

סילבוסים למדעי הנתונים תשפ"ד

תואר שני (M.Sc.) במדעי הנתונים

תוכן עניינים

3.....	תקצירי קורסי חובה
7.....	תקצירי קורסי ליבה
11.....	תקצירי קורסי בחירה

תקצירי קורסי חובה – תואר שני M.Sc.

במדעי הנתונים

99218 – למידה סטטיסטית Introduction to Statistical Learning

אופן הוראה: שיעור
שעות שבועיות: 3
נקודות זכות: 3
דרישות קדם: אין

מטרת הקורס

הקורס היא להקנות לסטודנטים ידע במתודולוגיות ושיטות של למידה סטטיסטית. הקורס יתחיל מסקירה של השיטות הקלאסיות, לאחר מכן, יוצגו שיטות רגולריזציה ושימושיהן בבעיות רגרסיה וסיווג. בנוסף לעיל, הקורס יציג נושאים כגון ניתוח אשכולות, עצי החלטה ובוטסטראפ. אם הזמן יאפשר, ילמדו נושאים הכוללים הסקה סיבתית והחלקות גרעין. המימוש של השיטות יעשה בשפת R. בסוף הקורס הסטודנטים יוכלו לממש את השיטות השונות על נתוני אמת, לנתח את התוצאות ולקרוא ספרות מחקרית מתקדמת בהתאם לצורך.

נושאים

סקירה של השיטות הקלאסיות, שיטות מודרניות בלמידה סטטיסטית: שיטות כיווץ (shrinkage), שיטות סיווג, SVM, ניתוח אשכולות, מטריצות קרבה, שיטות מבוססות עצים, עצי רגרסיה, שיטות בוטסטראפ, מבוא להסקה סיבתית, שיטות החלקת גרעין.

99002 – מתמטיקה למדעי הנתונים Mathematics for Data Science

אופן הוראה: שיעור
שעות שבועיות: 3
נקודות זכות: 3
דרישות קדם: אין

מטרת 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.

נושאים

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

שיטות סטטיסטיות – 99217 Statistical Methods

אופן הוראה: שיעור
שעות שבועיות: 3
נקודות זכות: 3
דרישות קדם: אין

מטרת הקורס

סטטיסטיקה היא הבסיס לניתוח, הסקה ובנית תחזיות מנתונים. בתחילת הקורס נחזור על מושגים בסיסיים בסטטיסטיקה כגון בדיקת השערות אמידה רווחי סמך התפלגות נורמלית ורב נורמלית מטריצת שונויות. המטרה העיקרית היא להקנות הבנה וכלים לנושא של רגרסיה ליניארית. רגרסיה היא מודל סטטיסטי המתאר קשר פונקציונלי בין משתנים. רגרסיה ליניארית מניחים שהמשתנה אותו מעוניינים ללמוד תלוי ליניארית במספר משתנים מסבירים, כאשר מקדמי המשתנים המסבירים אינם ידועים. יינתנו וינתחו דוגמאות מנתונים אמיתיים תוך שימוש בחבילות הסטטיסטיות של "פייתון"

נושאים

מציאת אמדים הטוב ביותר לקו הרגרסיה, תחזיות, חקירת ההתאמה של המודל לנתונים, מיקרים של רגרסיה לא ליניארית, בדיקת סטייה מהמודל, בחירת המשתנים המסבירים הטובים ביותר, פתרונות למקרים שבהם הנחות המודל לא מתקיימות.

99004 – למידת מכונה Machine Learning

אופן הוראה: שיעור
שעות שבועיות: 3
נקודות זכות: 3
דרישות קדם: עקרונות למדעי הנתונים ובסיסי נתונים

מטרת הקורס

למידת מכונה (או לימוד מכונה) היא מתודולוגיה העוסקת בפיתוח אלגוריתמים ומודלים לניתוח ועיבוד אוטומטיים של נתונים, החל מנתונים מובנים (structured data) כגון נתוני מסחר בכורסה, בדיקות דם ובדיקות רפואיות אחרות, נתונים פיננסיים, ועוד, וכן נתונים לא מובנים (unstructured data), כגון וידאו, תמונות ואותות דיבור. תחום זה חווה בשנים האחרונות התפתחות טכנולוגית משמעותית, בין היתר בעקבות הצטברות בסיסי נתונים גדולים ופיתוח של אלגוריתמים מתקדמים לעיבוד הנתונים, לרבות איסוף הנתונים, עיבוד מקדים וניקוי, ניתוח, מידול, הערכת ביצועים וויזואליזציה. בקורס זה נכיר את היסודות של למידת מכונה ותינתן סקירה מקיפה של מושגים, טכניקות ואלגוריתמים רבים בלימוד מכונה, החל מנושאים כמו סיווג ורגרסיה ליניארית, K-Nearest Neighbors, Support Vector Machine, עצי החלטה ורשתות נוירונים.

