

סילבוס קורס
מבוא לאיחזור מידע
4077006

פרטי הקורס

שנה אקדמית: תשפא
סוג הקורס: בחירה
רמת הקורס: תואר ראשון
צורת העברה: פנים אל פנים, למידה מרחוק
דרישות קדם:
דרישות במקביל:
שפת הוראה: עברית
סביבת עבודה: מעבדה
מתרגל/ים: מר יניב הדר
yanivha6@ac.sce.ac.il

קמפוס: באר שבע
מחלקה: הנדסת תוכנה
תחום: כריה ואחזור מידע
שנת לימוד: ג'
סמסטר: א
נקודות זכות: 3
נקודות ECTS: 4.5
מרצה/ים: ד"ר מרינה ליטבק
marinal@sce.ac.il

מטרה

הקניית ידע נרחב בעיבוד טקסט ואחזור מידע, כולל שיטות לחיפוש, סינון, דירוג וסיווג מסמכי טקסט ומידע באינטרנט.

תפוקות למידה

- עם סיום מוצלח של הקורס, הסטודנטים יהיו מסוגלים:
1. להתאים ולממש שיטה לבניית ייצוג טקסטואלי של מסמכים.
 2. להתאים ולממש שיטה לעיבוד מקדים של טקסט.
 3. להתאים ולממש שיטה לקיבוץ מסמכים.
 4. להתאים ולממש שיטה לסיווג מסמכים.
 5. להעריך איכות של שיטות לאחזור מידע.
 6. לעצב, להגדיר ולממש מערכות יישומיות לאחזור מידע.

תוכן הקורס

שבוע	נושא	מקורות רלוונטים
1	איחזור בולאני (Boolean retrieval)	[1] פרק 1
2	אוצר מילים (Term vocabulary)	[1] פרק 2
3	מילונים (Dictionaries)	[1] פרק 3
4	מודלים ווקטוריים (Vector Space Model)	[1] פרק 6
5	מודלים ווקטוריים (Vector Space Model)	[1] פרק 6
6	דירוג מסמכים במערכות חיפוש (Scoring and ranking in a search system)	[1] פרק 7
7	מדדים לאיחזור (Evaluation in IR: Precision vs. Recall)	[1] פרק 8
8	מודלים הסתברותיים (Probabilistic Retrieval)	[1] פרק 11
9	אלגוריתמים לסיווג אינפורמציה טקסטואלית (Text Classification)	[1] פרקים 13-15
10	אלגוריתמים לסיווג אינפורמציה טקסטואלית (Text Classification)	[1] פרקים 13-15
11	אלגוריתם לדירוג אתרים (Link analysis, PageRank Algorithm)	[1] פרק 21
12	יישומים באחזור מידע	
13	שיעור חזרה	

מקורות ספרות נדרשים ומומלצים

ספר הקורס:

1. [Manning, Christopher D., Hinrich Schutze, and Prabhakar Raghavan. Introduction to information retrieval. Cambridge university press, 2009](#)



פעילויות למידה מתוכננות ושיטות הוראה

שעות הרצאה שבועיות: 2, שעות מעבדה שבועיות: 2
ההוראה במסגרת הקורס הינה פרונטאלית/מקוונת. במהלך הקורס יינתנו פרויקטים תכנותיים. הסטודנטים יכינו את הפרויקט בזוגות, יגישו דוחות ויציגו אותו בכיתה.

שיטות הערכה וקריטריונים

הערות	אחוז	קריטריון
	40%	בחינה סופית:
יינתנו 6 תרגילים. משקלם זהה. הציון הנמוך ביותר מבין ששת התרגילים לא ישוקלל.	50%	תרגילים:
חובת נוכחות והשתתפות בלפחות 80% מהמפגשים, עם מצלמה פתוחה. בהיעדרות מעל ל- 20%, הציון יחושב באופן יחסי.	10%	נוכחות: