Mohid Tanveer

mohid.tanveer011@gmail.com | (731) 803-9164 | LinkedIn | Portfolio

Software engineer building scalable data and ML-driven systems, with experience deploying production services. Owned end-to-end pipelines/platforms spanning distributed data processing, system performance optimization, and applied ML.

SKILLS

Programming Languages: Python, C, Go, Java, JavaScript, TypeScript, SQL, R, Bash, HTML/CSS
Tools/Frameworks: Git, Node.js, React, REST APIs, Jira, AWS, Docker, Kubernetes, Linux (RHEL), Virtual Machines, PostgreSQL, AI Agents, GitHub Copilot, RAG
ML and Data Science: MapReduce, TensorFlow, PyTorch, LangChain, PySpark, NumPy, SciPy, Pandas, Matplotlib, Scikit-learn

WORK EXPERIENCE

St. Jude Children's Research Hospital **September 2024 - May 2025**
Student Machine Learning Engineer, Research Systems *Memphis, TN*

- Owned the end-to-end design and deployment of a production RAG-based local LLM question answering system for **~300 internal users**, architecting the full inference pipeline from query ingestion and knowledge-base retrieval through response generation, reducing ServiceDesk ticket volume by **80% over a one-month evaluation.**
- Owned monitoring and production reliability for the LLM service by defining SLIs and SLOs, with automated alerting and continuous benchmarking over **600 historical tickets** to guide latency, deflection, and failure-mode tuning in production.
- Designed and enforced data access controls and audit logging for the LLM system to meet internal security and research data governance requirements, ensuring sensitive datasets were never exposed in generated responses.
- Built and productionized end-to-end ML pipelines for **two research projects**, processing **multi-terabyte research datasets** into evaluated models and deployable workflows supporting ongoing computational biology research.

Applied Data Science Intern **May 2024 - August 2024**

- Built an automated network measurement and monitoring pipeline by benchmarking 10 GB data transfers across campus infrastructure, detecting a degraded network switch, and improving cluster-backed workload throughput by **~15%**.
- Owned a production microservice to asynchronously convert **bioimaging TIFF volumes ranging from multi-gigabyte to multi-terabyte**, eliminating manual workflows and improving conversion throughput by **~30%** via Python-based conversion wrappers and Slurm-orchestrated job execution and resource management.
- Built centralized logging, monitoring dashboards, and utilization reporting across Windows and Linux workstations, enabling data-driven resource reallocation and **reducing unused software licenses and underutilized systems**.

PROJECTS

Tubify - Typescript, Python, React.js, SQL, Librosa, Numpy **February 2024 - April 2025**
Open-Source Personal Project converted to Senior Capstone Project

- Led the design and development of a full-stack social music discovery platform used by **~150 students**, delivering core product features including playlist management, friend-based recommendations, and daily personalized music feeds.
- Engineered scalable backend data pipelines and GPU-accelerated audio analysis workflows to process and persist **user listening data and ~20,000 tracks** while managing server load and minimizing user-facing latency.
- Built and evaluated a batch-oriented recommendation system that generated daily recommendations at scale, balancing relevance and diversity through multiple ranking strategies.

ScreenSense - Python, PyTorch, scikit-learn, Numpy, Scipy, Pandas **September 2025 - December 2025**
Course Project

- Built a hybrid computer vision system to detect screen re-photography using a **500-image** mixed-authenticity dataset.
- Achieved **perfect class separation** on course benchmarks by fusing EXIF-based metadata classifiers with pixel-level signal analysis, outperforming baseline models and meeting inference-time constraints for real-world verification use cases.

EDUCATION

University of California, San Diego **September 2025 - December 2026**
M.S. Computer Science & Engineering, Overall GPA 3.9/4.0 *San Diego, CA*

- **Specialization in AI/ML with relevant courses in** Recommender Systems and Web Mining, Probabilistic Reasoning & Learning, ML: Learning Algorithms, Algorithm Design and Analysis, Unsupervised Learning, Computer Security.

Rhodes College **August 2021 - May 2025**
B.S. Computer Science, Overall GPA 3.91/4.0 *Memphis, TN*

- **Relevant courses**: Machine Learning, Artificial Intelligence, Mathematical Statistics, Data Structures and Algorithms