
2. I. Dagan, D. Roth, M. Sammons and F. Massimo Zanzotto. Recognizing Textual Entailment: Models and Applications. Morgan & Claypool Publishers 2013.
3. D. Jurafsky and J. H. Martin, Speech and Language Processing: an introduction to natural language processing, computational linguistics, and speech recognition, 2 nd edition, Prentice Hall 200
4. C. D. Manning, and H. Schutze. Foundations of Statistical Natural Language Processing. Cambridge, MA: MIT 1999.
5. List of current articles.

עיבוד שפה טבעית מתקדמת NLP - Natural Language Processing

קוד הקורס: 90203

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם: מבוא לתכנות בשפת PYTHON (קורס ההשלמה), למידת מכונה

סוג הקורס: שיעור

מטרת הקורס:

עיבוד שפה טבעית הוא אחד מאבני הבניין החשובים בבינה מלאכותית. יישומי עיבוד שפה טבעית הם מגוונים ונוגעים לממשקי תקשורת רבים בין אנשים, לרבות תקשורת דוא"ל, פרסומים, שירות לקוחות, אתרי אינטרנט, בוטים, דוחות רפואיים, מערכות לזיהוי דיבור ודובר, ועוד. בשנים האחרונות החלו לשמש בתחום זה טכניקות מעולם הלמידה העמוקה, שהביאו לשיפור משמעותי בביצועים. מטרת הקורס היא להקנות לסטודנטים ידע מעמיק במתודולוגיות ובטכניקות המשמשות לעיבוד שפה טבעית. הסטודנטים בקורס זה ירכשו, הלכה למעשה, את הכלים לפתח ולממש מודלים לעיבוד שפה.

הנושאים שיילמדו בקורס:

- הקדמה לוקטורי מילים (word vectors)
- וקטורי מילים ומשמעויות - Glove
- סיווג מילים בחלונות, רשתות נוירונים.
- טכניקת BackPropagation וגרפים חישוביים.
- מבנה לשוני- ניתוח תלות (dependency parsing)
- ההסתברות של משפט: מודלי RNN ומודלי שפה.
- בעיית הגרדיאנטים הנעלמים ו - Fancy RNNs
- מכונת תרגום seq2seq, ומודלי Attention
- מענה על שאלות (answering question)
- רשתות קונבולוציה לעיבוד שפה טבעית.
- מיצוי מידע מחלקי מילה, models subword - ייצוג קונטקסט.
- טרנספורמרים ו attention self - למודלים יצרניים.
- גרף ומידע (Knowledge graph) והנדסת מידע (Knowledge engineering)

ביבליוגרפיה:

1. D. Jurafsky and J. H. Martin, "Speech and Language Processing" (3rd ed. draft), 2019.

available at <https://web.stanford.edu/~jurafsky/slp3/>.

2. J. Eisenstein, "Introduction to Natural Language Processing is", MIT Press, 2018, available at <https://jacobeisenstein.github.io/>.

3. Y. Goldberg, "A Primer on Neural Network Models for Natural Language Processing", Journal of Artificial Intelligence Research, vol. 57, pp. 345-420, 2015, available at <https://arxiv.org/pdf/1510.00726.pdf>.

עיבוד תמונות Image Processing

קוד הקורס: 90103

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם : עיבוד אותות

סוג הקורס: שיעור

מטרות הקורס:

הקורס מיועד להקנות לסטודנט כלים קלאסיים בעיבוד בתמונה לפתרון בעיות שונות. הקורס מלווה בביצוע פרויקט מעשי.

נושאי הקורס:

- תמונה במחשב. תמונה בגווי אפור ותמונת צבע
- עיבודים מרחביים
- מסכות וקונבולוציה
- חילוץ שפות
- סינון בעזרת פעולות מרחביות
- היסטוגרמה ושיפור ניגודיות
- התמרת פורייה בדידה דו-ממדית
- סינון בעזרת פעולות במישור התדר
- סגמנטציה
- עיבוד תמונה מורפולוגי
- דחיסת תמונה
- מבוא לעיבוד וידאו

ביבליוגרפיה:

1. R.C. Gonzalez & R.E. Woods, **Digital Image Processing Using MATLAB**, Gatesmark Publishing; 2nd edition (2009).
2. G. Blanchet, M.Charbit, **Digital Signal and Image Processing using MATLAB**, ISTE 2006

עיבוד תמונות ביו-רפואיות Bio-Medical Signal Processing

קוד הקורס: 90210

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם: עיבוד תמונות

סוג הקורס : שיעור ולמידה מבוססת פרויקט PBL

נושאי הקורס :

במהלך הקורס הסטודנטים ירחיבו את הידע בעיבוד תמונה ביישומי הדמיה ביו-רפואיים במגוון רחב של שיטות לעיבוד וניתוח של נתונים ספציפיים למכשור שונה (למשל צילומי רנטגן, טומוגרפיה ממוחשבת-CT – אולטרסאונד ווידאו MRI, fMRI) הסטודנטים ילמדו את השיטות העיקריות לעיבוד וניתוח תמונות רפואיות הכוללות פילוח, רגיסטרציה, תצוגה, שחזור אובייקטים תלת-ממדיים. הסטודנטים יתנסו בבעיות מחקריות ויישומים אמיתיים תוך דגש על פרויקט מסכם של הקורס (בשיטת למידה מבוססת פרויקטים) ולפי מסלול ההתמחות.

מטרות הקורס:

הקורס מיועד להקנות לסטודנטים היכרות עם היסודות של עולם ההדמיה הרפואית ועם השיטות והאלגוריתמים הנפוצים לעיבוד וניתוח תמונות ביו-רפואיות.

