

202-2-5211

עיבוד שפה טבעית

ד"ר יובל פינטר
סמסטר א' תשפ"ד

לפני הכל

- המרחב המוגן הוא בחדר 328 ("צמוד לגרם מדרגות צפוני")
- מוקדי סיוע נפשי 08-6461088, psychserv@bgu.ac.il
- אחראים מחלקתיים לנושאי מילואימניקים: פרופ' גרא וייס, פרופ' אריה קנטרוביץ', גב' זהבה קובני, גב' בת אל אמר גולדשטיין
- מתנדבת לקשר עם הכיתה?

אמ;לק הקורס



- אנחנו רוצים לגרום למחשבים להתמודד עם האתגרים של שפה אנושית
- נלמד מהם האתגרים
- נלמד כמה אופיי התמודדות
- נלמד איך לדעת אם הצלחנו, ואיך נראה העולם שלנו בעקבות ההצלחות

קצת עליי

- רקע בלשני ומדמ"חי

- תחומי מחקר עיקריים בתוך נל"פ:

- **קומפוזיציונליות לקסיקלית:** האם ואיך הבפנוכו של המילים בונה את המשמעות שלהן
- **פרטיות:** איך אפשר לתשאל מודלי שפה גדולים בלי להסגיר פרטים חסויים
- **הסברנות:** "למה המודל עשה את מה שהוא עשה"

- לשון פניה: אתה/הוא

- **עמדתי על מגדר:** אני משתדל להתנסח בגוף שני/שלישי בלשון כוללנית (inclusive) - אתנם, אנשימות, סטודנט.ית. בדיבור קצת יותר קשה לי, אשמח להערות.

מנהלות

- כתובתי: uvp@cs.bgu.ac.il

- שעות קבלה: רביעי 12:00-14:00 (לפני השיעור), אלון (37) חדר 203;
ראשון זמין באופן כללי בזום, 9:00-11:00 תמיד, מומלץ ליצור קשר קודם;
מי שלא יכול.ה בשעות האלה, ליצור איתי קשר ונקבע ספציפית.

- נא לפנות בנימוס

- [הכלל השמרני, בגדול, באקדמיה: פניה ראשונה למי שלא מכירים אותי - בתואר ושם משפחה. במידה ונעניתם בחתימה עם שם פרטי - ניתן לאחר מכן לפנות בשם פרטי]
- בע"פ - "יובל" זה בסדר. שמי האמצעי מתעקש להופיע בכל מיני מערכות אך אני לא משתמש בו.
- רשתות חברתיות - עקבו אחריי לשיקולכם (ציוצי אינם מייצגים וכו' וכו'); אני לא אעקוב חזרה כל עוד הקורס מתקיים; נא לא לפנות בהודעה פרטית ברשתות.

- נוכחות

- אין חובת נוכחות מלבד בהרצאת האורח, אבל הולכות להיות לכם שאלות ולכן מומלץ מאוד להגיע.
בהתאם למתווה השנה ההרצאות היברידיות ומוקלטות אבל עד כמה שניתן נא להגיע פיזית.

מנהלות (המשך)

- שאלות במהלך ההרצאה

- כמה שיותר, אבל נסו בנקודות שמתאימות לכך

- פורום

- כרגע יש במודל
- פתוח להצעות

- משובים

- יהיו רשמיים
- מבקש לפנות אליי בכל עת - משובים זריזים על שיטות הוראה יכולים לטופל במהרה
- אפשר באנונימיות - ליצור כתובת מייל / פתק פיזי (בנין 37 קומה 1, תיבה 23)

קצת עליכם

<https://app.wooclap.com/SSJSPM>



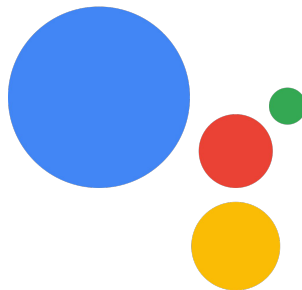
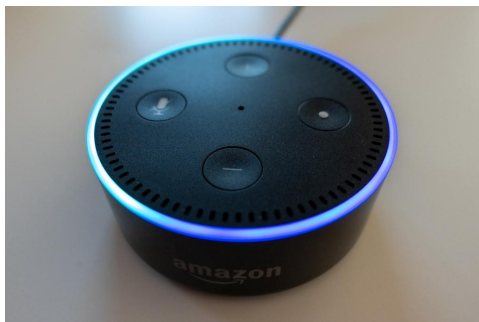
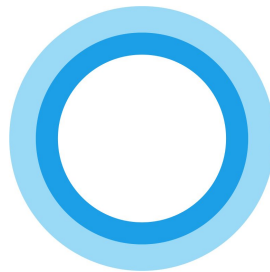
(שאלות 1-8)

