

# Mohid Tanveer

[mohid.tanveer011@gmail.com](mailto:mohid.tanveer011@gmail.com) | (731) 803-9164 | [LinkedIn](#) | [Portfolio](#)

## EDUCATION

- University of California, San Diego** September 2025 - December 2026  
San Diego, CA  
*M.S. Computer Science & Engineering, Overall GPA 3.9/4.0*
- **Relevant courses:** Recommender Systems and Web Mining, Probabilistic Reasoning & Learning, ML: Learning Algorithms, Algorithm Design and Analysis, Personal Genomics and Bioinformatics, Computer Security
- Rhodes College** August 2021 - May 2025  
Memphis, TN  
*B.S. Computer Science, Overall GPA 3.91/4.0*
- **Relevant courses:** Machine Learning, Artificial Intelligence, Mathematical Statistics, Statistical Analysis, Multivariable Calculus, Advanced Algorithms, Theory of Computation, Systems Programming and Computer Organization

## EXPERIENCE

- St. Jude Children's Research Hospital** September 2024 - May 2025  
Memphis, TN  
*Student Machine Learning Engineer, Research Systems*
- Reduced ServiceDesk ticket volume by 80% by analyzing support ticket data, deploying a production LLM-powered response system, and evaluating retrieval and response quality across real user queries.
  - Designed a Retrieval-Augmented Generation (RAG) inference pipeline utilizing embedding-based vector search and semantic retrieval, local LLM inference, and knowledge graphs; supporting secure model inference, improving answer precision, and reducing hallucinations on internal knowledge benchmarks.
  - Enabled active research projects by designing and integrating machine learning pipelines, conducting data preprocessing, feature engineering, model training, evaluation, and end-to-end pipeline integration in collaboration with domain researchers.
- Applied Data Science Intern** May 2024 - August 2024
- Collected and analyzed large-scale storage and network telemetry data to identify performance degradations, building automated monitoring, anomaly detection, and data visualization pipelines.
  - Streamlined researcher workflows by automating and scaling the preprocessing of large bioimaging datasets by building a backend service integrated with an internal imaging platform and Slurm-based batch processing.
  - Improved resource allocation decisions by implementing Splunk-based logging, monitoring, and observability for Windows and Linux systems, analyzing workstation utilization, and generating actionable system metrics and visualizations.

- Rhodes College Department of Computer Science** August 2023 - May 2025  
Memphis, TN  
*Peer Tutor*
- Tutored and mentored ~200 introductory and intermediate Computer Science students through twice-weekly sessions covering data structures, algorithms, object-oriented programming, memory management, pointers, and system-level design.

## PROJECTS

- Tubify** February 2025 - April 2025  
[GitHub Repository](#)  
*Senior Capstone Project*
- Designed and developed a full-stack web application enabling social music discovery by integrating Spotify and YouTube APIs to analyze user and friend listening histories, support shared playlists, and surface personalized recommendations.
  - Engineered a scalable backend and data pipeline to batch-process, store, and query audio features and user interaction data using a PostgreSQL-backed architecture, supporting efficient retrieval and long-term persistence.
  - Built a personalized music recommender system using collaborative and content-based filtering, taste clustering, and MMR reranking, leveraging audio feature extraction and CUDA-accelerated FFT analysis to improve recommendation quality.
  - Evaluated recommendation quality using offline ranking metrics and user-level preference similarity.

- ScreenSense** December 2025  
[GitHub Repository](#)  
*Course Project*
- Built a hybrid computer vision detection system utilizing Pytorch and signal processing to identify screen re-photography at verification time by combining EXIF metadata analysis, wavelet-based CNN models, and sub-pixel feature extraction.
  - Achieved state-of-the-art detection performance by fusing a metadata-driven random forest prior with learned pixel-level signals, demonstrating robustness across diverse image conditions.

## SKILLS

**Programming Languages:** Python, C, C++, Java, JavaScript, TypeScript, SQL, R, Bash, HTML/CSS

**Tools and Platforms:** Git, Node.js, React, REST APIs, Docker, Kubernetes, Linux (RHEL), Virtual Machines, PostgreSQL

**ML and Data Science:** Power BI, TensorFlow, PyTorch, Apache Spark, PySpark, NumPy, SciPy, Pandas, Matplotlib, Scikit-learn

**Languages:** Fluent in English, Urdu, and Punjabi; Conversational Proficiency in Spanish, Hindi | **Honors:** Eagle Scout