ביבליוגרפיה:

1. Najarian K., Splinter R., (2012) Biomedical Signal and Image Processing, 2nd Edition, CRC Press
2. Prince JL., Links J. (2014), Medical Imaging Signals and Systems Pearson 2nd edition
3. Shortliffe E. H., Cimino J. J. (2014), Biomedical Informatics: Computer Applications in Health Care and Biomedicine, Springer-Verlag
4. Venot A., Burgun A., Quantin C. (2014), Medical Informatics, e-Health: Fundamentals and Applications, Springer-Verlag

עקרונות למדעי הנתונים ובסיסי הנתונים

Data Science Fundamentals

קוד הקורס: 90001

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם: אין

סוג הקורס : שיעור

מטרות הקורס:

הצורך להתמודד עם כמויות גדולות של מידע הוליד בשנים האחרונות תפקידים רבים והתמחויות שונות כגון ה-DataAnalyst, ה-Big Data, וה-Business Intelligence. עם זאת, היכולת לשלב בין כל אלו ולהוסיף עליהם נדבך ייחודי של חיזוי, נותרה נחלתם של מעטים, ובשנים האחרונות ביסס עצמו ה-Scientist Data (מדען נתונים) כ"מקצוע הנחשק ביותר של המאה ה-21". מטרת הקורס היא לתת מענה לצורך הדוחק בתעשייה ובמחקר באנשי Data Science, להכשיר אנשים בעלי רקע רלוונטי לתחום, ולתת להם את השיטות והכלים הבסיסיים בעולם מדעי הנתונים.

נושאי הקורס:

- היכרות עם עולם הנתונים
- טיפול בסוגי נתונים שונים, ונתונים חסרים
- הסקה סטטיסטית בסיסית
- ויזואליזציות נתונים
- מבחנים סטטיסטיים
- עקרונות בסיסי הנתונים
- שימוש בבסיסי נתונים SQL
- NOSQL
- מבוא ללמידת מכונה

ביבליוגרפיה:

1. Python for Data Analysis: Data Wrangling with Pandas, NumPy and IPython, 2nd edition, Wes McKinsey, O'Reilly Media (2017)
2. The elements of statistical learning: data mining, inference and prediction. T. Hastie, R. Tibshirani and J. Friedman. Springer Verlag



מכון טכנולוגי חולון
Holon Institute of Technology

סילבוס מורחב לביצוע עבודת גמר לתואר שני במדעי הנתונים Final Assignment Project

נ"ד: 6 (קורס שנתי)

תוכן עניינים

37	<u>כללי:</u>
37	<u>הנחות יסוד:</u>
38	<u>שלבי הפרויקט ומטלות (סיום כל מטלה מותנה באישור המנחה):</u>
39	<u>אבני דרך (עדכונים ללוחות הזמנים באתר הקורס)</u>
41	<u>טיפים לעבודה נכונה:</u>
42	<u>נספח א' – מנחה</u>
42	<u>נספח ב': תחומים המתאימים לעבודת גמר:</u>
42	<u>נספח ג': כתיבת מסמכים בתוכנית לתואר שני במדעי הנתונים:</u>
42	<u>נספח ד': מבנה הצעת המחקר - מטלה 2</u>
45	<u>נספח ה': הנחיות לסקר -ספרות מטלה 3</u>
47	<u>נספח ו' - מבנה ותבנית לדו"ח עבודת גמר</u>
49	<u>נספח ז': דף כותרת להגשת עבודת גמר</u>
50	<u>נספח ח': רשימת תיוג</u>
52	<u>נספח ט' - בחינת הגמר</u>

כללי:

עבודת הגמר תהיה בתחומים הקשורים למדעי הנתונים תוך הדגשת הגישה המדעית והאנליטית. העבודה תחשב לבעלת ערך אם היא מקדמת את הידע וההבנה בתחום מדעי הנתונים. על העבודה להביא לידי ביטוי את החומר הנלמד, המתודולוגיות, והתכנים הנלמדים בתואר השני והיא חייבת להציג פן כמותי ולא רק איכותני. מטרת העבודה היא לפתח יכולות ומיומנויות לביצוע עבודת מחקר-פתוח במדעי הנתונים. העבודה צריכה להיות בדיקה לתחום ההתמחות של הסטודנט.

הנחות יסוד:

1. מעתה והלאה נתייחס לעבודה כאל פרויקט.
2. הפרויקט חייב להסתיים לאחר שנה ממועד התחלתו. לפרויקט הגמר יהיה לוח זמנים מפורט עם מטלות לאורך הדרך.
3. הפרויקט יבוצע בד"כ ע"י זוג סטודנטים. יש צורך בהבחנה ברורה בין התרומה האקדמית של כל סטודנט. במקרים מיוחדים ניתן לבצעו ע"י סטודנט יחיד.
4. בפרויקט גמר רצוי להשתמש בשיטות הנלמדות במסגרת הקורסים בתואר ובכלים לביצוע מחקר הנלמדים בקורס שיטות מחקר.
5. על פרויקט הגמר להביא לידי ביטוי את הנלמד בתואר השני ולכלול בנוסף לימוד עצמי יסודי של הנושאים הרלבנטיים לפרויקט.
6. יתכנו שני סוגי פרויקטים:
 - פרויקט יישומי להרחבת הידע בתחום מדעי הנתונים.
 - פרויקט מחקרי כמותי עם זיקה מעשית יישומית.
7. בפרויקט הגמר יושם דגש על הדרך. הוא חייב לכלול הצעת מחקר מסודרת, רמה לשונית גבוהה, סגנון וכללי כתיבה אחידים. יושם דגש על כתיבת חיבור ברמה גבוהה מבחינה לשונית ותחבירית.
8. הפרויקט יכלול סקר ספרות מקיף המבוסס על ספרות שפיטה. במסגרת סקר הספרות תוצג חשיבה ביקורתית, יכולת להבין וליישם מתודולוגיה נכונה.
9. חובה להגדיר קריטריונים לשיפוט והשוואה. יש לבסס את בחירת הקריטריונים על ספרות מקצועית או לוגיקה מוכחת. הקריטריונים יהיו כמותיים.