אמ;לק עיבוד שפה

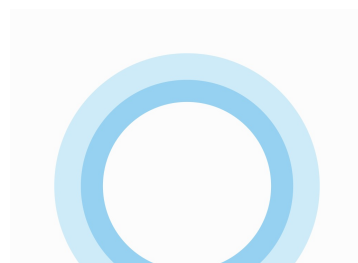
- תחום יישומי
- בעיות שמגיעות מחקר השפה או מיישומי קצה
- פתרונות שמגיעים (לרוב) מלימוד מכונה
- חוסה תחת מטריית בינה מלאכותית

(שאלה 9)

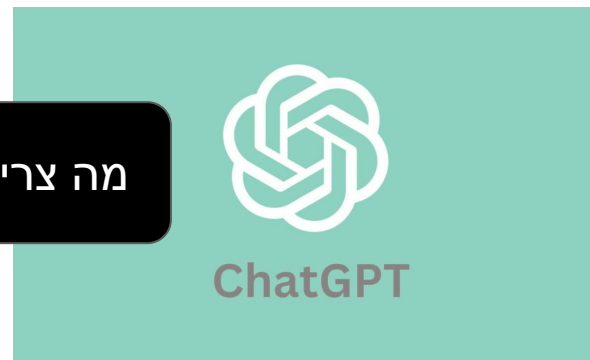
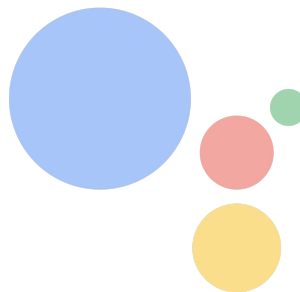
"תחום יישומי"



"תחום יישומי"



מה צריך לקרות כדי שמערכת כזאת תוכל לעבוד?





You

מה אפשר ללמוד בקורס עיבוד שפה טבעית באוניברסיטת בן גוריון?

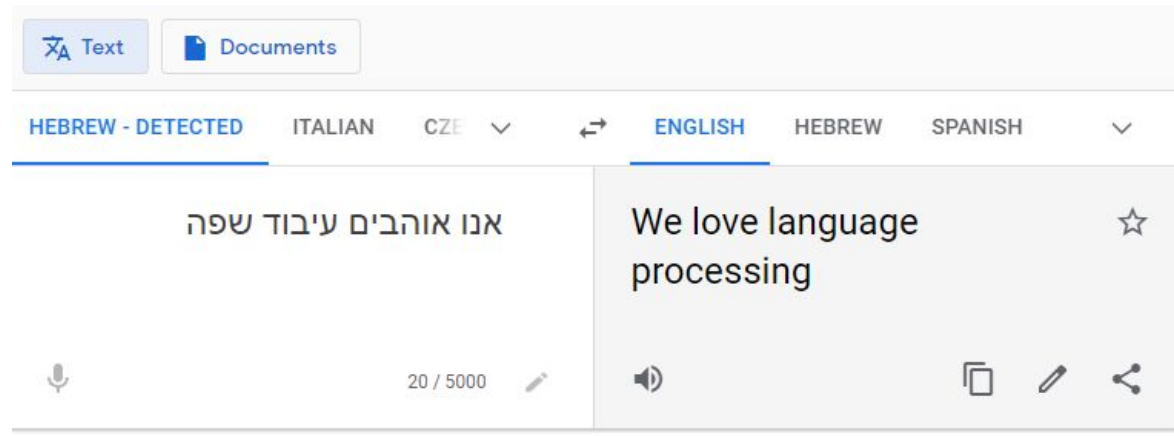


ChatGPT

לפי המידע הזמין לי עד לספטמבר 2021, אוניברסיטת בן-גוריון בנגב בישראל מציעה מספר קורסים בתחום עיבוד שפה טבעית במסגרת תוכנית הלימודים שלה. הנה כמה דוגמאות לקורסים שמתמקדים בתחום עיבוד שפה טבעית:

