Rohan Gupta

+1 (302) 363-8352 grohan@seas.upenn.edu LinkedIn GitHub

Education

University of Pennsylvania, *Philadelphia, PA* **Sep 2020 – May 2024**

B.S.E. in Computer Science (NETS program) GPA: 3.98

Notable Coursework: Discrete Mathematics (A+), Program Design (A+), Computational Linear Algebra (A+), Data Structures and Algorithms, Advanced Algorithms (A+, 2nd in class), Scalable and Cloud Computing, DevOps (A+), Artificial Intelligence (G) (A+), Graph Neural Networks (G), Algorithmic Game Theory (A+, 2nd in class). **G = Graduate Level*

Technical Skills

Languages: Python, Scala, Java, JavaScript, SQL, Go, LaTeX
Technologies/Frameworks: Git, Django, Flask, ML Python stack, Kubernetes, Docker, Emacs, Vim
General: Backend Development, Distributed Systems, DevOps/MLOps, Data Engineering, Unix

Experience

Stripe, *Seattle, WA* **May 2022 – Present**

Software Engineering Intern

- Develop in pure, functional, and highly concurrent **Scala** to integrate new, in-house **Memcached** cache with ML Feature Computation code, created to replace AWS-hosted **Redis**. Result in completing team's **quarterly OKR**.
- Single handedly architect previously unscoped, large-scale (500+ LoC) optimisations within 2 days leading to **9x end-to-end** latency improvement. Work on in-memory hot key batching after finishing project 3wks early.
- Work closely with distributed caching team to reproducibly test latency and consistency at 20k+ RPS, devise various locking mechanisms to improve write throughput, and proactively submit pull requests to improve their **Java SDK**.

NeuroFlow, *Philadelphia, PA* **June 2021 – Apr 2022**

Data Science/Engineering Associate

- Develop an end-to-end NLP and ML labeling, training and prediction pipeline (**spaCy, scikit-learn, Flask, FastAPI**) to classify patients' risk for severe anxiety/depression using journal entries
- Work part-time during the school year, leading transition of data stack to follow modern MLOps/DevOps (**Docker, K8s, Redis**) practices, with cache/model load optimisation using Redis, horizontal scaling, and automated retraining.
- Use ML (**spaCy, TextBlob, and GBDTs**) to detect a variety of wellbeing metrics from journal entries; productionize all data products using an internal and client-facing API for prediction (**FastAPI**).

Penn Labs, *Philadelphia, PA* **Oct 2020 – Present**

Co-Director/Team Lead/Backend Engineer

- Lead student organisation of 30+ engineers, designers, and business developers to maintain and develop new products (100k+ unique users) for the Penn community.
- Manage projects, recruiting and organise community events (tech talks / social functions); interface with external stakeholders (Penn Admin, Office of Student Affairs, Club Council) to keep operations running.
- Work at all levels of the stack for Penn's official club repository (Link: *Penn Clubs*), including developing the backend API (**Django/REST Framework**), optimising database/cache queries, and managing K8s infrastructure.

Projects

Distributed Machine Learning Pipeline | *dist-ml-pipeline* **Nov 2021**

- Develop a fully model and parameter agnostic Machine Learning infrastructure system (**Docker, Kubernetes, PyTorch**) to deploy and scale ML models with minimal overhead
- Offer a plug-and-play system with inputs as model class and hyparameters at runtime, continuous training using CronJobs, and automated logging to AWS S3 enabling DevOps-adherent ML engineering.
- Integrate support for distributed training workloads using Kubernetes custom resource PyTorchJob (by Google) integrated with CronJobs.

Computational Neuroscience Research | *neurogenesis-research* **June 2021 – Aug 2021**

- Work with *Prof. Vijay Balasubramanian* and David Kersen (MD/PhD student) to study adult neurogenesis in the olfactory bulb and its impact on odor perception.
- Develop computational models in **MATLAB** and **Python** to simulate the integration of new neurons into existing neural network topologies.
- Investigate multiple modes of neuron placement using metrics in **Graph Theory** and **Biostatistics**, and validate through mutual information and decorrelation between odors.