נושאים

הקדמה כללית, שימוש בכלי פייתון לעיבוד נתונים, אקספלורציה של בסיסי נתונים (data mining), עיבוד מקדים: נתונים חסרים, נרמול, התמודדות עם חוסר איזון בנתונים (SMOTE, upsampling, downsampling), ייצוג נתונים קטגוריאליים - embedding, one-hot, סווג K-Nearest Neighbors, Over-fitting, עצי החלטה, Random Forest, Gradient Boosting, רגרסיה ליניארית, Support Vector Machines, רגרסיה לא ליניארית, Logistic Regression, רגולריזציה, אופטימיזציה - עקרונות בסיסיים, פונקציות מחיר ושיטות אופטימיזציה, מדדי ביצוע ושיטות להערכת ביצועים - מטריצת מבוכה, recall ו-precision, וכו', K-Fold, Re-substitution, מידול בציר הזמן, שיטות להתמודדות עם בעיות של חוסר סטציונריות בנתונים (trend removal, differential features), ארכיטקטורות מתקדמות (LSTM, RNN), הקדמה לרשתות נוירונים. Unsupervised Learning, Kmeans, t-SNE, DBScan, Principle Component Analysis באמצעות

99006 – למידה עמוקה Deep Learning

אופן הוראה: שיעור

שעות שבועיות: 3

נקודות זכות: 3

דרישות קדם: מבוא ללמידת מכונה, סטטיסטיקה והסתברות, פיתון

מטרת הקורס

קורס זה עוסק בטכניקות העדכניות ביותר בלמידה עמוקה ולמידת ייצוג, תוך התמקדות בלמידה עמוקה מפוקחת ובלתי מפוקחת, רשתות קונבולוציה (CNN) והיזון חוזר (RNN), רשתות Transformers, ומודלים גנרטיבים (Autoencoders, GANS) עם יישומים לראייה ממוחשבת, הבנת שפה טבעית ורפואה.

נושאים

מבני רשתות עצביות, אלגוריתם ה-backpropagation, אלגוריתם stochastic gradient descent ותכונותיו, פונקציות מחיר ושיטות אופטימיזציה מסדר ראשון, למידה עמוקה עבור בעיות סיווג ורגרסיה, מעבר מסיווג לזיהוי עצמים, מודלים יוצרים – variational autoencoder, autoencoder.

99005 – ניהול והנדסת תשתיות נתוני עתק בענן Management and Engineering of cloud data infrastructure

אופן הוראה: שיעור

שעות שבועיות: 3

נקודות זכות: 3

דרישות קדם: אין

מטרת הקורס

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

נושאים

זהו קורס מבוא לעולם התכנות, המלמד את עקרונות שפת התכנות Python. הנושאים שיידונו בקורס הם: הצורך בתכנות (מושגי יסוד: בעיה חישובית, אלגוריתם, תכנית), תכנות בסיסי בשפת Python, בקרת זרימה, לולאות, פונקציות, מחלקות ואובייקטים, רקורסיה, ניתוח סיבוכיות, מיון וחיפוש ומבוא למבני נתונים.

99305 – סמינריון מחקר Research Seminar

אופן הוראה: שיעור

שעות שבועיות: 2

נקודות זכות: 2

דרישות קדם: סטטיסטיקה, למידת מכונה

מטרת הקורס

כחלק מתהליך המחקר, בקהילה המדעית, יש חשיבות לתת לסטודנט כלים מחקרניים. בקורס זה נכיר את הכלים צריכים לאפשר להגדיר את שאלת המחקר, לעשות סקירת ספרות, לדעת לאסוף ידע שתומך בשאלת המחקר, ולנתח את התוצאות במתודולוגיות סטטיסטיות המתאימות ביותר, ולדעת להתייחס לתוצאות בצורה ביקורתית. סמינריון מחקר זה, נועד לתת לסטודנט את הכלים האלו, כולל כלים חדשים של מציאת בבילוגרפיה מתאימה, שימוש בכתיבה באמצעות בינה מלאכותית בצורה ביקורתית. בהרצאות מבוא של הקורס עוברים על מאמרים עדכניים ובוחרים בצורה ביקורתית את טיב הכתיבה, סקירת הספרות, הערכה לטיב השיטות שכותבי המחקר השתמשו, ובחינה האם המסקנות של המאמר נתמכות בתוצאות. רוב הדוגמאות בכתה יהיו מתחום הרפואי, בתחום של עיבוד מידע רפואי, וקבלת החלטות רפואיות, אך לא מוגבלות אליו. במסגרת הקורס יתנסו התלמידים עם כתיבת סקירת ספרות, ובהצגת ניתוח של מאמר בצורה ביקורתית.