1. קורס זה מתמקד בשיטות ובטכניקות (Natural Language Processing) עיבוד שפה טבעית לעיבוד, ניתוח והבנת שפה טבעית באמצעות מחשבים. הקורס מתמקד בנושאים כמו דקדוק וסמנטיקה של השפה הטבעית, מודלים חישוביים לעיבוד שפה, ויישומים שונים של עיבוד שפה טבעית כמו תרגום אוטומטי, תגיות טקסט, ואחרים.
2. יישומים של עיבוד שפה טבעית: קורס זה עוסק ביישומים מעשיים של עיבוד שפה טבעית בתחומים רחבים כמו חיזוי טקסט, שימושים בסוגיות של חיפוש מידע, תגובות אוטומטיות למשתמש, ויישומים נוספים בתחום האינטרנט והמחשוב.
3. עיבוד שפה טבעית מתקדם: קורס מתקדם זה עוסק בתכנות מתקדם וטכניקות יישום בעיבוד שפה טבעית. הקורס עשוי לכלול

עוד דוגמא



חקר שפה (בלשנות)

- מהן אבני הבניין היסודיות של שפה?
- איך מרכיבים אותן לתיאור מצבים ולביטוי רעיונות מסובכים ו/או מקוריים?
- אילו שיטות נוקטות שפות שונות לבטא פנים שונות בעולם?
- **השפה כתופעה חברתית**
 - איך החברה משפיעה על השפה?
 - איך שפה משתלבת באמצעי תקשורת אחרים?

לימוד מכונה

- איך מזהים **תבניות** בנתונים?

- עד כמה ניתן ללמוד **מדוגמאות**?

- דאטא אימון מול דאטא מבחן (train vs. test)

- למידה תחת מגבלות שונות:

- נתונים מעוטים

- כוח חישוב מועט

- זמן-אמת

- "למידה עמוקה"

**ספוילר: הרבה אלגברה לינארית,
הרבה הסתברות,
פה ושם חדו"א מרובת-משתנים**

בינה מלאכותית

- איך בונים מכונות שיודעות לעשות דברים שבני אדם עושים?
- אילו מערכות יכולות לעזור לבני אדם בהתנהלות היומיומית?
- מהי ההשפעה של פרישת מערכות כאלה על העולם והחברה?
 - הטיה
 - נגישות
 - פרטיות

תחומים שכנים / משיקים

- עיבוד שפה טבעית עוסק, בכללותו, ב**טקסט**
- זיהוי וייצור קול נכנס תחת תחום **דיבור** (Speech)
 - ASR - זיהוי קול אוטומטי
 - TTS, STT - טקסט לדיבור וההיפך
 - למה זה בנפרד? "סיבות היסטוריות" (מדמ"ח מול הנדסה)
- שפות סימנים - לצערנו אין הרבה, מה שקיים לרוב נכנס ב-NLP
- "טכנולוגיות שפה אנושית" (HLT) - שם נרדף לנל"פ, הרבה פחות בשימוש
- בלשנות חישובית - **ייכלל חלקית בקורס**
- פסיכולשנות - איך **אנחנו** מעבדים שפה? (קשר הדוק למדעי המוח וקוגניציה)
- סוציולשנות - איך אנחנו, כחברה, **משתמשים** בשפה?
- מדעי המידע (data science)
- כריית מידע/טקסט (data mining / text mining)
- ...

עצירה לשאלות

דוגמאות לשאלות מרכזיות בעיבוד שפה

- בהינתן מסמך (נניח ביקורת על ספר), האם **היחס** שלו למושא הכתיבה חיובי או שלילי?
- בהינתן משפט, מהו **הפועל** העיקרי שלו?
- בהינתן טקסט, אילו **ישויות** מופיעות בו?
- בהינתן כינוי גוף בטקסט (pronoun), **לאיזו** ישות הוא מתיחס?
- בהינתן צורת מקור של מילה ותכונות התאם (גוף, זמן), מהי הצורה ה**נטויה** המתאימה?

