

Alexander KLAPHEKE

Boston-based data scientist with linguistic expertise

M.A. Linguistics, Harvard University 2019

B.A. Linguistics, Columbia University 2012

Cambridge, Mass.

(860) 605-6150

alexklapheke@gmail.com

alexklapheke.github.io

EXPERIENCE

Data science fellow, General Assembly 2020

Created data science reports & applications during 480-hour intensive training, including: data sourcing & cleaning, exploratory data analysis, feature engineering, regression and classification modeling, and machine learning.

Instructor, Harvard University 2014-2017

Taught 5 seminars in logic, semantics, discourse analysis, and philosophy of language.

Received Distinction in Teaching Award in fall 2017 for universally positive feedback.

Research assistant, Columbia University 2011-2012

Wrote BASH & AWK scripts to search 100-million-word linguistic corpus for NLP application.

PROJECTS

Essay scoring github.com/alexklapheke/essay

Analyzed essay scoring patterns, and trained a dual-input recurrent neural network on 350,000 words, achieving over 60% accuracy.

Menu analysis github.com/alexklapheke/menu

Analyzed trends from over 17,000 menus going back to the 1870s; trained a Bayesian model to categorize menu items by decade.

COVID-19 github.com/alexklapheke/covid-sentiment

Analyzed sentiment from two subreddits to gain insight into public perceptions of statewide public health policies.

Ames, Iowa github.com/alexklapheke/ames

Predicted housing prices using linear regression models in Python.

NPI intervention DOI: 10.15496/publikation-32629

Gathered linguistic data in 3 languages using Mechanical Turk, and analyzed using linear mixed models in R.

SKILLS

Python Keras, Matplotlib,

NLTK, NumPy, Pandas,

scikit-learn, spaCy, Seaborn

R dplyr, ggplot2, lme4, lsmeans

Haskell

Scala

BASH

SQL MySQL

Git

LaTeX Beamer, TikZ

HTML5 & CSS

Apache Spark

Amazon Web Services EC2, S3

Amazon Mechanical Turk