

שם הקורס: מבוא למדעי הנתונים

שם הפקולטה: מדעי הטבע

המחלקה: מדעי המחשב

שם המרצה: ד"ר רועי יוזביץ

מתכונת הקורס: הרצאה

שנת לימודים: ג סמסטר: א היקף שעות: 3 ש"ס + 1 תרגול

נקודות זכות: 3.5 נ"ז

א. **מטרות הקורס** (מטרות על / מטרות ספציפיות): הקורס מציג לסטודנט ראיית רוחב על תחום ה-DATA SCIENCE. בין היתר, הסטודנטים ייפגשו כאן עם שפת התכנות פייתון, ספריית הפייתון המדעי ולמידת המכונה (SKLEARN). הקורס מספק לסטודנט הכרות עם המתמטיקה ההכרחית לתחום כגון אלגברה לינארית, הסתברות וסטטיסטיקה. כמו כן, ילמדו הסטודנטים על מודלים שונים בלמידת מכונה כמו רגרסיה לינארית, KNN, עצי החלטה ו-NAIVE BAYES.

מהלך השיעורים: הקורס הינו קורס פרונטלי – הרצאה. ושעת תרגול אחת.

ב. **תוכן הקורס:**

תכנית הוראה מפורטת לפי נושאים

יחידת שיעור	נושא השיעור	הערות
1	פתיחה, מבנה הקורס, בסיסים בינאריים ומבוא להסתברות	אקסיומות ההסתברות ושאלות בסיס. ההבדל בין גישות שונות בהסתברות. מה פירוש מטבע מוטה (BIASED) משתנה אקראי (מ"א)
2	הסתברות מותנית וחוק בייס ומשתנים אקראיים. תכנות בשפת פייתון ו-NUMPY	
3	סטטיסטיקה	ממוצע, סטיית תקן, שונות התפלגות נורמאלית התפלגות ברנולי חישובי PDF

מטריצות, וקטורים ודדטרמיננטות הכפלת מטריצות. הגדרת טרנספוז ופיתרון של מערכת משוואות	המשך סטטיסטיקה – אלגברה לינארית	4
What is DataFrame, Index, Reindex, Rank & Sort, Data Alignment, Missing data, Reading csv files	ספריית PANDAS חלק 1	5
כולל מבוא לאלגברה לינארית	ספריית PANDAS חלק 2	6
	ויזואליזציה באמצעות MATPLOTLIB	7
Overfitting vs Underfitting Correctness The Bias-Variance Trade-off Feature Extraction ML vs DL What is Reinforcement Learning?	מבוא ללמידת מכונה	8
Regression vs Classification	Linear Regression/Logistic Regression	10
Regression vs Classification	KNN	11
Type 1, Type 2, F1 score, Accuracy level, chi square	מודלים של שגיאה חלק 1	12
	מודלים של שגיאה חלק 2 הסבר על המטלה	13

ג. מרכיבי הציון הסופי :

שיעורי בית – 30%

עבודת גמר – 70%

ד. **ביבליוגרפיה:** (חובה/רשות) - מסודרת לפי נושאי
הקורס.

1. Grus, Joel. *Data science from scratch: first principles with python*. " O'Reilly Media, Inc.", 2015.
2. McKinney, Wes. *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython*. " O'Reilly Media, Inc.", 2012.