דוגמאות לשאלות הרבה פחות מרכזיות בעיבוד שפה

- איך מזהים מבע סרקסטי?
- איך נוצרים הלחמים (blends)? איך אפשר לפרק הלחמים?
- איך פותרים הגדרה בתשבץ הגיון?
- איך משחזרים סימנים דיאקריטיים בטקסט איטלקי?
- האם אפשר לחזות אילו מילים שהאקדמיה ללשון מציעה ייקלטו בשפה?

שאלות מרכזיות, שאלות לא מרכזיות

העיקר להגדיר את הבעיה נכון

אתגרים בעיבוד שפה

- **מפתח בהשערה** "א" ⇔ "יש לי רעיון"
- **מערכת יגתמומי תוכן** (domains)
- **דמישועי אלזמן, דמקום**; ועוד
- **יחגמיל לעולמס** בתיק כי הוא [גדול/קטן] מדי."
- **דמבשה שמועות מהדאטא**

עוד אתגרים בעיבוד שפה

- דו-משמעות

- (נדון בהרחבה)

- מילים לא-מוכרות

- התפלגות מילים בגוף טקסט גדול היא **זיפיאנית** (Zipfian): עבור המילה ה-n בשכיחותה, $Pr(w_n) \propto \frac{1}{n^s}$

- קומפוזיציונאליות והעדרה

- דיבר+ת = דיברת; למד+ת = למדת; ברק+ת = ברקת?

- בית-הוריי \Leftrightarrow בית-ספר

- אכל את הראש = אכל את הראש??!

עיבוד שפה והחברה

- יש יישומים של עיבוד שפה בעלי השלכות משמעותיות על החיים
 - פיננסים
 - משפטים
 - רפואה
- איך דואגים שהמערכות בשימוש יהיו הוגנות?
 - איך מגדירים מערכת "הוגנת"?
- האם ואיך אפשר להסביר למשתמש-קצה למה המערכת החליטה מה שהחליטה?
 - עד לפני 7-8 שנים?
 - ומאז?