שלבי הפרויקט ומטלות (סיום כל מטלה מותנה באישור המנחה):



1. הצעה ראשונית – ההצעה (כעמוד) מבהירה את תחום העבודה. ההצעה תכלול את שם המחקר בעברית ובאנגלית, תיאור הבעיה שהובילה לעבודה, רציונל, ומטרת הפרויקט. נושא ההצעה ייקבע במשותף עם המנחה וחייב לקבל את אישור רכז עבודות הגמר. הצעות לתחומי מתאימים לעבודת הגמר. ההצעה תוגש כחודש לאחר תחילת הסמסטר. רכז עבודות הגמר רשאי לזמן את הסטודנט להצגת נושא הפרויקט לצורך החלטה סופית על קבלת הנושא והמנחה
2. כתיבת הצעת מחקר (עד 10 עמודים) הכוללת את המוטיבציה לפרויקט, המצב הקיים, המתודולוגיה המוצעת, המצב בספרות (לפחות 3 מקורות) ותוכנית המחקר בפרויקט. ניתן להתחיל בשלב זה ולהגיש את מטלה מס' 2 (ההצעה המורחבת) לאישור הנושא והמנחה במקום את מטלה מס' 1 (ההצעה הראשונית והמצומצמת)
3. כתיבת סקר ספרות מפורט של 15 פריטים לפחות ותיאור מפורט של המתודולוגיה לפי נושאים/שלבים. סקר הספרות יבצע אינטגרציה של כל פריטי הספרות שנקראו. הסקר יחולק לתת נושאים הקשורים לבעיית המחקר בפרויקט. ראה נספח ה'.
4. דו"ח התקדמות: הדו"ח מתאר מה בוצע בתוכנית המחקר, פיתוח אלגוריתמים, תיאור הניסויים, תוצאות וביצוע העבודה בפועל (כשני עמודים).
5. הגשת פרויקט הגמר
 - הגשת הפרויקט הסופי למנחה לאישור, לאחר אישורו הגשת 3 העתקים לשיפוט.
 - בחינת הגנה על הפרויקט בפני ועדת הבדיקה (מנחה + שני בוחנים)
 - תיקונים נדרשים והגשה סופית של הפרויקט לאחר אישור המנחה.

אבני דרך (עדכונים ללוחות הזמנים באתר הקורס)

הלוח מתאר את מועדי הסיום המאוחרים ביותר של כל אבן דרך. לוח הזמנים המוצג להלן הינו ביחס למועד התחלת הלימודים. לדוגמא אבן דרך ראשונה מדברת על מפגש היכרות בתחילת סמסטר אביב של השנה הראשונה בלימודי המוסמך. חובה לבצע את המטלות המודגשות בצהוב.

כל מטלה חייבת להיות מאושרת ע"י המנחה לפני הגשתה לרכז עבודות הגמר!

אבן דרך	תאריך	תיאור	מספר המטלה	הערות
א	בתחילת סמסטר שני בשנה א.	מפגש היכרות ראשון		מפגש מתוקשב.
ב	אבן דרך א 30+ יום	מפגש עם מנחים פוטנציאליים		פרטים יימסרו בהמשך
ב	אבן דרך א 60+ יום	הגשת הצעת מחקר ראשונית באישור מנחה. ראו פרק 1 בנספח ד'	1	חובת הגשה לרכז עבודות הגמר לא ניתן להמשיך ללא אישור המטלה ע"י רכז עבודות הגמר
ג	אבן דרך א 90+ יום	מפגש בנושא כתיבת סקר ספרות ומתודולוגיה		מפגש מתוקשב.
ד	אבן דרך א 100+ יום	הגשת הצעת מחקר מלאה כולל סקירת ספרות ראשונית ומתודולוגיה מוצעת. ראו נספח ד'	2	הגשה למנחה (אפשר בנוסף לרכז עבודות הגמר). אם מטלה מס' 1 לא הוגשה לאישור רכז עבודות הגמר, לא ניתן להמשיך ללא אישור מטלה זו ע"י רכז עבודות הגמר

ה	אבן דרך א	הגשת סקר ספרות ומתודולוגית הפרויקט. ראו נספח ה'	3	הגשה למנחה (אפשר בנוסף לרכז עבודות הגמר).
	120+ יום			
ו	תחילת שנה ב	הגשת דו"ח התקדמות	4	הגשה למנחה (אפשר בנוסף לרכז עבודות הגמר).
ז	סיום סמסטר 3 ללימודים	הגשת דו"ח עבודת הגמר (כולל תקציר מנהלים של 5-10 עמודים). ראו נספח ו'+ז' לגבי מבנה העבודה. בנוסף, הגשת מסמך תיוג – נספח ח'	5.1	חובת הגשה לרכז עבודות הגמר + הגשת חוברת (3 עותקים קשים) לרכזת תואר שני לצורך העברה לבוחנים
ח	במהלך סמסטר 4 ללימודים	בחינות הגמר. ראו נספח ט'	5.2	
ט	סיום סמסטר 4 ללימודים	הגשת דו"ח סופי, לאחר התיקונים	5.3	חובת הגשת עותק אלקטרוני ו-CD לרכזת תואר שני

טיפים לעבודה נכונה:

