סילבוס קורס

עקרונות מדע הנתונים

0000000

פרטי הקורס

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  | | --- | --- | --- | --- | |  |  |  |  | | קמפוס: | באר שבע | שנה אקדמית: | תשפ"ב | | **מחלקה:** | הנדסת תוכנה | **סוג קורס:** | בחירה | | **תחום:** |  | **רמת קורס:** | תואר ראשון | | **שנת לימוד:** | ד' | **צורת העברה:** | פנים אל פנים או זום (תלוי במצב האפידומיולגי) | | **סמסטר:** | א' | **דרישות קדם:** | מבוא לתכנות פייתון, מבני נתונים, אלגברה ליניארית, סטטיסטיקה | | **נקודות זכות:** | 3 | **דרישות במקביל:** | - | | **נקודות** ECTS**:** | 4.5 | **שפת הוראה:** | עברית | | **מרצה/ים:** | מר אביעד אלישר | **סביבת עבודה:** |  | | **מתרגל/ים:** |  | **תאריך אישור:** |  | | **דוא"ל מרצה/ים:** |  |  |  | |  |  |  |  | |

מטרה

|  |
| --- |
| הקורס מלמד עקרונות בסיסיים של big data, מטודולוגיות לאנליזה של נתונים מובנים ובלתי מובנים עם דגש על היחס בין מדען הנתונים וצרכים עסקיים. היום יותר יותר אירגונים משתמשים בנתונים שלהם ככלי לקבלת החלטה ולבניית שירותים ומוצרים מלאים במידע. על מנת לתמוך בכך, אוסף הכלים הנדרשים ע"י הארגון נלמד בתחום "Data Science" – מדע מבוסס נתונים. |

תפוקות למידה

|  |
| --- |
| עם סיום מוצלח של הקורס, הסטודנטים יהיו מסוגלים:   1. לנרמל נתונים. 2. לבנות מודלים לסיווג וקיבוץ נתונים. 3. לבצע חיפוש תבניות בנתונים. 4. לעשות חיזוי אירועים. 5. לבצע כריית הטקסט והרשת. |

תוכן הקורס

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | | שבוע | נושא | מקורות רלוונטים | | 1 | מבוא ל-Big Data. סיבות למהפכת ה-AI, יתרונות, חסרונות, הסבר מונחי יסוד בתחום | [1] chapter 1 | | 2 | משימות ושימושים מרכזיים במדעי הנתונים | [1] chapter 1 | | 3 | Data Objects and Attribute Types | [1] chapter 2 | | 4 | תהליך הלמידה: שלבים, אתגרים, פתרונות | [1] chapter 2 | | 5 | תהליך איסוף המידע – בניית Crawlers, שימוש ב-API, ניתוח דאטאסטים, פרטיות וסודיות המידע | [1] chapter 4 | | 6 | עיבוד מקדים של נתונים (Preprocessing) – ניקיון המידע, נרמולים | [1] chapter 3 | | 7 | Data Transformation Strategies – Normalization, Smoothing | [1] chapter 3 | | 8 | Classification – Decision Tree Induction | [1] chapter 8 | | 9 | Classification - עצי החלטה CART, Pruning | [1] chapter 8 | | 10 | מדדי הערכה, Random Forest, Bagging | [1] chapter 8 | | 11 | Artificial Neural Networks | [1] chapter 9 | | 12 | Instance-Based Learning and SVM | [1] chapter 9 | | 13 | נתונים לא מפוקחים: Clustering, K-Means, זיהוי חריגות | [1] chapter 10, 12 | |

מקורות ספרות נדרשים ומומלצים

|  |
| --- |
| ספר הקורס:  1. Jiawei Han, Micheline Kamber and Jian Pei. Data Mining: Concepts and Techniques. Third Edition, 2011.  מקורות נוספים: |

פעילויות למידה מתוכננות ושיטות הוראה

|  |
| --- |
| הקורס יילמד במסגרת של 3 ש"ש הרצאה. |

שיטות הערכה וקריטריונים

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | | בחינה סופית: | 70% | קיימת חובת מעבר בציון 56 בבחינה | | בחנים: | - |  | | תרגילי בית: | 30% | חובת הגשה | | דוחות: | - |  | | פרויקט: | - |  | | נוכחות: |  |  | |  |  |  | |