Text Documents

DETECT LANGUAGE HEBREW ITALY ↕ ENGLISH HEBREW SPANISH

tbu tuvcho ghcus apv

tbu tuvcho ghcus apv ☆

Did you mean: תביאו תוכוחו גהכוס אפּו

20 / 5000

Text Documents

HEBREW - DETECTED ITALIAN CZE ↕ ENGLISH HEBREW SPANISH

אנו אוהבים עיבוד שפה

We love language processing ☆

20 / 5000

Text Documents

DETECT LANGUAGE ENGLISH HEB ↕ ENGLISH HEBREW SPANISH

I was a secretary until yesterday ×

הייתי מזכירה עד אתמול ☆

33 / 5000

Text Documents

DETECT LANGUAGE ENGLISH HEB ↕ ENGLISH HEBREW SPANISH

I was a physician|until yesterday ×

הייתי רופא עד אתמול ☆

33 / 5000

Text Documents

DETECT LANGUAGE ENGLISH HEBREW ITALY ↕ ARABIC SPANISH ENGLISH

אני עורכת-דין ×

أنا محام ☆

'ana muham

13 / 5000

Text Documents

DETECT LANGUAGE ENGLISH HEBREW ITALY ↕ ARABIC SPANISH ENGLISH

אני מזכירה ×

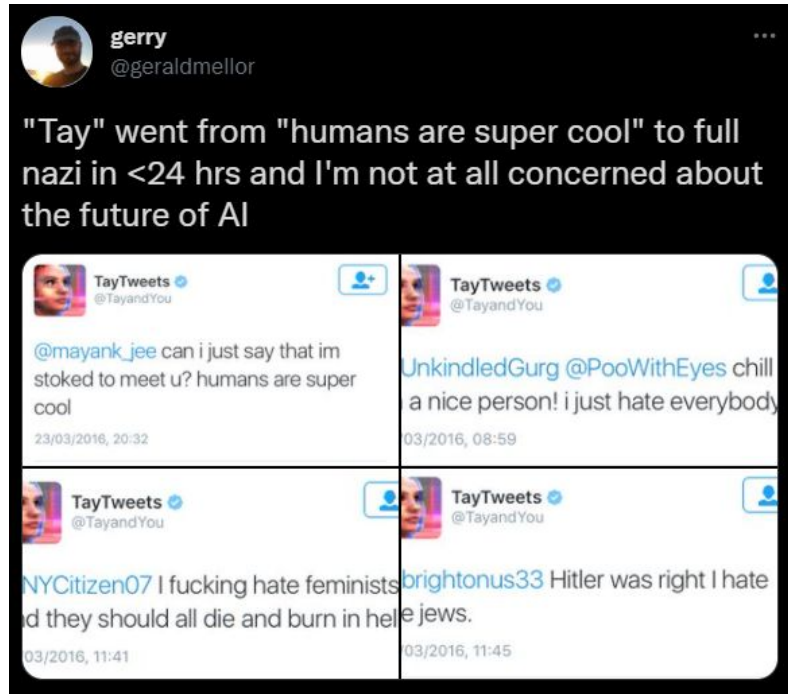
أنا سكرتيرة ☆

'ana sikritira

10 / 5000

Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day

By James Vincent | Mar 24, 2016, 6:43am EDT



[illegible]

The guinea pigs are the only ones who have a grudge
It is as if they are not the case
Have a good look at some of the things you need
The throne has been disregarded by the throne
Declaration of receipt of a contract
The call will be accepted by the Executive Board
We need to get rid of it
The bribe in the case
The sacrilege of the exchequer
The Krebschins geckasaxenzag
The arches are classified as an e
The doorway to the doorway to t

Hebrew – detected ↕ English ▾

✕ קצח Foolish

ספריית חב"ד ליובאוויטש

ספריית חב"ד ליובאוויטש

דף הבית / ספרי כ"ק אדמו"ר הזקן / שולחן ערוך / אורח חיים חלק א / הלכות תפילין: / לב דין כתיבת התפילין ובו פ'... / כ /

ספרי כ"ק אדמו"ר הזקן < שולחן ערוך < אורח חיים חלק א < הלכות תפילין: < לב דין כתיבת התפילין ובו פ' ... > כ

44

32

וְכַאֲשֶׁר נִסְמְקָה אַחַת מֵאֲחֻתוֹת הַפְּשׁוּטוֹת שֶׁאֵין לָהֶם אֵלָּא יָד אַחַת³⁷ כְּגוֹן וְאִיזוֹ אוֹ וְיִיזוֹ אוֹ נִיזוֹ מִטְּסֵה וְכִיכּוּזָה בְּרֹאן אֵין שִׁיעוֹר הַכֶּשֶׁר כְּמֵאלֵה אֹת קִטְעָה שֶׁהִיא יִיזוֹ כְּגוֹן בְּרֵבֶל הָחַיִּי וְכִיכּוּזָה בָּהּ שֶׁהִיא וְאִיזוֹ שֶׁלֹּא תִשְׁתַּיֵּר מִמֶּנּוּ אֵלֶּם כְּמֵאלֵה יִיזוֹ אִזּוֹ אוֹ וְאִיזוֹ כֻּלָּם וְכֵן שֶׁבֶּאֱרַח אֲחֻתוֹת הַפְּשׁוּטוֹת אֵלֶּם הַשְׁמֵר תִּלְוִי בְּקִבְרֵיָתָהּ הַתֵּינִיק שֶׁאֵין חֵטֶם וְלֹא טֶשֶׁשׁ שֶׁאֵם הַתֵּינִיק יוֹדֵעַ לְקִלְרוֹתוֹת הַחֲלֵתוֹן כִּשְׁתֵּרָתוֹ וְאֵם אֵלֶּם³⁸ מִטְּסֵה אֵלֶּם הַשְׁמֵר חֲכָם שֶׁבְּרֵבֶל הָחַיִּי³⁹ דִּהְיִי מִירוֹשׁ הַמִּלּוֹת הִרִי הוּא אֲחֵרֵת הַתֵּינִיק מִתּוֹךְ הַבְּנֵי שֶׁמִּתְפַּרְשִׁים לֵבֵי יַלְיֵי הַתֵּינִיק וְאִין שֶׁמִּסְמְקִים אֵלֶּם בְּעַל אִיזוֹ חֲכָם כִּד כֹּךְ לְהַבִּיז הַתֵּינִיק אֵיזוֹ שֶׁהוּא חֲכָם בְּחִבְרַת הָאֲחֻתוֹת הַיִּטִּיב⁴⁰ שֶׁמִּכְבֹּן עֲלֵיו וְהוּא לֹא מִטְּסֵה

