Jerry Sun

js2746@cornell.edu | 480-512-2414 | jerrysun14 | jsun14 | jsun14.github.io

Education

Cornell University, Ithaca, New York
M.Eng, B.S., Computer Science, GPA: 3.76, Deans List Aug 2019 - May 2023

- **Coursework**: Advanced ML Systems, Intro to ML (Head TA), Computer Vision, Probability & Stats, Databases, Computer & Operating Systems, Algorithms, Linear Algebra, Multivariable Calculus

Skills

Languages and Packages

- Python, Terraform, Helm, SQL, PyTorch, Tensorflow, Sklearn, XGBoost, Java

Frameworks and Technologies

- GitOps, DevOps, MLOps, Agile, Kubernetes, ArgoCD, Azure/AWS, Mlflow, Databricks

Experience

MunichRE US Life Insurance | *Data Science Intern* Mar - Aug 2022

- Utilized routine GitOps and DevOps practices to create cloud infrastructure (Terraform, Helm, Kubernetes, Azure, ArgoCD) and deployed 2 new products for clients.
- Iterated on dashboard visualizations for clients to create a feedback loop on mortality model training.
- Utilized pointwise mutual information to create custom disease embeddings for anomaly detection through a multi-armed bandit feedback system in identifying areas of interest for Underwriters.
- Generated disease embeddings using a contrastive learning approach with K-Means Clustering and implemented an XGBoost mortality prediction model.
- Developed an Active and Continual Learning framework to be released as open-sourced software.

Amazon | *SDE Intern (AWS)* Jun - Aug 2021

- Developed an internal tool (Python) to automatically collect and store EC2 Nitro hardware system metadata inside a MySQL database, reducing query times from hours to seconds
- Created an internal tool's testing infrastructure by adding unit and integration tests, greatly reducing the frequency at which the service crashes

Cornell Data Science | *Subteam Lead & Onboarding Chair* Jan 2020 - Present

- Managed the creation of data science tutorials, workshops, and all other technical projects
- Onboarded 50+ members through a new system of technical lectures and culture building events
- Lectured for the student-led Intro to Data Science course with a class size of 60+

Projects and Accolades

CoalescenceML (Cornell Data Science) Feb 2022 - Present

- Built an open-source MLOps framework to easily produce industry-grade production ML pipelines by coalescing various MLOps technologies under one umbrella.

1st Place Project X (Cornell Data Science) Sep 2021 - Feb 2022

- Derived insights on the virality of COVID-19 misinformation on Twitter through the use of BERT and other NLP models by building a misinformation-classification framework to combat it.

4th Place Election Prediction Kaggle Competition Dec 2020

- Created SVM, KNN, and Neural Network models to predict 2016 election results with 86% accuracy.