נושאים

1. הקדמה כללית להערכה של עבודה מדעית וגיבוש שאלת המחקר. מתן דוגמא מבניית מערכת של ממשק אדם-מכונה באמצעות חישה מתקדמת [3-1]
2. סקירת ספרות – המטרה של הסקר, מבנה הסקר, חיפוש, שימוש בכלי ענן לכתיבה בבילוגרפיה, וכלים מבוססים בינה מלאכותית
3. קריאה ביקורתית של מאמרים: בחינה של איכות סקר ספרות, התאמת שיטות המחקר לדתה, ויזואליזציה, התאמת תוצאות המאמר למסקנות
4. שיטות להצגה והנגשת המחקר
5. שיטות עיבוד מידע לחישה וסיווג של תנועה [9-4]
6. שיטות עיבוד מידע לכריית מידע מהבעות פנים [10,11]
7. שיטות לחיזוי של מצב רפואי עבור סיווג של מחלות ומצבי קיצון [12-14]
8. שימוש בבינה מלאכותית עם ידע מומחה באבחון רפואי ובלמידת חיוויים (RL)

99300 – עבודת גמר - פרויקט Final Assignment Project

אופן הוראה: שיעור

שעות שבועיות: קורס שנתי, 6

נקודות זכות: 6

דרישות קדם: אין

מטרת הקורס

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 בזיקה לתחום ההתמחות של הסטודנט.

תקצירי קורסי ליבה – תואר שני M.Sc. במדעי הנתונים

99101 – עיבוד שפה טבעית Telemedicine

אופן הוראה: שיעור

שעות שבועיות: 3

נקודות זכות: 3

דרישות קדם: למידת מכונה, למידה עמוקה (במקביל)

מטרת הקורס

תלמידי הקורס יכירו את המורכבות של המידול ועיבוד שפות אנושיות, (דיבור וטקסט) בתוספת ידע מעשי על פתרון משימות שונות בתחום כגון מערכות צ'ט ושיח, עיבוד טקסט ודיבור, תרגום מכונה, זיהוי רגשות בשפה וכד. בקורס נדון באספקטים שונים של NLP נסביר גישות וטכניקות בתחום ובעיקר נשוחח על האתגרים ומדוע זו נחשבת למשימה הקשה ביותר בתחום הבינה המלאכותית. לדוגמא, התלמידים ילמדו להבין את השלבים השונים בהבנת ויצירת לשיח, ניהול דיאלוג אדם מכונה, ומהי הארכיטקטורה הנדרשת לפתרונות NLP אמיתיים. הסטודנט ייחשף לכלים עכשוויים המאפשרים לו לטפל בצורות שונות של עיבוד שפה, חלקם דורשים פיתוח וחלקם דורשים אינטגרציה.

נושאים

מבוא ומושגי יסוד – סקירה, השכבות השונות של מידול שפה, תיוג, ניתוח ותחביר של טקסט, עיבוד דיבור ומערכות שיח, עיבוד דיבור ומערכות דיאלוג, עיבוד רגש וסנטימנט, יצירת שפה טבעית ותרגום מכונה, ארכיטקטורה של פתרונות NLP, איך בודקים?

99102 – כריית טקסטים Text Mining

אופן הוראה: שיעור

שעות שבועיות: 3

נקודות זכות: 3

דרישות קדם: למידה עמוקה

מטרת 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 NLP (עיבוד שפה טבעית) ניתוחים סטטיסטיים ושיטות עדכניות של למידה עמוקה.

נושאים

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 ו *MODELLING TOPIC*, תרגום טקסטי

99103 – עיבוד תמונות Image Processing

אופן הוראה: שיעור
שעות שבועיות: 3
נקודות זכות: 3
דרישות קדם: עיבוד אותות

מטרת הקורס

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

נושאים

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.