שאמרו

- ☐ ספרי הבעל שם טוב
- ☐ ספרי הרב המגיד ממעזריש
- ☒ ספרי "ק' אדמו"ר הזקן
- ☐ תניא
- ☐ תורה אור
- ☐ דלקוטי תורה
- ☐ מאמרי אדה"צ
- ☒ שלחן ערוך
- ☐ שו"ת
- ☐ פסקי אדמו"ר
- ☐ אגרות קודש
- ☐ הלכות פסח עם ביאור דברי שלום

שרשור הסבר למתעניינים

מה... קורה פה?

● ופה?

מונחי יסוד בעיבוד שפה

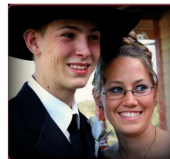
- **מטריקה (metric):** שיטת מדידה שקובעת את ביצועיה של מערכת M על משימה T בהינתן דאטא D
 - למשל: אחוזי דיוק
 - חיוני כחלק מהגדרת הבעיה
- **קורפוס (corpus):** אוסף מסמכים בשפה (או שפות) שמשמש לניתוח לשוני או סטטיסטי ולא־מיון בלתי-מבוקר (unsupervised)
 - למשל: אוסף כל הויקיפדיות של שפות עם למעלה מ-100,000 ערכים
- **דאטא משימתי (task data):** אוסף מסמכים **ותגיות** (annotation, labeling, tagging) שמשמש ללימוד משימה מבוקר (supervised) או חצי-מבוקר (semi-supervised)
 - למשל: אוסף ביקורות הסרטים מ-IMDb בשילוב הציונים שהוענקו
 - למשל: אסופת משפטים שכל הישגיות (שמות אנשים, שמות מקומות, שמות ארגונים) מסומנות בהם

עיבוד שפה - הנגדות

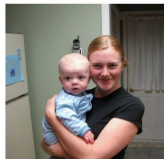
- "הבנה" מול "יצירה" (NLU vs. NLG)
- "מסמכים" מול "אטומים" (document-level vs. sequence tagging)
- "רצפים" מול "שק מילים"
- ארכיטקטורות גנריות מול ארכיטקטורות עם ידע מוטמע (inductive bias)

עיבוד שפה ועוד

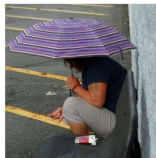
Who is wearing glasses?
man



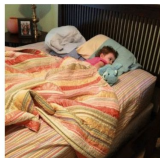
Where is the child sitting?
fridge



Is the umbrella upside down?
yes



How many children are in the bed?
2



Goal: "Rinse off a mug and place it in the coffee maker"



● מולטימודאליות (multimodal)

● קרקוע (grounding)

● תחומים שהתקדמו מאוד מאז שיש

למידה עמוקה
○ (למה?)

מידול שפה (language modeling)

- משימה קלה להגדרה, בעלת שם מבלבל ומידת יישומיות חמקמקה
- הגדרה קלאסית: **בהינתן רצף מילים, מהי המילה הבאה?**
- עבור מסמך שלם D עם n מילים (ונניח סימן מיוחד לתחילת המסמך כ- w_0):

