Timothy James Leffel

updated jan31/2020

✉ tjleffel@gmail.com

☎ +1(xxx)xxx-xxxx

http://lefft.xyz 🌍

github.com/lefft

ABOUT

I'm a generative linguist by training, and a data scientist and researcher by profession. I have worked on tasks ranging from document classification and topic modeling to the development of automated data ingestion and reporting systems. I approach NLP problems with industry standard tooling and the intuitions of an empirically-focused language scientist. Publications, teaching materials, etc. are available at http://lefft.xyz.

WORK EXPERIENCE

2019– Research Scientist. Amazon.

>> Applied machine learning for advertisement
>> Ad sourcing and relevance modeling

2016–19 Data Scientist (Research Scientist 2016-18). NORC at the University of Chicago.

>> social media text mining (sentiment analysis, document classification, topic modeling, ...)
>> pre-processing; classifier training and evaluation; simulation and resampling; visualization

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)

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

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

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

TECHNICAL SKILLS

core languages and computing stack
>> **Python**: pandas, numpy, scikit-learn, gensim, nltk, keras
>> **R**: dplyr, ggplot2, lme4, shiny, text2vec. Strong command of both base and hadleyverse dialects
>> **tools**: AWS EC2/S3; basic web dev (Django); (R)markdown; bash scripting; git/github; html/css; LaTeX
>> **other**: basic literacy in C/C++, Javascript, and Haskell. Enthusiastic and quick (but careful) learner

tasks and analysis techniques
>> deep experience with classification of short texts and construction of term embeddings from corpus data
>> neural network-based modeling for document classification and term representation/clustering tasks
>> mixed-effects regression modeling; permutation and randomization testing for experimental data
>> time-series analysis of psycholinguistic data for eye-tracking and neuro-imaging (MEG) studies