1. בטרם תיגשו לעבודה – בררו שאכן יש בידכם או באפשרותכם להשיג את הנתונים הנדרשים. אם אתם זקוקים גם לנתונים מארגון מסוים דאגו להבטיח אותם על ידי איסופם מוקדם ככל האפשר. (עבודה בה חסרים נתונים לא תתקבל)
2. בהרבה תחומים קיימים מאגרי מידע ציבוריים (למשל, הלשכה המרכזית לסטטיסטיקה בכל מדינה, ארגונים ממשלתיים וכדומה). בדקו מראש אם תוכלו להשיג את הנתונים או לרכשם. לגבי רכישה יש להגיש בקשה מסודרת למכון לצורך בדיקת האפשרות לממן זאת ע"י המכון
3. נספח ג' מתאר את הפורמט האחיד להגשת כל המטלות.
4. סקר ספרות חייב לתאר את כל המקורות לפי נושאים משותפים, לא לפי המחברים. לא יופיע ברשימת המקורות שלכם מקור אותו לא קראתם בעצמכם.
5. אסור להעתיק חלקים ממאמרים – יש לצטט בהתאם לנלמד בקורסים.
6. בדו"ח הסופי יש להקפיד על התבנית שתוארה לעיל ולכלול את כל מרכיבי החובה המתוארים בנספח ו'.
7. סדר הגודל של עבודה (לא כולל תקציר מנהלים ונספחים) כ-50 עמודים.
8. ייעוץ סטטיסטי – למעוניינים תינתן עזרה בהכנת שאלונים ובהרצת סטטיסטיקות (SPSS, R)
9. כל עבודות הגמר מופיעות בספרייה. עבודות גמר לדוגמה מופיעות באתר הקורס moodle
10. לאתר הקורס ב moodle יועלו כל החומרים הרלוונטים. נא הקפידו לעקוב ולהתעדכן.

עבודה פורה ומוצלחת!

נספח א' – מנחה

לכל סטודנט יהיה מנחה. על המנחה להיות חבר סגל במכון, בדרגת מרצה לפחות, או בעל תואר שלישי בתחום רלוונטי, חבר סגל באוניברסיטה או מכללה מוכרת, או מומחה מהתעשייה שאינו מקורב לסטודנט במקום עבודתו.

כאשר יש שני מנחים, נושא העבודה ייקבע במשותף. רכז עבודות הגמר רשאי למנות מנחה על הפרויקט אשר בו ניתן להיוועץ במהלך הכנת הפרויקט. על המנחה האקדמי לבדוק שהפרויקט עונה על הדרישות המקצועיות והאקדמיות לפני שהוא מאשר את הפרויקט.

רכז עבודות הגמר לתואר שני יפרסם רשימת מנחים ונושאים מעודכנת. דרישות והיקף הפרויקט ייקבעו על ידי המנחה.

נספח ב': תחומים המתאימים לעבודת גמר:

הנושאים לעבודת הגמר יהיו בהתאם להתמחויות השונות, ותחומי המחקר והעניין של המנחים.

נספח ג': כתיבת מסמכים בתוכנית לתואר שני במדעי הנתונים:

כל הדוחות יסודרו לפי התבנית הבאה:

- מרווח מימין ומשמאל - 25 מ"מ או 1 אינץ'
- מרווח מלמעלה 3 ס"מ ומלמטה 3 ס"מ
- גופן: עברית: Narkisim אנגלית: Times New Roman
- גודל הגופן : 12
- מרווח בין שורות 1.5
- כותרות ממוספרות בשיטה של 1.1.1 1.1.1 תוך שימוש בHeading 1, Heading 2, Heading 3
- גודל הגופן לכותרת ראשית 16 bold אחריה 14 bold ושלישית 12 bold אם צריך רביעית אז היא תהיה ב11 bold and Underline
- יש למספר את כל הטבלאות והתרשימים עם כותרת לכל טבלה ותרשים בנפרד. למשל "טבלה 3 ציוני הסטודנטים בקורס פרויקט הגמר"
- יש להכין תוכן עניינים, רשימת טבלאות ורשימת תרשימים באופן אוטומטי

נספח ד': מבנה הצעת המחקר - מטלה 2

כללי:

- היקף – עד 10 עמודים
- להקפיד על כותרות משנה וחלוקה לפסקאות
- רווח שורה וחצי בין השורות. פונט - גודל 12 (עברית: נרקיסים)
- להימנע מהצהרות שאין להן גיבוי מחקרי או ספרותי

פרק 1 – מבוא *

1.1 **רציונל** – רקע על הנושא הנחקר. מה מסקרן? מה התופעה שמעניינת? למה המחקר חשוב? מהי דילמת המחקר (מהי הבעיה שבגללה יש צורך במחקר? על מה צריך להחליט?) (חלופות של פתרונות אפשריים).

2.1 **מטרת העבודה** – ניסוח אפשרי: המחקר דן ב.../עוסק ב.../מטפל ב... מטרת המחקר היא... להבין את.../ללמוד על... / לבחון את ... / לאתר את.../לזהות את המאפיינים/הקשיים/הגורמים... יתכן שעל בסיס ממצאי המחקר יהיה ניתן להמליץ על ...

3.1 **שאלות המחקר** – שאלות המחקר נגזרות ממטרת/מטרות המחקר. (רצוי להימנע משאלות "האם") אפשר להשתמש ב: "באיזו מידה" (כמותי) ו"מה קורה" (איכותני).
4.1 חשיבות המחקר בהיבט התיאורטי והמעשי.

פרק 2 - סקר ספרות ראשוני

סקירת הידע הקיים בעולם בנושא המחקר; ממצאי מחקרים דומים. הסקירה גם באה להבטיח שהחוקר ייקח בחשבון את כל המשתנים הרלבנטיים.