$$Pr(D) = \prod_{i \in 1..n} Pr(w_i | w_0, \dots, w_{i-1})$$

עצירה לשאלות

מבנה הקורס

- עיבוד שפה הוא תחום של **בעיות**
- לפני שניגשים לפתור בעיה צריך **להגדיר** אותה
- לכן רוב היחידות (~שבוע) יכללו:
 - חצי הרצאה עד הרצאה של הגדרת הבעיה (לשונית) ואיך נמדוד פתרון
 - שיטה מרכזית מתחום למידת המכונה (ML) ויישומה על הבעיה שהוצגה
 - (לפעמים גם שיטה "פשוטה" יותר, לפעמים מספר שיטות, לפעמים עוד בעיות עם אותה שיטה)
- יחידות נוספות:
 - ~~מיתח טקסט בסיסי~~
 - ~~לבוש שלום (מבטוח, מוכן, מוכן, מוכן, מוכן)~~
 - ~~שבת 10-19 דלי שפה גדולים~~
 - השלכות של עיבוד שפה (ובינה מלאכותית ככלל) על החברה האנושית
 - יישומים של עיבוד שפה בתחומים קשורים (מי ה"לקוחות" שלנו)
- לו"ז מפורט ברמת השיעור, לרבות שינויים במוד ההוראה, נמצא במודל

שבוע אחרון

תאריך	יחידה	תוכן	SLP	בוחן	תרגיל
2/1/24	מבואות	מבוא לנל"פ			
3/1		עיבוד טקסט	2	0 (ללא ציון)	
9/1	סיווג מסמכים	סנטימנט			
10/1		בייז התם + רגרסיה לוגיסטית	4	1	1 יוצא
16/1		חלקי דיבר וישויות			
17/1	תיוג רצפים	מרקוב, ויטרבי	4-5	2	
23/1		רשתות נוירונים	8		1 מוגש
24/1		תיוג רצפים נירוני	8	3	2 יוצא
30/1	ניתוח תחבירי	תחביר	7		
31/1		אלגוריתמי ניתוח תחבירי	9	4	
6/2	חזרה ועיבוי	השלמות	17-18		1 חוזר, 2 מוגש
7/2		חזרה	18		3 יוצא
13/2	שיכונים	סמנטיקה לקסיקלית, דמיון מילים	23		
14/2	(Embeddings)	סקיפגראם, שימושים קלאסיים	6	5	
20/2	חילול	תרגום מכונה, סיכום טקסט	(21)		2 חוזר, 3 מוגש
21/2		רובוטריקים (Transformers)		6	4 יוצא
27/2		טכניקות אימון מתקדמות	13		
28/2	רובוטריקים מתקדם	דגימה והנחיה	10-11	7	5 יוצא
5/3		מולטימודאליות			3 חוזר, 4 מוגש
6/3	יישומי	הרצאת אורח (הבנת שפה טבעית)		8	
12/3		חברה + אתיקה			
13/3		חזרה		9	4 חוזר, 5 מוגש
17/3					5 חוזר
29/3	מועד א				
10/4	מועד ב				
???	מועד ג				

בחירות?

חומר

- ספרי לימוד עיקריים:

- Dan Jurafsky & James H. Martin: *Speech and Language Processing* (3rd edition)
 - זמין חינם לחלוטין ברשת, פרקים מתעדכנים, גרסאות קודמות זמינות
- Jacob Eisenstein: *Natural Language Processing*.
 - יש ספר מודפס, אנחנו נשתמש בגרסת הטיוטה הזמינה ברשת (ובמודל) מתאריך 13/11/2018

- מספרי פרקים בראש כל חלק במודל (ובשקף הבא) - לא בהכרח חד-חד ועל

- חומר נוסף יינתן לעתים פר נושא (סרטונים, טוטוריאלים, מאמרים)

- עדיין, חשוב לעקוב אחרי השיעורים.

יחידות מרכזיות (שבועות 2-8)

תופעה לשונית	בעיה אלגוריתמית	גישות לפתרון
אפיון מסמכים (documents)	סיווג (classification)	גרסיה לוגיסטית (LR), בייז נאיבי (NB), פרספטרון (perceptron) (SLP 4-5, E 2-4)
תכונות מילים (word classes)	תיוג רצפים (sequence tagging)	מודל מרקוב סמוי (HMM), רשת נוירונים נשנית (Recurrent NN) (SLP 7-9, E 7-8)
תחביר (syntax)	ניתוח תחבירי (parsing)	תכנות דינמי לעצים (CKY), הזחה-צמצום (shift-reduce) (SLP 12-14, E 10-11)
סמנטיקה לקסיקלית (lexical semantics)	דמיון מילים (word similarity)	שיכון מילים התפלגותי (distributional word embeddings) (SLP 6;18, E 14)
תרגום (MT) / תמצות טקסט (summarization)	חילול טקסט (text generation)	רצף-לרצף (seq2seq), מקודד-מפענח (encoder-decoder), מודולי צומי (attention), מודלי שפה גדולים (LLMs = BERT, GPT) (SLP 10, E 18-19) (SLP 9.7-8,11)

