Dr. Ekaterina Kravtchenko

Located in Berlin, Germany.
Address redacted for privacy.
Please contact me by email or on LinkedIn!

Email: eskrav@gmail.com
Homepage: http://eskrav.github.io
Citizenship: USA (Ger-EU permanent residence)

Education

Ph.D. Language Science and Technology; *Saarland University*, Germany; 2022.

M.A. Linguistics; *University of California, Santa Cruz*, USA; 2013.

B.A. Linguistics & Germanic Studies; *University of Massachusetts, Amherst*, USA; 2008.

Research and Work Experience

Fraugster, Berlin, Germany; Machine Learning Team

Senior Data Scientist; 2019–present.

- Developed and deployed machine learning classifiers for fraud detection (logistic regression, ensemble methods), resulting in a decrease in chargeback rates for e-commerce sellers.
- Led cross-team projects to enhance company ML tools, including implementing oversampling for imbalanced class distributions and creating Python modules for automated feature selection, as well as developing a custom method of cross-validation for time series data.
- Analyzed and improved existing ML practices and processes, leading projects to implement necessary changes.
- Onboarded and mentored junior colleagues.

Saarland University, Saarbrücken, Germany; Department of Language Science and Technology

Research Associate; 2014–2019.

- Created and programmed predictive probabilistic models of empirical language phenomena.
- Designed, programmed (JavaScript), and statistically analyzed data from crowdsourced behavioral experiments using multilevel modeling (Python, R).
- Created novel web-based method for empirically measuring the strength of listeners' pragmatic inferences from complex utterances, resulting in 3 papers, 6 conference presentations, and 2 invited talks.
- Organized statistics reading group focused on mixed effect models in R; PhD representative 2014-2017.

Intel Corporation, Hudson, MA, USA; Linguistic Resource Group;

Software Engineer: Linguist; 2013–2014.

- Created and co-managed a cross-site team of 6 linguists and computer scientists in systematically testing and guiding the development of linguistic software prototypes using Agile development methodology. Launched products include the Oakley Radar Pace.
- Co-developed an automated framework (Python) for rapid evaluation of software release functionality.

University of California, Santa Cruz, USA; Linguistics Department

Masters Student; 2010–2013.

- Performed manual collection and detailed annotation of texts, and conducted advanced statistical analyses and visualization of experimental and corpus data (R).
- Provided experimental evidence for efficiency-based omission of linguistic elements in predictive contexts, resulting in 1 paper, 2 conference presentations, and 2 invited talks.

Harvard University, Cambridge, MA, USA; Polinsky Language Processing Lab

Lab Manager and Research Assistant; 2008–2010.

- Programmed, recruited for, and designed stimuli for web- and lab-based language experiments. Coded, organized, and statistically analyzed data, resulting in 3 co-authored publications and 3 conference presentations.

Skills and Projects

Computer & Technical

Programming and Markup Languages: Python (intermediate; scikit-learn, pandas, matplotlib), R (intermediate; dplyr, ggplot2, lme4), SQL (intermediate), JavaScript (beginner), go (beginner), Julia (beginner), LaTeX

Machine Learning and Statistics: classical machine learning models (linear and logistic regression, random forest, XGBoost), multilevel (mixed effect) linear and logistic models, hypothesis testing

Data Collection and Analysis: Jupyter, RStudio, Tableau, Amazon MTurk

Miscellaneous: git, Amazon SageMaker, Amazon Athena, docker, bash, JIRA

Operating Systems: Linux (Ubuntu, Fedora), Windows XP-11, Mac OS X

Independent Coursework

Udacity Machine Learning Nanodegree: Used linear regression, supervised learning, unsupervised learning, reinforcement learning, and deep learning (scikit-learn, Keras) to gain insights into a variety of datasets.

Udacity Data Analyst Nanodegree: used R (tidyverse), Python (pandas, matplotlib), and Tableau to explore several datasets and compose sample reports.

Coursera Data Structures and Algorithms Specialization (expected completion 2023).

Languages

Native: English, Russian

Proficient: German

Recent Research Publications

Demberg, V., **Kravtchenko, E.** & Loy, J.E. (in press). A systematic evaluation of factors affecting referring expression choice in passage completion tasks. In *Journal of Memory and Language*.

Kravtchenko, E., and Demberg, V. (2022). Informationally redundant utterances elicit pragmatic inferences. *Cognition* 225.

Kravtchenko, E., and Demberg, V. (2022). Modeling atypicality inferences in pragmatic reasoning. In *Proceedings of the 44th Annual Meeting of the Cognitive Science Society* (pp. 1918–1924).

Kravtchenko, E., and Demberg, V. (2015). Semantically underinformative utterances trigger pragmatic inferences. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 1207–1212).

Kravtchenko, E. (2014). Predictability and syntactic production: Evidence from subject omission in Russian. In *Proceedings of the 36th Annual Meeting of the Cognitive Science Society* (pp. 785–790).

PhD Dissertation

Kravtchenko, E. (2022). Integrating pragmatic reasoning in an efficiency-based theory of utterance choice. Grade: *magna cum laude*.

Awards & Honors

National Science Foundation Graduate Research Fellowship Program, Honorable Mention; 2012.

Hobbies

Hiking, rock climbing, {board|role-playing|computer} gaming.