1.2 מבוא לפרק - מה נכלל בסקירה ולמה? איך הסקירה רלבנטית למחקר? רצוי לחלק את הפרק ל 2-3 חלקים בהתאם למספר המרכיבים בשאלות המחקר. להקפיד על כותרות משנה.

2.2 סקר ספרות על נושא א'.

סקר הספרות איננו סיכום כרונולוגי של ספרים/מאמרים. הפרק צריך להיות כתוב במילים של הכותב. יש לקחת כל נושא ולפתח אותו תוך התבססות על ספרות מחקרית.

בתחילת כל פסקה/עמוד ו/או בסיום כל פסקה/עמוד יש לקשור את הסקירה עם מטרת/שאלת המחקר ולנמק מדוע ואיך מה שהובא רלבנטי למטרת/שאלת/נושא המחקר מותר ורצוי שסקר

הספרות יהיה ביקורתי. לא כל מה שבספרות הוא 'תורה למשה מסיני' להעביר את כל סקר הספרות דרך ה"מסננת" של מטרת/שאלת המחקר

2.3 סקר ספרות על נושא ב'....(ובך הלאה)

2.4 סיכום הפרק

פרק 3 – מתודולוגיה ראשונית

בפרק זה תתואר שיטת המחקר / העבודה.

אם המחקר מתבסס על אוכלוסיה מסוימת – יש לתארה ולציינה.

יש לתאר את מחקר הגישוש אם קיים כזה.

שיטת המחקר – מה תהיה שיטת המחקר (ניסוי, סקר, מחקר קורלטיבי וכד'?)

המדגם: תאור המדגם, גודלו, שיטת הדגימה.

כלי איסוף הנתונים – מדידות: היכן ואיך? שאלונים: כמה? איזה? מסמכים? סוג הראיונות (למשל, חצי

מובנים) וסוג התצפיות(למשל, הצופה המשתתף) עם נימוקים להצדקתם .

כלי מחקר נוספים (מה שרלוונטי)– סימולציות, מודלים, פיתוחים מתמטיים, תוכנות, כלי קבלת החלטות

ובחירת חלופות, אלגוריתמים וכד'.

בסקר – תכנון שלבי בניית השאלון (או השאלונים).

תוקף ומהימנות – כיצד יתוקפו כלי המחקר הכמותיים וממצאי המחקר האיכותיים? איך תיבדק מהימנות

כלי איסוף הנתונים? תוקף ומהימנות של כלי המחקר הנוספים.

ניתוח הנתונים - מה תהיה שיטת ניתוח ועיבוד הנתונים ?

תכנון לוחות זמנים ושלבי העבודה – הסבר ותרשים גאנט

רשימת מקורות References

את רשימת המקורות References יש לערוך לפי תקן. APA ראו למשל:

<http://owl.english.purdue.edu/owl/resource/560/01>

נספחים:

שימו לב – שפרק 1 הוא למעשה מטלה מס' 1 (הצעת מחקר מצומצמת).

נספח ה': הנחיות לסקר ספרות מטלה 3

סקר הספרות הוא תאור הרקע לפרויקט כפי שמשתקף בצורה אינטגרטיבית מכל הספרות שנקראה. את תוכן העבודה (המחקר) יש לפרק למספר נושאי משנה, יש לתאר כל נושא משנה בנפרד. בתיאור נושא המשנה יש להמציא את ההתייחסויות של המקורות השונים לאותו תחום. למשל: נניח כי בכוונתם לחקור את הלמידה באמצעות משחק עסקים. לפיכך עליכם לפרק את סקר הספרות לתחום משחקי העסקים, עקרונותיהם, יישומיהם ומרכיב הלמידה, למידה אקטיבית, יכולת משוב וכדומה. בכל מקרה, סקר הספרות איננו סיכום כרונולוגי של ספרים/מאמרים. יחד עם סקר הספרות יש להגיש מחדש את מתודולוגיית הפרוייקט, כפי שתוקנה והותאמה לפי סקר הספרות.

לדוגמא, בנושא משחקי עסקים - יש להגדיר את התחום ולתאר אסכולות:

The history of business games can be traced back to war games originating in China (Sun Tzu) nearly 2,500 years ago and the use of board games. However, the modern business simulation game dates back to 1950-1960 (Faria, 1998), when several business games were developed by the American Management Association (AMA), and when Greene and Andlinger's Business Management Game was developed for the consulting firm of McKinsey & Company.

The computer-based business simulation industry is now nearly 50 years old. Nevertheless, it has certainly undergone a remarkable transition. Currently, corporations, managers, educators, development consulting firms, and collegiate business programs all use simulations for training and teaching purposes (Faria, 1987, 1998).

Business simulations support three types of applications: educational programs, training firms and assessing business strategies. The present article will concentrate on the educational applications.

Business simulations are increasingly becoming an instructional methodology in business schools. Keys and Wolfe (1990) cited more than 50 empirical studies that provide evidence

about business games' problematic educational efficiency. In this article, we extended this line of research by seeking answers to the following interrelated questions:

1. Which success criteria should be used in business games?
2. What are the relationships between business game courses and other graduate level study courses?
3. What other parameters influence the successes and failures in business games?

In one of the first studies about the evaluation of business games, McKenney (1962) compared case studies of educational study programs that included using business games with study programs that did not. The comparison took place ...

A recent review (Baldissin et al., 2007) indicates that the market of educational business games offers a large number and a wide variety of management games. This includes 110 business game reviews with new technologies.