99104 – עיבוד אותות Signal Processing

אופן הוראה: שיעור
שעות שבועיות: 3
נקודות זכות: 3
דרישות קדם: מתמטיקה למדעי הנתונים

מטרת הקורס

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

נושאים

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

99105 – כריית נתונים מתקדמת Advanced Data Mining

אופן הוראה: שיעור
שעות שבועיות: 3
נקודות זכות: 3
דרישות קדם: מכונה, אחזור וכריית מידע, שימושי למידה עמוקה בדימות רפואי (במקביל – מומלץ)

מטרת הקורס

להכיר לסטודנטים נושאים מתקדמים מתחום כריית הנתונים, בזיקה ללמידת מכונה ולמידה עמוקה

נושאים

הקורס יתמקד בנושאים מתקדמים בתחום כריית הנתונים, תוך זיקה ללמידת מכונה ולמידה עמוקה, כגון: כרייה של נתוני רשת וגרפים, ניתוח קהילות, כריית נתונים מבוססי מיקום, כריית נתונים עיתיים (לפי ציר זמן) זיהוי אנומליות ומערכות המלצה.

99106 – סדרות עתיות ושיטות חיזוי Time series and prediction methods

אופן הוראה: שיעור

שעות שבועיות: 3

נקודות זכות: 3

דרישות קדם: שיטות הסתברותיות וסטטיסטיות מתקדמות

מטרת 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

נושאים

דוגמאות מתחומים שונים, גישות כלליות לניתוח נתוני עיתים, טיפול בטרנד ועונתיות, תהליך סטוכסטי סטציונרי, פונקציות אוטוקורלציה, אמידה של פונקציית תוחלת, תכונות של אוטוקורלציה, תהליכים ליניאריים, הערכת של פונקציית אוטוקורלציה, נוסחת ברלט, מודלים ARMA, פונקציית אוטוקורלציה חלקית, מבוא לפילטרים ליניאריים וורת ספקטראלית, צפיפות ספקטאלית של תהליך ARMA, הערכת תהליך 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.

99107 – למידה לא מונחית מתקדמת Advanced Topics in Supervised Machine Learning

אופן הוראה: שיעור

שעות שבועיות: 3

נקודות זכות: 3

דרישות קדם: למידת מכונה

מטרת 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.

נושאים

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

אופן הוראה: שיעור

שעות שבועיות: 3

נקודות זכות: 3

דרישות קדם: למידת מכונה

מטרת הקורס

למידת חיזוק היא מסגרת למידול האינטראקציה של סוכן אוטונומי עם עולם לא ידוע. מטרת הסוכן היא ללמוד את ההשפעות של פעולותיו ולשנות את מדיניותו על מנת למקסם פרס עתידי. למידה מונחית חיזוקים מבוססת על בינה מלאכותית. בניגוד ללמידה מפקחת (*supervised learning*), בלמידה מונחית חיזוקים לסוכן לא נתונות תוויות (*labels*) כלשהן, אלא עליו ללמוד רק על בסיס ניסוי וטעייה. מסגרת זו משמשת בבעיות רבות, לדוגמא משחקי מחשב (אטארי, שש-בש, שחמט וכיו'), רובוטיקה, רכבים אוטונומיים, ועוד. רעיונות מעולם הלמידה מונחית-החיזוקים משמשים גם לצורך חקר והסבר של למידה בבעלי חיים, הבנת מנגנון הדופמין במוח האנושי ועוד.

נושאים

קורס זה מציג את התשתית לרעיונות הבסיסיים של למידה מונחית חיזוקים מודרנית, כולל תהליכי החלטה של מרקוב, פונקציות ערך, אומדן מונטה קרלו, תכנות דינמי, למידה על בסיס הבדלים זמניים וקירוב פונקציות. קורס זה יפתח הבנה אינטואיטיבית של מושגים אלה, תוך התייחסות לפרספקטיבה של הסוכן והתמקדות בתיאוריה המתמטית של למידה מונחית-חיזוקים.

תקצירי קורסי בחירה – תואר שני M.Sc. במדעי הנתונים

מערכות תבוניות

99201 – שיטות מתקדמות בלמידת מכונה Machine Learning Advanced Supervised

אופן הוראה: שיעור

שעות שבועיות: 3

נקודות זכות: 3

דרישות קדם: למידת מכונה, סטטיסטיקה

מטרת 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 מכונת 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.

נושאים

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

99202 – עיבוד אותות על גרפים Graph Signal Processing

אופן הוראה: שיעור

שעות שבועיות: 3

נקודות זכות: 3

דרישות קדם: עיבוד אותות

מטרת 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.

