

LOGAN LAWRENCE

(817) 729-5668 ◊ Amherst, MA

✉ ◊ in ◊ 🌐 ◊ loganlawrence.info

EXPERIENCE

Senior Data Engineer

June 2020 - Current

Saifr (Fidelity Labs)

Boston, MA

- Project lead for GOST Crawl, a proprietary search engine for indexing derogatory web content. To date, GOST Crawl has indexed over 14 billion webpages, totaling in ~ 4 petabytes of uncompressed HTML. It is a highly-scalable, highly-available ETL pipeline and storage system hosted in AWS which houses several machine learning models for tasks such as topic modeling, sentiment prediction, and Named Entity Recognition (NER).
- Research Sponsor for the Clemson University School of Computing. Managed 4 teams of Computer Science seniors to create state-of-the-art methods for NER and entity-centric summarization on long range text used within GOST Crawl. Supervised a paper on the interplay between knowledge distillation and efficient attention patterns (up to 4096 tokens) with respect to model inference costs, which was accepted to EMNLP 2023.
- Science Team lead. Conducted over 24 production-impacting experiments, assisted over 32 experiments, and contributed to the methodology of almost all experiments conducted by FTE Science members. Led the hiring of 5 employees, contributed to the hiring of 5 additional employees, and managed the priorities and task assignments for the Science team.
- Lead developer for a contracted web forum stock market data and sentiment product. Created an ETL pipeline which processes over 100k Reddit, Twitter, and Chan articles daily, detects mentioned stock tickers, predicts market sentiment, and aggregates to a internet-wide market sentiment. The pipeline is hosted in AWS and delivers hourly files via SFTP.
- Previously Giant Oak Inc., acquired in February 2024 by Fidelity Investments.

Graduate Research Assistant

Nov 2020 - August 2022

Prof. Nuno Vasconcelos, SVCL, University of California San Diego

San Diego, CA

- Built an Active Learning method for maximizing model fairness under cost constraints. The method proposes a new acquisition function and post-annotation filtering method which produces breakout performance in annotation cost-efficiency. Working paper available upon request.
- Led weekly research discussions and conducted over 180 experiments totaling in ~ 1.2 years worth of GPU compute time. Aided all other experiments with debugging efforts, coordinated internal team meetings, and structured submitted papers. Built an Active Learning library to interact with Nautilus, a west coast compute cluster available through the NSF Pacific Research Platform.

Undergraduate Research Assistant

Nov 2017 - May 2019

Team ASTRO + Wearables Lab, Rice University

Houston, TX

- Presented an algorithm and drone system for tracking and following a targets movement over live video feed at Rice's Undergraduate Research Symposium. Spoke to over 150 people on the drone system's design, potential applications, the implications of drone tracking technology, and the logistics of pruning and deploying neural networks on low-power devices. PI: Prof. Edward Knightly, Team ASTRO.
- Implemented algorithms for room impulse estimation and robust measurment of user heart rate through a video camera. Used existing signal processing method methods to create working Apple Watch and IOS apps which could function within given time budget and error rate constraints. These implementations were used by subsequent teams in proof-of-concept projects dedicated to non-invasive health screening. PI: Prof. Ashutosh Sabharwal, Wearables Lab.

EDUCATION

Ph.D., Computer Science, University of Massachusetts Amherst

Present

M.S., Electrical and Computer Engineering, University of California San Diego

June 2021

B.S., Computer Science, Rice University

May 2019

B.S., Electrical Engineering, Rice University

May 2019

PUBLICATIONS

Efficient Transformer Knowledge Distillation: A Performance Review. Nathan Brown, Ashton Williamson, Tahj Anderson, Logan Lawrence. [\[EMNLP 2023\]](#)

ALoud: Active Learning of Unbiased Datasets. Logan Lawrence, Shunkai Yu, Debalina Chowdhury, Yichen Jia, Yi Li, Nuno Vasconcelos. [Preprint]

COMPUTER SKILLS

Languages: Python (torch, transformers, cv2), Terraform, Ansible, SQL, L^AT_EX

Environment: Linux, Windows, Neovim (not EMACS), tmux, VSCode

Cloud: AWS (experience with most popular storage, DB, compute, logging, security, and networking services)

Miscellaneous: Jenkins, Rundeck, K8s + Helm, Docker, Elasticsearch, Kibana