

Shahzaib Saqib Warraich

✉ warraich@usc.edu

🌐 <https://shahzaib-s-warraich.github.io>

🐦 [@Shahzaib598](https://twitter.com/Shahzaib598)

☁️ [@Shahzaib598](https://github.com/Shahzaib598)

🔄 [shahzaib-s-warraich](https://github.com/shahzaib-s-warraich)

🌐 [shahzaib-saqib-warraich](https://github.com/shahzaib-saqib-warraich)

I am interested in AI interpretability and evaluation: understanding what language models truly know and building rigorous methods to measure it. My work probes internal representations, uncertainty, and predictable failure modes, bridging interpretive insight with principled benchmarking.

EDUCATION

University of Southern California
M.S. in Applied Data Science

Jan 2024 - Dec 2025

National University of Sciences and Technology
B.E. in Electrical Engineering and Computer Science

Sep 2017 - June 2021

PUBLICATIONS

Sample, Align, Synthesize: Graph-Based Response Synthesis with ConGrs ([URL](#))
S Ghosh, [SS Warraich](#), D Tarsadiya, G Yauney, S Swayamdipta COLM 2026

How Reliable is Language Model Micro-Benchmarking? ([URL](#))
G Yauney, [SS Warraich](#), S Swayamdipta ICLR 2026 (Oral)

RESEARCH EXPERIENCE

Sample, Align, Synthesize: Graph-Based Response Synthesis with ConGrs *Supervisor: Prof. Swabha Swayamdipta — COLM 2026*

- Developed and validated application of consensus decoding framework to factuality-critical tasks, conducting systematic evaluations across long-form biography generation and PopQA benchmarks with five independent replications per model, establishing robust baseline performance metrics and characterizing the factuality-information tradeoff
- Designed and implemented high-throughput batch processing infrastructure for FActScore evaluation, engineering performance optimizations including response caching that resolved critical processing bottlenecks and improved evaluation efficiency
- Architected comprehensive ablation studies investigating scaling dynamics of sample size during alignment and response synthesis phases, along with temperature-based sensitivity analysis, providing empirical insights into framework robustness
- Extended framework to refusal-aware settings across three challenging domains (Numerical False Presuppositions, Scientific Attributions, Historical Events), integrating sophisticated LLM-as-a-judge abstention detection for more nuanced assessment

How Reliable is Language Model Micro-Benchmarking? *Supervisor: Prof. Swabha Swayamdipta — ICLR 2026 (Oral)*

- Proposed and implemented confidence-based stratified sampling methodology for micro-benchmark selection in meta-evaluation frameworks, adapting principled statistical techniques from computer vision [literature](#) to language model evaluation

- Developed comprehensive evaluation infrastructure integrating multiple micro-benchmark selection approaches including stratified sampling, TinyBenchmarks (P-IRT and GP-IRT), and task reduction [baselines](#), enabling systematic comparison of benchmark selection strategies
- Maintained and extended research codebase, implementing modular components for reproducible meta-evaluation experiments across diverse benchmark selection methodologies

INDUSTRY EXPERIENCE

Turon AI, Co-Founder — Solution Architecture, Technical Sales Jan 2026 - Present
Applied AI R&D and Production Engineering

- Own solution architecture and technical pre-sales: lead discovery, scope system designs, and convert them into signed engagements and delivery plans.
- Architect and ship production agentic systems and AI infrastructure for enterprise clients.
- Built and run the firm's commercial and operating layer, from entity structure and GTM to engineering infrastructure and technical hiring.

Paramount, Data Scientist Intern May 2025 - Oct 2025
Entertainment Analytics & AI

- Enhanced content performance forecasting accuracy by up to **20%** across critical business KPIs through strategic integration of multi-platform social media engagement signals into predictive modeling frameworks
- Architected a production-grade GenAI explainability system leveraging ensemble LLM-as-a-judge methodology with majority voting consensus, enabling granular title-level performance attribution and democratizing advanced ML insights for non-technical stakeholders across the organization

Retrocausal, Research Engineer II (AI) Dec 2022 - Dec 2023
Computer Vision & Manufacturing Intelligence

- Designed and deployed precision assembly verification solutions achieving **98%** detection accuracy for Fortune 500 manufacturing clients including Schneider Electric, Carrier, and Honda through state-of-the-art object detection and activity recognition algorithms ([Assembly Copilot](#))
- Architected an automated model training infrastructure incorporating advanced cut-paste-mix augmentation techniques to generate photorealistic synthetic datasets, reducing manual annotation overhead by **80%** while maintaining rigorous performance standards
- Conceptualized and led development of a computer vision-enabled ergonomics risk assessment module for Kaizen Copilot, establishing a differentiated product capability addressing workplace safety optimization and regulatory compliance ([Kaizen Copilot](#))

Adlytic AI, AI Engineer June 2021 - Dec 2022
Edge AI & Recommender Systems

- Spearheaded comprehensive model optimization initiatives, implementing knowledge distillation and NVIDIA TensorRT optimization to achieve **66%** reduction in edge inference latency while maintaining performance within **1.5%** of baseline accuracy across production deployments
- Architected and operationalized a scalable cloud-based recommender system achieving **97%** prediction accuracy, driving a **3x** increase in product conversion rates and generating measurable revenue impact through enhanced customer engagement

AWARDS & HONORS

- USC Andrew J. Viterbi Scholarship 2024 - 2025
- William J. Fulbright Scholarship 2024 - 2025

TEACHING

Course Producer, *University of Southern California – Viterbi School of Engineering*

- CSCI 544: Applied Natural Language Processing Aug 2024 - Dec 2024
- CSCI 544: Applied Natural Language Processing Jan 2025 - May 2025

LEADERSHIP

Viterbi Graduate Student Association (VGSA), Vice President May 2025 - Present

- Led student advocacy initiatives representing 6,000 graduate students as primary liaison between student body and university administration
- Managed strategic partnerships with internal stakeholders and external organizations to enhance graduate student resources and professional opportunities
- Delivered monthly progress reports to Viterbi School Dean, communicating student concerns and actionable recommendations

SKILLS & INTERESTS

Skills

- Languages: English, French
- Programming Languages: Python, C/C++, SQL
- Deep Learning: PyTorch, Hugging Face, vLLM
- Machine Learning: Scikit-Learn, OpenCV
- Data Analysis: Pandas, Dask, PySpark
- Data Visualization: Matplotlib, Seaborn, Plotly
- Cloud: AWS Sagemaker, GCP Vertex AI
- Others: Docker, Kubernetes, MongoDB, Git, LaTeX, Figma

Interests

- AI Ethics and Responsibility
- Hackathons
- User Experience (UX) Design