נושאים

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

אופן הוראה: שיעור

שעות שבועיות: 3

נקודות זכות: 3

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

מטרת הקורס

עיבוד שפה טבעית הוא אחד מאבני הבניין החשובים בבנייה 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 וזובר, 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.

נושאים

- 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)

אופן הוראה: שיעור
שעות שבועיות: 3
נקודות זכות: 3
דרישות קדם:

מטרת 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.

נושאים

1. מבוא למושג של נהיגה אוטונומי. רמות שונות של הנהיגה
2. סקירת מערכת 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

אופן הוראה: שיעור

שעות שבועיות: 3

נקודות זכות: 3

דרישות קדם: למידת מכונה; שיטות הסתברותיות וסטטיסטיות מתקדמות

מטרת 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).

נושאים

- 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

99207 – מערכות תומכות החלטה Decision Support Systems

אופן הוראה: שיעור

שעות שבועיות: 3

נקודות זכות: 3

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

מטרת הקורס

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

נושאים

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. הכרת 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.

99212 – אבטחה והגנה על מערכות נתוני עתק Protection Security of Big Data

אופן הוראה: שיעור

שעות שבועיות: 3

נקודות זכות: 3

דרישות קדם: אין דרישת קדם

מטרת הקורס

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

נושאים

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

99213 – ניתוח רשתות חברתיות ויישומים מסחריים Social Networks Analysis

אופן הוראה: שיעור

שעות שבועיות: 3

נקודות זכות: 3

דרישות קדם: כריית נתונים מתקדמת או למידה לא מונחית מתקדמת

מטרת הקורס

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

נושאים

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.

99214 – למידת מכונה למערכות סייבר Machine Learning for Cyber Systems

אופן הוראה: שיעור

שעות שבועיות: 3

נקודות זכות: 3

דרישות קדם: כריית נתונים מתקדמת או כריית טקסטים

מטרת הקורס

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

נושאים

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.

אופן הוראה: שיעור

שעות שבועיות: 3

נקודות זכות: 3

דרישות קדם: עיבוד אותות

מטרת 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.

נושאים

- 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

99207 – מערכות תומכות החלטה Decision Support Systems

אופן הוראה: שיעור

שעות שבועיות: 3

נקודות זכות: 3

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

מטרת הקורס

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

נושאים

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. הכרת 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.

99208 – יישומים של מדעי הנתונים לאפידמיולוגיה ובריאות הציבור Applications of Data Science in Epidemiology and Public Health

אופן הוראה: שיעור

שעות שבועיות: 3

נקודות זכות: 3

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

מטרת הקורס

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

נושאים

במהלך הקורס הסטודנטים ילמדו מגוון רחב של יישומים של מדעי הנתונים בבריאות הציבור ובאפידמיולוגיה. הנושאים המרכזיים הם: מקורות הנתונים הבריאותיים והרפואיים השונים, ארגוניהם, 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.

99216 – עיבוד תמונות ביו רפואיות Bio-Medical Signal Processing

אופן הוראה: שיעור
שעות שבועיות: 3
נקודות זכות: 3
דרישות קדם: עיבוד אותות

מטרת 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 מצרי מידע. השפעת האינטרנט על שרשרת הערך. יישומים טיפוסיים. מודלים מבוססים מסחר אלקטרוני כגון (B2C, B2B) טכניקות שיווק וסוגיות של אבטחת מידע.

נושאים

- 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

99211 – עיבוד אותות ביו רפואיות Bio-Medical Signal Processing

אופן הוראה: שיעור
שעות שבועיות: 3
נקודות זכות: 3
דרישות קדם: עיבוד תמונות

מטרת הקורס

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

נושאים

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

אופן הוראה: שיעור

שעות שבועיות: 3

נקודות זכות: 3

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

מטרת 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.

נושאים

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

אופן הוראה: שיעור

שעות שבועיות: 3

נקודות זכות: 3

דרישות קדם: אין

מטרת 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.

נושאים

- שינויים סוציו אקונומיים, תרבותיים, טכנולוגיים ואתיקה
- הגדרת נתוני עתק, נתונים אישיים ונתונים אנונימיים
- נתונים אישיים, חוקיות, יושרה, סודיות, הוגנות, שקיפות ו- (GDPR) General Data Protection Regulation.
- 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.

****המחלקה שומרת לעצמה את הזכות לשנות את היצע קורסי הבחירה.**