Timothy James Leffel

updated dec17/2017

✉ tjleffel@gmail.com
☎ +1(614)306-7814
http://lefft.xyz
github.com/lefft

ABOUT

I'm a generative linguist by training, and a research/data scientist by profession. I've worked on projects ranging from document classification and topic modeling to data ingestion and automated reporting. I approach NLP problems with industry standard tooling and the intuitions of an empirically-focused language scientist. Enthusiast of: R, Python, the NBA, the history of math/logic, math/logic.

WORK EXPERIENCE

2016– Research Scientist, Academic Research Centers. NORC at the University of Chicago.

>> social media text mining (sentiment analysis, document classification, topic modeling)
>> pre-processing; training and evaluation; regression; simulation and resampling; visualization
>> development of data ingestion pipelines and automated reporting systems for RCTs

2015 Data science intern (p/t). Vantage Sports Inc.

>> development and implementation of daily fantasy sports lineup optimization algorithm
>> evaluation of competing predictive models for daily fantasy performance of NBA players

2014–16 Postdoctoral Researcher. Language Processing Lab, Department of Linguistics. University of Chicago

>> experimental design and implementation (surveys, eye-tracking for visual world studies)
>> data analysis and modeling (mixed-effects and logistic regression, RM-ANOVA, others)
>> dissemination of research results (conference presentations, proceedings and journal papers)

2010–16 Adjunct Instructor and Teaching Assistant in Linguistics and Psychology.

>> primary instructor or TA for various undergrad courses in linguistics and psychology
>> 2016 at the University of Chicago; 2010–2014 at New York University

EDUCATION

2014 PhD, Linguistics. New York University

2011 MA, Linguistics. New York University

2009 BA with honors, *magna cum laude*, Linguistics/Philosophy. The Ohio State University

TECHNICAL SKILLS

programming and scientific computing

>> **R**—core packages: `dplyr`, `ggplot2`, `purrr`, `lme4`, `rstan`, `quanteda`. Happy in base or hadleyverse.
>> **Python**—core libraries: `pandas`, `numpy`, `scipy`, `nltk`, `spacy`, `scikit-learn`. Usually use iPython/Jupyter.
>> basic competency in Javascript, Haskell, Julia, Matlab. Quick and enthusiastic learner.

algorithms, analysis techniques, etc.

>> NLP: multinomial naive Bayes; decision trees; LDA; *n*-gram probability models; morphological analysis
>> regression modeling (OLS, logistic regression, censored models); NHST; Bayesian estimation

other skills

>> AWS cloud computing (EC2/S3); (R)markdown; Shiny; basic web development; git; github; LaTeX