Jeremy Li

Bellevue, WA 98006 · (425) 647 0813 · h.jeremy.li@gmail.com

Summary

I am an accomplished researcher and engineer with extensive experience in fast-paced startup environments, leading and rapidly scaling data/research teams, and interfacing with executive, engineering, and product teams. I have a strong technical background in statistical modeling, machine learning, computational genomics, and software engineering.

Skills

Languages: Python, R, bash, some C++, Rust

Tools: Docker, AWS (EC2, S3, ECR, Batch), git, GNU/Linux, cluster computing

Experience

Consultant, Research & Engineering   12/2024 – present
Stealth biotech startup, USA

- Consulting work for a seed-stage biotech startup, engaged to optimize a class of highly specialized machine learning models for genomic data
- Unifying disparate multi-TB genomic datasets comprising >500k whole genome sequences

Director of Research   02/2022 – 12/2024
Head of Research   10/2021 – 02/2022
Gencove, New York, NY

- Oversaw company-wide ML and R&D strategy, scaling a team from 1 to 7 PhD-level researchers in under 12 months to develop novel computational genomics solutions at scale
- Spearheaded deep learning initiatives (PyTorch) to replace expert-driven models, reducing manual overhead by 30%+ while accelerating analysis pipelines
- Collaborated with product and engineering teams to build proof-of-concept solutions, resulting in multiple $100k-$1M partnerships
- Alongside engineering, designed and maintained genomics pipelines processing petabytes of genomic data in a high-throughput AWS environment, decreasing compute costs by >25%

Senior Research Engineer   03/2021 – 10/2021
Research Scientist (employee # 5)   02/2019 – 02/2021
Gencove, New York, NY

- Developed and implemented novel statistical and computational methods for sparse NGS data, enabling POC demonstrations that helped secure 15+ enterprise deals
- Collaborated with engineering teams to build out the initial production workflows for genomic data, achieving stable parallel processing of multi-TB datasets
- Led massive-scale biobank analyses, designing novel pipelines and methods to handle 100s of thousands of whole-genome sequences to extract biological insights
- Authored three first-author publications and presented at major genomics conferences

Lead Research Scientist (employee # 4)   07/2017 – 02/2019
Genomic Prediction, North Brunswick, NJ

- Led 3-person team to develop and deploy the cloud-based NGS analysis pipeline underpinning the first revenue-generating product of the company

Publications

Published 14 research articles (h-index: 5), presented findings at major genomics conferences

Education

University of Washington for a B.S. in Physics and B.A. in Chemistry   2013 – 2017