Srikar Talluri

214-449-5920 | tallurium.com | srikartalluri@berkeley.edu | srikartalluri | srikartalluri

Education

University of California, Berkeley May 2025
Bachelor Degree in Computer Science Berkeley, CA
Bachelor Degree in Applied Mathematics

- **GPA:** 3.92
- **Selected Coursework:** Algorithms, Database Systems, Operating Systems, Deep Learning, Optimization Models, Computer Graphics, Probability Theory, Discrete Math, Linear Algebra, Real/Complex Analysis, Numerical Analysis
- **Activities:** Math Undergraduate Student Association, Competitive Programming, Berkeley Debate, Pickleball Team

Honors & Awards

Honors: Upsilon Pi Epsilon (CS Honors fraternity), Math Honors Program (graduate classes + thesis), Dean's List
Awards: Google Code Jam (Top 500), International Collegiate Programming Contest/ACM ICPC Top 15 Division 2
Certifications: Postman API Design Fundamentals, AWS Certified Cloud Practitioner

Work Experience

Machine Learning Engineering Intern May 2024 – Aug 2024
Talroo *Austin, TX*

- Fine tuned ensemble transformer models for job requirement extraction in TensorFlow at 0.98 f1, an increase by 0.26 f1
- Supported throughput of 120 tps measured on one GPU by designing end-to-end Triton serving architecture
- Boosted client conversion rate from 4% to 6% by designing a model-serving RESTful API for the core search platform

Data Engineering Intern Jun 2023 – Aug 2023
Vogue Magazine *New York City, NY*

- Processed 60+ petabytes of daily user data by building real-time multi-threaded Spark/Kafka pipelines
- Deployed parallelized stream architecture in Terraform to ingest 2,000,000+ subscriber data, achieving 4x speedup
- Developed service to smart-ship overstock, earning potential of $175k+; won 1st place in intern pitch competition

Software Engineering Intern Oct 2022 – May 2023
SC Electric Automation *Alameda, CA*

- Reduced latency on electric grid chips by 67ms (from 150ms) by building automated communication systems in C++
- Developed production-scale React web application on 200+ GPS devices with TCP/IP and UDP networking protocols

Software Engineering Intern Jun 2022 – Aug 2022
ANB Systems *Houston, TX*

- Built Recognition & Redaction API to secure sensitive data on 500+ documents using OCR detection for client security
- Cut debugging costs by 83% by designing GraphQL API to display extraction status on user dashboards

Selected Projects

ChessFormer | *Python, PyTorch, GPT Transformer Models, Onnx, Lichess API*

- Designed searchless tranformer-based chess agent with 1900+ rating trained on 2,000,000+ board states with pytorch
- Achieved 92% accuracy when fine-tuned to predict expert moves from PGN datasets with integrated move legality
- Served quantized model using Onnx and linked to bot API for open play on Lichess servers

PhySolve | *Rust, Python, PyO3, Rayon, Numpy*

- Designed physics engine in Rust to simulate physics problems: kinematics, n-body problem, fluid dynamics, etc
- Wrote parallelized versions of Range-Kutta & Navier Stokes resulting in 70% speedup and 40x speedup from Python
- Built robust Python wrapper library around Rust core using PyO3 enabling seamless integration with Python scripts

PintOs | *C++, C, x86 assembly, Docker, Bochs*

- Developed operating system kernel in C (without stdlib) that implements preemptive multi-threading, synchronization primitives, advanced priority scheduling, and system calls.

Technical Skills

Areas: Machine Learning, Model Fine-Tuning, Data Engineering, Distributed Systems, Backend Development
Languages: Python, C/C++, Java, JavaScript, TypeScript, Scala, SQL, Rust
Tech Stacks/Tools: AWS, Google Cloud, Docker, Django, Next.js/React, REST, gRPC, PostgreSQL
Libraries: Pytorch, Tensorflow, Pandas, SKLearn, MatPlotLib, OpenCV, Selenium, OpenGL, GLSL