ABDULLAH HASHMAT

AI/ML Researcher • LLMs & Agentic Systems • NLP • AI4Science
Portfolio | abdullah.hashmat02@gmail.com | LinkedIn | GitHub | +92 3035551871

EDUCATION

BS Computer Science | Lahore University of Management Sciences (LUMS) ***Aug 2021 – May 2025***
Relevant Coursework: CS6304 Advanced Machine Learning*, CS5302* Speech Processing with Generative AI*, CS6303* Large Language Models*, CS331* Artificial Intelligence*, CS487* Cloud Development*, CS437* Deep Learning

SKILLS

Languages: Python, C/C++/C#, Java, SQL, JavaScript, MATLAB, Haskell, LaTeX
ML & LLMs: PyTorch, Hugging Face Transformers, LLM Fine-Tuning, Knowledge Distillation, Agentic AI/Tool Calling, LLM Evaluation & Benchmarking, NLP, Deep Learning, RDKit
MLOps & Data: AWS (Lambda, S3, ECS/Fargate, DynamoDB), Docker, Serverless, Model Deployment & Inference, ETL, Web Scraping, PostgreSQL, MERN, React

EXPERIENCE

Research Fellow | Max Planck Institute for Software Systems (MPI-SWS) | Kaiserslautern, DE ***Apr 2026 – Present***
Supervised by Dr. Abhilasha Ravichander
- Built an LLM-based agentic framework with dynamic retrieval and experience memory that learns from feedback.
- Connected RDKit, docking, and QSAR tools to the LLM in a closed loop, improving validity, diversity, and binding affinity.

NLP Researcher | Center for Speech and Language Technologies (CSaLT), LUMS | Lahore, PK *(Remote)* ***Sep 2024 – Present***
Supervised by Dr. Agha Ali Raza & Dr. Usman Naseem
- Led a research team building culturally aware, self-evolving LLM agents to improve reasoning and tool use and reduce Western bias.
- Built end-to-end LLM evaluation pipelines and datasets (PakBBQ, Pak3H, PakDecept, IRTIQA); first-author EMNLP Main (2025, 2026).

Lead Researcher | German Research Center for Artificial Intelligence (DFKI) & RPTU | Kaiserslautern, DE ***Jul 2025 – Mar 2026***
Supervised by Dr. Nabeel Asim
- Built a multi-granular transformer (atom-, functional-group-, branch-level) for SMILES-based molecular property prediction.
- Pre-trained BERT, RoBERTa, DeBERTa, and state space models; distilled into one compact student via multi-teacher KD (MLM + FitNet, dynamic gating), achieving state-of-the-art on all 22 ADMET tasks.
- Built a 500M+ record ETL pipeline (large-scale web scraping, 150+ properties) feeding distributed multi-GPU training.

MLOps Engineer | Darcheeni (AI Healthcare Startup, LUMS) | Lahore, PK ***May 2024 – May 2025***
- Led backend for a production AI clinical triage assistant integrating ASR, RL-based decision models, and LLM response generation.
- Boosted ASR via fine-tuning + LLM post-processing; deployed autoscaling serverless inference on AWS (Lambda, DynamoDB, Auth0).

Applied ML Researcher | Center of Urban Informatics, Technology and Policy (CITI) | Lahore, PK ***Sep 2024 – Mar 2025***
Supervised by Dr. Muhammad Tahir
- Designed confidence-aware multi-teacher KD (entropy-based weighting) to cut texture/shape bias and improve OOD generalization.
- Balanced cue reliance (shape 0.49 / texture 0.51), beating single-teacher and naive KD on ImageNet-1k and Stylized-ImageNet.

Sales Performance Intern | Delivery Hero (Foodpanda) | Lahore, PK ***Jun 2024 – Aug 2024***
- Cut QC rejection rate by 15% via large-scale EDA on vendor onboarding and QC data.
- Built an agentic Python workflow (Google APIs) automating data pulls, invoicing, and WhatsApp distribution; cut process time 35%.

Data Science Intern | Pakistan Television Network (PTV) | Lahore, PK ***Jun 2023 – Jul 2023***
- Cut headline redundancy by 20% with ML text-similarity models, built a content-analysis model that lifted viewer engagement by 10%.

PUBLICATIONS

- **PakBBQ: Culturally Adapted Bias Benchmark for QA** — EMNLP 2025 Main Conference Link
- **Pak3H: Evaluating the Cost of Cultural Mismatch in LLM Alignment** — EMNLP 2026 Main Conference Link
- **IRTIQA: Evolving in English, Evaluated in Urdu — Does Agent Self-Evolution Transfer Across Languages?**— *(Committed)* Link
- **PakDecept: Benchmarking LLM Deception in Urdu** – (Under review) Link
- **Rethinking the Architecture of Chemical Language Models: ADMET Property Prediction with State Space Models** – (Under peer blinded review)
- **MolScientist: Agentic Inverse Molecular Design With Scientific Experience** – (Under peer blinded review)

SELECTED PROJECTS

AI-Driven Stock Insights Platform (AWS + Claude API) Link – Full-stack real-time platform (Lambda, S3, ECS Fargate, Neon Postgres) with an LLM chatbot for investment queries; <200ms latency, serverless autoscaling.
AI for All – Evaluating Transformer and LLM Models in Low Resourced Settings Link – Benchmarked open-source LLMs on Urdu, Sindhi, and Pashto; jailbreak prompts reached up to 40% success on Llama-3B.
Bias Mitigation in Deep Learning Vision Models Link – Multi-teacher KD (ViT + VGG16) achieving 12% bias reduction on cue-conflict datasets under distribution shift.
Agentic Data Science Chatbot Link – Fine-tuned Llama 3 agent for automated EDA, statistics, and cleaning across 10+ file formats; deployed on Hugging Face (Gradio).

TEACHING EXPERIENCE

AI600 Machine Learning (MS AI) | CS5302 Foundations of Generative AI | CS535 Machine Learning | CS100 Computational Problem Solving | LUMS

LEADERSHIP

TechStars Startup Weekend Lahore — Convenor ***Jul 2022 – Mar 2023***
- Led a team of 60 with National Incubation Center & Google for Startups; 40+ startups, 150+ attendees, 800K+ PKR sponsorship.

SPADES (PSIFI) — Assistant Director, Socials & Operations ***Sept 2022 – May 2023***
- Led organisation of PSIFI's flagship event, coordinating social functions and concerts for 1,400+ attendees