נספח ו' - מבנה ותבנית לדו"ח עבודת גמר

ספר הפרויקט צריך לכלול פרקי מבוא המשקפים את הגדרת הבעיה וניתוחה, המצב הקיים בספרות המקצועית ובתעשייה; ניתוח אלטרנטיבות; פתרונות אפשריים. חלקו העיקרי של הספר יכלול פרקים ייחודיים המציגים פיתוחים מתמטיים, עבודות סימולציה ותוצאותיהן, תוצאות ניסיוניות ואופן השגתן, דיון וניתוח מקיף של התוצאות ומסקנות. העבודה חייבת להתבסס על מחקר ביבליוגרפי מקיף. הפרקים הייחודיים בפרויקט הגמר מיועדים לבחון את יכולתו של הסטודנט לעבודה עצמית בפתרון מקורי של בעיות ניהוליות טכנולוגיות. הסטודנט יצטרך לעבוד עם ספרות מקצועית וללמוד בעצמו את החומר שלא נלמד במסגרת הקורסים שבתכנית הלימודים לקראת התארים הראשון והשני.

תבנית של עבודת הגמר:

דף כותרת: ראה נספח ז'

תודות (לא הכרחי)

תקציר מנהלים (5-10 עמודים)

תוכן עניינים (אוטומטי) - חובה

רשימת טבלאות (אם יש טבלאות)

רשימת תרשימים (אם יש תרשימים)

רשימת קיצורים ומילון מושגים (אם יש ספציפי לעבודה)

פרק 1: מבוא - תיאור מפורט של הבעיה שהובילה לעבודה, הרציונל לפרויקט, מטרות הפרויקט, (שאלות והשערות המחקר היכן שרלוונטי), חשיבות העבודה בהיבט התיאורטי והפרקטי.

פרק 2: סקר ספרות ביקורתי - מטרת הסקר הינה להציג את מצב הידע הקיים בעולם בתחום הנבדק תוך התבססות על מקורות עדכניים וממצאי מחקרים. בסקר יש לתאר את מרכיבי המחקר ולכל מרכיב את השיטות בהן נוקטים בספרות המקצועית. הסקר יכלול רק נושאים ייחודיים למחקר. נושאים שיש לסקור אותם, על פי הנחיית המנחה, ואשר מופיעים בספרי הלימוד או בתכני הקורסים, יכללו בדוח הפרויקט בתת פרק נפרד. איכות הפרויקט תיבחן גם על ידי רמה ואיכות של המאמרים הנסקרים. חשוב שהסקר יהווה אינטגרציה של כל הספרות הרלוונטית ולעולם לא יכלול רשימת מקורות ומה מתואר בכל מקור.

פרק 3: מתודולוגיה – שיטת העבודה לרבות שלבי העבודה, מערך המחקר, אוכלוסיית המחקר ואופן בחירת

מדגם הנבדקים (אם רלוונטי), שיטות וכלי איסוף הנתונים ונימוקים לגבי תקפות ומהימנות הכלים (אם רלוונטי), אלגוריתמים לפיתוח, (אם רלוונטי).

פרק 4: תוצאות - פרק המכיל את הממצאים מתוך העבודה והרלבנטיות שלהם.

פרק 5: סיכום ודיון - הפרק יכלול ביו היתר התייחסות ישירה של ממצאי הפרויקט לסקירת הספרות,

להשערות שהיו בבסיס הפרויקט ולמטרות הפרויקט, המלצות ומגבלות הפרוייקט.

רשימת מקורות- רשימת מקורות תיעשה על פי כללי APA יש להקפיד שלכל פריט ברשימת המקורות יש

אזכור בתוך הפרויקט, ולכל אזכור בתוך גוף הפרויקט יש אזכור ברשימת המקורות.

נספחים: נספחים ממוספרים באותיות א-ב.

מכון טכנולוגי חולון
תוכנית מוסמך במדעי הנתונים

"שם העבודה"

עבודת גמר לשם מילוי חלקי של הדרישות לקבלת תואר מוסמך במדעים

מאת:

שם עם מספר זיהוי

בהנחיית: *שם המנחה*

נספח ח': רשימת תיוג

רשימת תיוג של סיום עבודת הגמר בתואר שני – מדעי הנתונים

לכבוד

ראש החוג לתואר שני במדעי הנתונים

המכון הטכנולוגי בחולון

הנדון: אישור על השלמת דרישות עבודת הגמר

במסמך זה הנני חותם ומאשר כי ביצעתי את עבודת הגמר כראוי וכמתבקש בדרישות הסילבוס של הקורס "עבודת הגמר".

להלן חתימתי על כל אחד מהמרכיבים הנדרשים:

<u>הערות והבהרות</u>	<u>סמן ב V</u>	<u>הנושא</u>
		הגירסה המוגשת היא הסופית ולאחר אישור המנחה
		דף השער תואם את הדוגמא בנספח ו' בסילבוס
		המסמך כתוב בגופן ורווחים לפי נספח ג' בסילבוס
		תקציר מנהלים קיים ב 5-10 עמודים
		תוכן עניינים קיים ומתעדכן אוטומטית
		רשימת טבלאות קיימת ומתעדכנת אוטומטית
		רשימת איורים קיימת ומתעדכנת אוטומטית
		פרק מבוא - קיים
		פרק סקר ספרות - קיים
		הסקר רשום בצורה אינטגרטיבית מכוונת נושא ולא סכום מאמרים
		הציטוט לפי חוקי APA (כמוצג בסילבוס הקורס)
		רישום המקורות לפי אב... ולפי חוקי APA
		פרק בנושא שיטת המחקר (מתודולוגיה) – קיים (לא רלוונטי לעבודה בתכן הנדסי)
		פרק תוצאות – קיים (לא רלוונטי לעבודה בתכן הנדסי)

		פרק סיכום ודיון מבוססות על ממצאי העבודה והספרות - קיים
		כל הטבלאות ממוספרות + כותרות
		כל האיורים ממוספרים + כותרות
		איור או טבלה המועתקים ממקור אחר – מצויין המקור בתחתית הטבלה/איור
		כל עמודי העבודה ממוספרים באמצע החלק התחתון
		אין העתקה והדבקה בעבודה המוגשת למעט פסקה לכל היותר עם הכללתה במרכאות וציון המקור
		בעבודה שנעשתה ע"י זוג סטודנטים מוצגת חלוקת העבודה בניהם

ההערות והבהרות מיועד למקרים שאינך עומד במרכיב זה בחלקו או במלואו. הסבר את הסיבה לכך.

