

ALI AHMAD QURESHI

ali.ahmad@ciitwah.edu.pk | +92 310 5053646 | linkedin.com/in/ali-ahmad-qureshi
Portfolio: ali-ahmad-qureshi.pages.dev | github.com/AliAhmad1997 | ORCID: 0009-0005-9130-605X

RESEARCH PROFILE

MS Computer Science researcher (CGPA 3.91/4.00; expected April 2027) with 7+ years of engineering experience across LLM/RAG systems, embedded software, C/C++, FPGA/HLS, and full-stack development. Current research focuses on reliable retrieval and memory for LLM agents. **Research interests:** Large language models; information retrieval and RAG; agentic and adaptive AI; reliable/trustworthy ML systems; long-horizon memory; ML systems and efficient AI.

EDUCATION

COMSATS University Islamabad Expected Apr 2027
MS in Computer Science | CGPA: 3.91/4.00
Thesis: EA-MEM: Reliable Memory for LLM Agents | Advisors: Dr. Kashif Ayyub, Dr. Ehsan Ullah Munir
NUST College of Electrical & Mechanical Engineering (CEME), Rawalpindi Jul 2019
BE in Computer Engineering | CGPA: 3.33/4.00 | Rector's Gold Medal (Final-Year Project)

PUBLICATIONS & MANUSCRIPTS

Syed, S., **Qureshi, A. A.**, Ahmad, S. G., Ayyub, K., Munir, E. U., & Khan, I. "A Priority-Aware Load-Normalized VM Scheduling Algorithm for Efficient Cloud Task Allocation." *SN Computer Science* - submitted, under review (June 2026).
Qureshi, A. A., Ahmad, S. G., Ayyub, K., & Munir, E. U. "EA-MEM: Failure-Tolerant Retrieval and Maintenance for LLM Agent Memory." *Complex & Intelligent Systems* (Springer Nature), manuscript submitted (October 2026).

RESEARCH & APPLIED AI EXPERIENCE

MS Thesis - EA-MEM: Reliable Memory for LLM Agents (*COMSATS, ongoing*)

- Extends A-MEM (NeurIPS 2025) with failure-tolerant metadata handling, hybrid dense + BM25 retrieval with graph-aware re-ranking, importance-based forgetting, consolidation with link inheritance, and targeted fault-injection/evaluation pipelines.
- Evaluation: LoCoMo F1 0.291 -> 0.344 on 159 informational questions (improved 9/10 conversations; $p = 0.004$); identifier-focused retrieval nDCG@10 0.896 -> 0.982; metadata-failure nDCG@10 0.503 -> 0.745.

Applied RAG System - On-Premises Knowledge-Base Assistant (*NECOP, 2022 - Present*)

- Built and deployed a locally hosted RAG application over organizational documentation using SentenceTransformers, persistent Chroma retrieval, FastAPI, local LLM inference, and an Electron desktop client for privacy-sensitive, cloud-independent deployment.
- Implemented multi-format ingestion, incremental re-indexing, retrieval/context construction, streaming chat, and packaged local deployment; the system's retrieval and grounding challenges directly motivated later thesis research.

PROFESSIONAL EXPERIENCE

National Electronics Complex of Pakistan (NECOP) Jan 2022 - Present
Assistant Manager (Tech) - Cadence Stratus HLS -> Verilog RTL; developed GENYSYS, automating FPGA SoC configuration, constraints/address generation, Vivado builds, bitstream generation, monitoring, and artifact delivery.

Trivor Software Pvt. Ltd. Jul 2021 - Jan 2022
Software Engineer (C++) - Desktop applications for Bentley Systems.

Avanza Solutions Pvt. Ltd. Jan 2021 - Jun 2021
Software Engineer - E-banking application development and transaction verification.

Simplicity Technologies Pvt. Ltd. Nov 2019 - Sep 2020
Embedded Systems Engineer - NVIDIA Jetson AI applications; web/edge-device communication.

Bentley Systems Pvt. Ltd. Jul 2019 - Oct 2019
QA Automation Intern - Automated GUI-based testing and repeatable software validation.

SELECTED PRIOR PROJECT

FPGA-Based Sign Language Recognition - NUST Final-Year Team Project

- Designed and implemented a resource-aware sign-language-recognition CNN on Xilinx Zynq-7000/ZedBoard, using Python/Keras for training and C/C++, Vivado HLS, and custom Verilog RTL for hardware implementation; integrated FPGA IP, validated the design on hardware, and optimized it around device resource constraints. Rector's Gold Medal; open-sourced with archival Zenodo DOI.

TECHNICAL SKILLS

AI/ML & Retrieval: Transformers, SentenceTransformers, BM25, Chroma, FAISS/Qdrant, TensorFlow, scikit-learn
Software & Data: Python, C/C++, C#, Java, Embedded C, FastAPI, Angular/Node.js, MySQL/SQLite
Hardware & Systems: Verilog, Vivado, Cadence Stratus HLS, FPGA/SoC workflows, Linux

HONORS & AWARDS

Rector's Gold Medal, NUST (Final-Year Project) | Merit Award - IoT & Digital Systems, COMPEC'19