יחידות שלא יילמדו השנה ואפשר להשלים עצמאית

תופעה לשונית	בעיה אלגוריתמית	גישות לפתרון
סמנטיקה "גבוהה"	ניתוח תפקידים סמנטיים (SRL); הסקה (inference)	משאבים מילוניים (semantic graphs), (15,19, E 12-13)
חקר השיח (discourse) (analysis) ודיאלוג (dialogue)	סייענים אישיים (personal) (assistants), בוטי-שיח (chatbots)	(שתי השורות לעיל +) שמירת מצבים (state) (tracking), מילוי חריצים (slot-filling) (SLP 21-22,24, E 15-16)

שפת הוראה

- עיבוד שפה טבעית הוא תחום מחקר **אנגלוצנטרי** באקסטרים
- עם זאת, אנחנו בישראל ורובנו דוברים עברית כשפת אם
- כל עוד קל ונוח להסביר עם **דוגמאות** בעברית, כך אעשה
- כאשר לא נדרש ידע בשפה, יהיו דוגמאות משפות שאינן עברית או אנגלית
- בכל יתר המקרים, הדוגמאות יהיו באנגלית (וכך גם בספרי הלימוד)
- **המונחים** בתחום, כמו ברוב המדע, מוכרים בשמם האנגלי. אשתדל להשתמש בתרגומם לעברית (לפעמים מקורי שלי) אבל תמיד אתן גם את גרסת המקור.
- להתפלגות כישורי האנגלית יש תומך רחב, ולכן:

Know what you don't know

למה לצפות מהקורס

- האם אבין יותר מהי שפה אנושית?
 - כן.
- האם אדע להגדיר בעיות הנוגעות להבנה של שפה, ולגשת לפתרון?
 - כן.
- האם אדע **לפתח** מודל שפה גדול?
 - כנראה שלא.
- האם אדע **לעבוד** עם מודל שפה גדול?
 - כנראה שכן.
- האם אדע לעבוד **בלי** מודל שפה גדול?
 - כן.

דרישות הקורס

- מבחן מסכם - 50% מהציון (מועד א' 29/3, מועד ב' 10/4, מועד ג' טרם נקבע)
 - אין חובה לעבור כדי לעבור את הקורס - הציון המשוקלל הוא הקובע; ציון מעבר (תואר מתקדם) הינו 65.
- חמישה תרגילי בית - 40%
 - לרוב כל שבועיים
 - כמעט שבועיים להגשה
 - לפחות שניים יכללו מרכיב תכנותי
 - היה אמור להיות בודק, נאחל לשובו בשלום ובבטחה מהמילואים
- מיני-בחינים שבועיים - 8%
 - לא בזמן השיעור
 - לפחות 96 שעות להגשה, ולרוב מרביעי עד שלישי
 - ללא חומר עזר - אבל יש שני נסיונות (שביניהם מותר לחזור על חומר)
 - הבוחן בעל הציון הנמוך לא ייחשב (ובוחן 0 ללא ציון)
- נוכחות בהרצאת אורח.ת - 2%
 - רשמו ביומנים - 6/3/24
 - ניתן להמיר בציון המיני-בוחן הנותר

מילואים והקלות

- הסיבה המרכזית שיש יותר תרגילים השנה (והם קצרים יותר) היא לאפשר גמישות למילואימניקים לקבל פטור מהגשה של חלק, וזה יהיה מוד ההקלות העיקרי. מרכיב התרגילים ייקבע על-סמך הציונים בתרגילים שיוגשו.
- מועדי ההגשה המתוכננים כרגע הם: 23/1, 6/2, 20/2, 5/3, 13/3. נסו לכוון מוקדם ככל הניתן לפי העומס הצפוי, אם יש אפשרות לדעת מראש.
- לגבי המיני-בחנים, תהיה אפשרות להגשה מאוחרת אבל הם ייפתרו בכיתה מיד עם מועד ההגשה המקורי ואני מצפה ליושרה.