לסיום, הנני מאשר בשמי ובחתימתי כי עבודה הגמר המוגשת הינה פרי עבודתי בלבד .

שם הסטודנט _____ ת.ז. _____ .

חתימה _____ תאריך _____ .

נספח ט' - בחינת הגמר

רכז עבודות הגמר ימנה וועדת בדיקה בת שלושה בוחנים או יותר (אחד מהם הוא המנחה) לכל עבודת גמר. לאחר שוועדת הבדיקה קראה את דו"ח עבודת הגמר והדו"ח אושר על ידה, יקבע מועד לבחינת גמר בעל-פה. בנוסף לבוחנים, ייתכן ויוזמנו לבחינה חברי סגל נוספים.

הרכב הבחינה:

- 20 דקות: הצגת העבודה ע"י הסטודנט.
- 20 דקות: שאלות מועדת הבדיקה
- 20 דקות: דיון ועדת הבדיקה (לא בנוכחות הסטודנט) לשם סיכום חוות דעת מוסכמת.

במהלך ההצגה והשאלות הסטודנט יערוך רישום של כל ההערות שניתנו ע"י הבוחנים לתיקון. במידה ולא יגיעו הבוחנים לחוות דעת מוסכמת יעבירו את חוות דעתם לראש החוג להמשך טיפול. במקרה זה יכריע ראש החוג לאחר שיתיעץ עם הבוחנים וחברי סגל אקדמי נוספים. ועדת הבדיקה רשאית לדרוש מן הסטודנט להכניס תיקונים בדו"ח אחר בחינת הגמר. הסטודנט ידרש להגישם תוך חודש ממועד הבחינה. הדו"ח המתוקן יועבר למנחה אשר יבדוק באם תוקנו כל הערות הבוחנים. במידה וכן הוא יעביר את הסכמתו בכתב לרכז עבודות הגמר. במידה והסטודנט נדרש להגיש דו"ח חדש, יעביר המנחה עותק לבוחנים אשר יתבקשו שוב להגיש לראש התוכנית חוות דעת מנומקת בכתב. רק לאחר ביצוע כל התיקונים יעביר המנחה האחראי לרכז עבודות הגמר את חוות הדעת המסכמת של הבוחנים. בהתאם לחוות דעת זו יחליט ראש החוג ללימודי תואר שני באם לקבל את הדו"ח. כנשל הסטודנט בבחינה, יורשה להיבחן פעם נוספת לא יאוחר מאשר שנה אחרי הבחינה הראשונה. כשלוש שני יביא להפסקת הלימודים.

רכב אוטונומי Autonomous Vehicle

קוד הקורס: 90204

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם : אין דרישת קדם

סוג הקורס: שיעור

מטרות הקורס:

פיתוח רכב אוטונומי הוא תחום שנמצא היום בקדמת הטכנולוגיות המשלבות חומרה ותוכנה. פיתוח רכב אוטונומי משתמש באופן מסיבי בשיטות של מדעי הנתונים, בראש וראשונה - בלמידה עמוקה. הקורס מהווה הכרה בשיטות מתקדמות לפיתוח מערכות נהיגה אוטונומי.

נושאי הקורס:

1. מבוא למושג של נהיגה אוטונומי. רמות שונות של הנהיגה
2. סקירת מערכות GNSS (מערכות ניווט לוויינית גלובלית)
 - שגיאות ודיוק GNSS
 - איתור בזמן אמת
 - היתוך מידע סנסורי
3. חישה לנהיגה אוטונומית
 - איסוף מידע ממקורות שונים
 - זיהוי
 - סגמנטציה
 - זרימת נתונים
4. למידה עמוקה בנהיגה אוטונומית
 - Convolutional Neural Networks
 - למידה עם חיזוק
 - שימוש בלמידה עמוקה לזיהוי, סגמנטציה וניתוח זירה
5. תכנון, בקרה וקבלת החלטות
 - תהליך החלטה מרקובי
 - תכנון תנועה
 - בקרה עם משוב

1. S. Liu, L. Li, J. Tang, S. Wu, J. Gaudiot, **Creating Autonomous Vehicle Systems**, Morgan & Claypool ,2018.
2. H. Yu, X. Li, R. Murray, S. Ramesh, C. Tomlin, **Safe, Autonomous and Intelligent Vehicles**, Springer 2019.

