Eliot Seo Shekhtman

🔖 eliotshekhtman.com ✉ eliotshekhtman@gmail.com 🔗 linkedin.com/in/eliot-shekhtman 🎓 Google Scholar

Education

University of Pennsylvania   Sept 2024 - present
Ph.D. in Computer and Information Science   *Philadelphia, PA*

Cornell University   Sept 2022 - May 2024
Master of Science in Computer Science (GPA: 3.968 / 4.3)   *Ithaca, NY*
- **Relevant Coursework:** Machine Learning Theory*, Machine Learning in Feedback Systems*, Deep Generative Models*[†], Computer Vision*, Honors Introduction to Analysis, Game Theory I

Cornell University   Sept 2018 - May 2022
Bachelor of Science in Computer Science (GPA: 3.943 / 4.3)   *Ithaca, NY*
- **Relevant Coursework:** Advanced Topics in Machine Learning*, Learning with Big Messy Data*, The Structure of Information Networks*, Principles of Large-Scale Machine Learning*, Advanced Language Technologies*, Natural Language Processing*, Intro to Computer Vision*[†], Intro to Machine Learning*[†], Intro to Analysis of Algorithms
  Graduate level courses   [†] Teaching assistant

Technical Skills

Languages: Python, Java, R, C++, Shell, OCaml, SQL (MySQL, SQLite), Kotlin, Julia, Dart
Technologies: PyTorch, Tensorflow, Keras, Huggingface, Pandas, Scikit Learn, Git, R Studio, Android SDK
Concepts: Machine Learning, NLP, CV, Diffusion, Transformers

Experience

Goldman Sachs   Sep 2022 - Mar 2025
Machine Learning Engineer   *Remote*
- Creating and developing machine learning algorithms as a contract worker for industry use.
- Productionized my NLP training pipeline, utilizing test-driven development and passing intensive code reviews.

Goldman Sachs   Jun - Aug 2022, Jun - Aug 2021
Engineering Summer Analyst   *New York City, NY*
- Trained BERT-based transformer models to classify text documents, achieving high recall and workload reduction.
- Developed a semi-supervised model to highlight sections of text influential for a classification task.
- Implemented a generalized pipeline for training and evaluating large NLP models for future projects.
- Implemented Hadoop MapReduce jobs to efficiently filter, tokenize, label, and embed ~10 million articles stored in Google Protocol Buffer format.
- Researched and developed LSTM and attention models to predict on a Named Entity Recognition task, achieving high accuracy over ~5 million sentences.

Projects

Cornell University Sustainable Design (CUSD) | *Python, Unsupervised learning, Kernel density estimation*
- Developed an online algorithm to predict room usages and optimize energy consumption, forecasted to reduce unecessary HVAC operation by over 70%.
- Coordinated with two other subteams in 20+ meetings to gather information.
- Led the development of a script to generate accurate dummy data for initial algorithm training.

Publications

Eliot Shekhtman*, Yichen Zhou*, Ingvar Ziemann, Nikolai Matni, and Stephen Tu. *Nearly Instance-Optimal Parameter Recovery from Many Trajectories via Hellinger Localization*. arXiv preprint.

Eliot Shekhtman, and Sarah Dean. *Strategic Usage in a Multi-Learner Setting*. AISTATS 2024.

J. Lovelace, V. Kishore, C. Wan, **Eliot Shekhtman**, and K. Weinberger. *Latent Diffusion for Language Generation*. NeurIPS 2023.

Yuxuan Zhao, Eric Landgrebe, **Eliot Shekhtman**, and Madeline Udell. *Online Missing Value Imputation and Correlation Change Detection for Mixed-type Data via Gaussian Copula*. AAAI 2022.