



אוניברסיטת בן-גוריון בנגב  
הפקולטה למדעי הרוח והחברה

שנת הלימודים תשע"ז-ח – 2017-18

שם הקורס: מבוא לעיבוד שפה טבעית למדעי הקוגניציה והחברה

מס' קורס: 12920228

שם המרצה: פרופ' יאיר נוימן

שעות קבלה: בתיאום מראש

מטרות הקורס: ללמוד מספר מושגים בסיסיים בעיבוד שפה טבעית ואת האופן בו ניתן לעשות בהם שימוש להבנת מערכות קוגניטיביות וחברתיות.

מבנה הקורס:

הנושאים שידונו במסגרת הקורס:

1. tf-idf

2. n-gram

3. Word sense disambiguation

4. POS tagging

5. Parsing

6. Information retrieval

דרישות והרכב ציון הקורס\*

**משקל בציון הסופי \*\***

נוכחות \_\_\_\_\_ % - 100 \_\_\_\_\_

בחנים \*\*\* \_\_\_\_\_ % -

עבודות \_\_\_\_\_ % - 100 \_\_\_\_\_

מבחנים \_\_\_\_\_ % -

סיורים \_\_\_\_\_ % -

סה"כ - 100%

רשימה ביבליוגרפית (ראשונית וניתנת לשינוי):

\*References:

Major reference for the first ten meetings:

Manning, C. D., & Schütze, H. (1999). *Foundations of statistical natural language processing* (Vol. 999). Cambridge: MIT press. Topics to be discussed: tf-idf, collocations, n-gram, word sense disambiguation, POS tagging, parsing, IR, clustering, text categorization, text summarization, ML algorithms, dependency parsing etc.

References for the following 10 meetings:

1. Turney, P. D. (2002, July). Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th annual meeting on association for computational linguistics* (pp. 417-424). Association for Computational Linguistics.
2. Gabrilovich, E., & Markovitch, S. (2007, January). Computing semantic relatedness using wikipedia-based explicit semantic analysis. In *IJCAI* (Vol. 7, pp. 1606-1611).
3. Koppel, M., Schler, J., & Argamon, S. (2009). Computational methods in authorship attribution. *Journal of the American Society for information Science and Technology*, 60(1), 9-26.
4. Dagan, I., Dolan, B., Magnini, B., & Roth, D. (2009). Recognizing textual entailment: Rational, evaluation and approaches. *Natural Language Engineering*, 15(4), i-xvii.
5. Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., ... & Ungar, L. H. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. *PloS one*, 8(9), e73791.
6. Neuman, Y., Assaf, D., Cohen, Y., Last, M., Argamon, S., Howard, N., & Frieder, O. (2013). Metaphor identification in large texts corpora. *PloS one*, 8(4), e62343.
7. Pakhomov, S., Chacon, D., Wicklund, M., & Gundel, J. (2011). Computerized assessment of syntactic complexity in Alzheimer's disease: a case study of Iris Murdoch's writing. *Behavior research methods*, 43(1), 136-144.
8. Dodds, P. S., Harris, K. D., Kloumann, I. M., Bliss, C. A., & Danforth, C. M. (2011). Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. *PloS one*, 6(12), e26752.
9. Neuman, Y., Turney, P., & Cohen, Y. (2012). How language enables abstraction: A study in computational cultural psychology. *Integrative Psychological and Behavioral Science*, 46(2), 129-145.
10. Neuman, Y., & Cohen, Y. (2016). A Novel Methodology for Automatically Measuring Psychological Dimensions in Textual Data. *The Computer Journal*, 59(9), 1408-1414.

\*The list is tentative and may be changed.

יקבעו בתחילת הקורס ולכל המאוחר בשבועיים הראשונים של הסמסטר ולפני תום תקופת השינויים. \*

אין לשנות את מרכיבי הציין הסופי במהלך הסמסטר. \*\*

על המורה להודיע מראש על סוג הבחנים ומועדם. \*\*\*