שיטות בייסיאניות בלמידה

Bayesian Methods in Machine Learning

קוד הקורס: 90205

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם: למידת מכונה; שיטות הסתברותיות וסטטיסטיקות מתקדמות

סוג הקורס: שיעור

מטרות הקורס:

תחום המחקר המתהווה של למידה עמוקה בייסיאנית משלב את היתרונות של שיטות למידה עמוקה מודרניות עם היתרונות של שיטות סטטיסטיות בייסיאניות מודרניות כדי להעריך את ההסתברויות ולקבל החלטות תחת אי ודאות. מטרת הקורס היא להביא את הסטודנטים לחזית הידע בתחום זה. במהלך הקורס ייחשפו הסטודנטים לגישות החדשות בתחום המידול (למשל מודלים לפונקציות ומודלים גנרטיביים עמוקים), פרדיגמות למידה (למשל MCMC ו- variational inference) ופלטפורמות תכנות הסתברותיות (למשל TensorFlow, PyTorch, PyMC)

נושאי הקורס:

- מבוא לגישה בייסיאנית, אמידה בייסיאנית, קבלת החלטות בייסיאנית
- הסקה בבעיות הקירוב: Laplace approximation, variational Bayes, expectation propagation
- דגימה, שיטות דחייה וקבלה, MCMC
- מודלים פרמטרים: רגרסיה ליניארית בייסיאנית, רגרסיה לוגיסטית, מודל דריכלה
- מודלים אי פרמטרים: תהליכי גאוסיינים
- רשתות עצביות בייסיאניות
- מודלי מונטה קרלו לעבור רשתות עצביות בייסיאניות
- מידול של אי ודאות

ביבליוגרפיה:

1. Gaussian Processes for Machine Learning. Rasmussen and Williams, MIT Press, 2006.
2. Ian Goodfellow and Yoshua Bengio and Aaron Courville (2016) Deep Learning Book.
3. Bayesian Methods for Data Analysis (2008), B.P. Carlin, CRC press.
4. Statistical Decision Theory and Bayesian Analysis by James Berger.
5. Bayesian Reasoning and Machine Learning by David Barber.

שיטות הסתברותיות וסטטיסטיות מתקדמות למדעי הנתונים Probabilistic methods and statistics for data science

קוד הקורס: 90003

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם: מבוא להסתברות וסטטיסטיקה

מטרות הקורס:

שיטות הסתברותיות וסטטיסטיות הן חלק חיוני בתחום של מדעי הנתונים. הקורס מיועד להקנות לסטודנט ידע בסיסי בסימולציה, שרשראות מרקוב ובהסקה סטטיסטית ולשמש בסיס לקורסי המשך, כגון: למידת מכונה, למידה עם חיזוק, שיטות בייסיאניות ועוד.

נושאי הקורס:

מבוא לתהליכים סטוכסטיים:

- תוחלת מותנה
- שרשראות מרקוב בזמן בדיד
- הסתברויות גבוליות וסטציונריות
- Text analysis, monte carlo, pagerank

מבוא לסימולציה:

- עקרונות סימולציה
- סימולציה של מספרים אקראיים
- סימולציה של תהליכים אקראיים
- אמידה של פרמטרים של תהליכים אקראיים בעזרת סימולציה

הסקה סטטיסטית:

- התפלגויות רב מימדיות
- מטריצת קווריאנס
- התפלגות רב נורמלית
- רגרסיה לינארית רב משתנית
- רגרסיה לא ליניארית

ביבליוגרפיה:

1. תלמה לויתן ואלונה רביב, **מבוא להסתברות וסטטיסטיקה**, עמיחי, 2005
2. M. Mitzenmacher, Probability and Computing: Randomization and Probabilistic Techniques in Algorithms and Data Analysis, 2nd edition.
3. N. Matloff, **Probability and Statistics for Data Science**, CRC Press Taylor & Francis Group, 2020
4. Daniel J. Denis, **Univariate, Bivariate, and Multivariate Statistics Using R (Quantitative Tools for Data Analysis and Data Science)**, John Wiley, 2020
5. S. Ross, Simulation , 2012
6. J.V. Guttag, Introduction to Computation and Programming Using Python: With Application to Understanding Data 2016
- R. Dobrow, Introduction to Stochastic Processes with R. 2016

שיטות מתקדמות בלמידת מכונה Machine Learning Advanced Supervised

קוד הקורס: 90201

היקף הקורס: 3 ש"ס, 3 נ"ז

דרישות קדם: למידת מכונה, סטטיסטיקה

מטרות הקורס:

זהו קורס מתקדם המציג את היסודות התיאורטיים של למידת מכונות מודרנית, כמו גם שיטות ומסגרות מתקדמות המשמשות בלימוד מכונות מודרני. הקורס מניח כי הסטודנטים למדו קורסים מבוא לתואר שני בלימוד מכונה וכן סטטיסטיקה. הקורס מתייחס הן לתכנון אלגוריתמי למידה טובים, כמו גם לניתוח תכונות חישוביות וסטטיסטיות של האלגוריתם וזמני ביצוע. האלגוריתמים מוצגים יחד עם היבטים מעשיים של מתודולוגיה ואינטואיציה כדי לסייע לתלמידים לפתח כלים לבחירת שיטות וגישות מתאימות לבעיות במחקר שלהם. נעסוק בשיטות למידת מכונות מתקדמות כגון גישות לא-פרמטריות ועמוקות להלחנה מכוונת לאומדן צפיפות ורגרסיה; תיאוריה מתקדמת כמו יסודות של אשכולות, סיווג, הגברה; תיאוריה ושיטות בצומת יעילות סטטיסטית וחישובית, כמו גם תצלומים של תוצאות תיאורטיות בכמה נושאים חמים כמו איתנות והסבר.

נושאי הקורס:

- Nonparametric Prediction
- High-dimensional Prediction
- Nonparametric Density Estimation
- Deep Density Estimation
- Computational vs Statistical Efficiency Tradeoffs, Optimization and Estimation
- Prediction and Games

ביבליוגרפיה:

1. Kevin P. Murphy. Machine Learning: A Probabilistic Perspective. MIT Press 2012
2. Shalev-Shwartz, Shai, and Shai Ben-David. Understanding Machine Learning: From Theory to Algorithms. Cambridge University Press, 2014. ISBN: 9781107057135