Alec Stashevsky

alecstashevsky.com

alec@alecstashevsky.com

Applied machine learning scientist and technical leader with a track record of leading teams shipping machine learning / AI solutions at scale. Leader of Fetch's core machine learning team building solutions powering the Fetch app handling over $150 billion in annual gross merchandise volume, an equivalent to the third largest retailer in the United States. Technical expertise in deep learning for document AI, semantic search/entity resolution, ads ranking, recommendations systems, and fraud detection.

PROFESSIONAL EXPERIENCE

Head of Search and Discovery / Fetch (10/2024)-Present

- Leading four teams responsible for in-app search, recommendations, ads ranking, and location services.
- Engineering manager and technical leader for 30+ engineers in discovery group. Built teams and ML-driven search + discovery features from 0 to 1.
- Building core ML infrastructure to support large scale real-time workloads including streaming data aggregation via Apache Kafka and Flink, feature store, GPU cluster management and deployment, and hybrid search infrastructure.

Tech Lead Manager, Core Machine Learning / Fetch (10/2022)-(10/2024)

- Lead a team of 15+ scientists and engineers building world-class ML/AI technology powering the core of the Fetch app. Our systems extract information from over 11 million receipts in real-time every day, and process over $150 billion in gross merchandise volume annually.
- Core contributor and technical leader for Fetch's largest product launch to date with over $200M in annual revenue attributable to our core document AI technology.
- Research, build, and deploy deep neural networks touching computer vision, natural language processing, and graph machine learning.
- Pre-training and fine-tuning of large language models (LLMs), vision encoder-decoders, and graph neural networks.
- Engineering leader supporting AI/ML microservices, edge device deployments, and serving optimizations, including quantization, compilation, pruning, and knowledge distillation.
- Lead partnerships with open-source and academic communities including Stanford University, Hugging Face, PyTorch, PyTorch Geometric, AWS SageMaker, and Streamlit.

Senior Machine Learning Scientist / Fetch **(10/2021)-(10-2022)**

- Research and develop document understanding models including CNNs and Vision Transformers for optical character recognition, large language transformers, and graph neural networks for named entity recognition and key information extraction.
- Use tools like PyTorch, Hugging Face, PyTorch Geometric, scikit-learn, SQL, R, Docker, AWS and Snowflake.

Financial Economist / NERA Economic Consulting **(09/2020)-(09/2021)**

- Look inside the books of some of the largest financial institutions in the world to estimate damages and predict the performance of complex financial instruments responsible for high-profile banking crises, securities fraud, and market-manipulation cases.
- Build probabilistic financial models using Markov chain Monte Carlo methods, random matrix theory, and time-series forecasting.
- Identify, explain, and value litigation involving mortgage-backed securities (RMBS), collateralized debt obligations (CDOs), swaps, and other derivatives underpinning trillions of dollars in assets.
- Provide evidence and expert economic testimony for Fortune 500 companies, SEC, DOJ, and FINRA.

Economist / The Cadmus Group **(11/2019)-(09/2020)**

- Lead the design and evaluation of demand-side management programs, including a $600k+ randomized control trial on smart thermostat direct load-control.
- Forecast electric vehicle adoption, demand elasticity, and electrification for budgeting hundreds of millions of dollars under diverse energy industry clients' management.
- Design difference-in-difference models, demand-elasticity programs, and causal inference mechanisms to provide gold-standard reporting to regulators and operators responsible for most of the United States energy supply.

Data Scientist/ Kyrgies **(08/2017)-(10/2019)**

- Designed marketing mix models, audience creation, and targeted advertising solutions for a bootstrapped e-commerce startup.
- Deployed recommendation systems and probabilistic engagement models (pCTR, pConversion) for targeted advertising that led to 8% reduction in CAC.

PUBLICATIONS AND TALKS

- Speaker, "Optimizing Sentence Transformers for Entity Resolution at Scale." *Convergence*, May 2024. (recording)
- Speaker, "Graphing Groceries: Understanding Receipts with Transformer Hydranets." *Microsoft AI Graph Learning Group*, January 2024. (recording)
- Speaker, "Graph Transformers for Semantic Link Prediction at Scale." *Stanford Graph Learning Workshop*, October 2023. (recording)
- Assessing Strategies to Reduce the Carbon Footprint of the Annual Meeting of the American Academy of Ophthalmology. *JAMA Ophthalmology*, August 2023. doi:10.1001/jamaophthalmol.2023.3516
- Speaker, "Production Graph ML at Fetch." *PyTorch Geometric Town Hall*, January 2023. (recording)
- Teaching to Our Time: a Survey Study of Current Opinions and Didactics About Climate Mental Health Training in US Psychiatry Residency and Fellowship Programs. *Academic Psychiatry*, July 2022. doi:10.1007/s40596-022-01680-7
- Estimation of the Carbon Footprint Associated With Attendees of the American Psychiatric Association Annual Meeting. *JAMA Network Open*, January 2021. doi:10.1001/jamanetworkopen.2020.35641

ORGANIZATIONS

Statistician / Group for the Advancement of Psychiatry, Climate Committee **(06/2020)-Present**

- GAP is a think tank of top psychiatric minds whose thoughtful analysis and recommendations serve to influence and advance modern psychiatric theory and practice. I serve as a resident statistician supporting climate-related research (*see* publications above).

OPEN-SOURCE SOFTWARE DEVELOPMENT

- Author and maintainer of {**blocklength**}, an R package for advanced statistical analyses of time-series and other dependent data with *over 12 thousand downloads*. {blocklength} provides tooling to optimize the block-length parameter of a dependent bootstrap procedure (the block-bootstrap) commonly used in time-series modelling. Available on CRAN and at my GitHub.

EDUCATION

B.A. Mathematics-Economics / Reed College, Portland OR