

שם הקורס: מבוא למדעי הנתונים

שם הפקולטה: מדעי הטבע

המחלקה: מדעי המחשב

שם המרצה: ד"ר אור חיים אנידג'ר

מתכונת הקורס: הרצאה +

תרגול

שנת לימודים: א 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). כמו כן, ילמדו הסטודנטים על מודלים שונים בלמידת מכונה כמו רגרסיה לינארית, KNN, עצי החלטה, SVM ועוד.

מהלך השיעורים: הקורס הינו קורס פרונטלי – 3 שעות הרצאה. ושעת תרגול אחת.

ב. **תוכן הקורס:**

תכנית הוראה מפורטת לפי נושאים

יחידת שיעור	נושא השיעור	הערות
1	פתיחה, מטלות ומבנה הקורס, מבוא בסיסי לשפת פייתון + הכרת NUMPY	
2	היכרות עם עולם הלמידה – איך פותרים בעיה?	Supervised and Unsupervised Learning
3	רגרסיה לינארית	
4	רגרסיה לוגיסטית	
5	KNN	
6	עצי החלטה + Random Forest	
7	SVM	
8	Cross Validation + Overfitting +	

	Bootstrap	
	מבוא ל-Unsupervised Learning + Clustering + PCA	9
	K-means + PCA	10
	Hierarchical Clustering	11
	מבוא ל-Loss Functions	12
Type 1, Type 2, F1 score, Accuracy level	מודלים של שגיאה	13

הערה: הסילבוס הינו דינמי, ועשוי להשתנות במהלך הסמסטר כפונקציה של התקדמות בחומר.

ג. מרכיבי הציון הסופי :

מטלות בית – רשות

בחינה סופית -

100%

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.