Abdullah Hashmat

abdullah.hashmat02@gmail.com • linkedin.com/in/abdullah-hashmat • github.com/Hashmat02 • +92 3035551871

EDUCATION

Lahore University of Management Science - *BS Computer Science* *Aug 2021-May 2025*

Relevant Coursework: *CS6304* Advanced Machine Learning, CS*5302* Speech Processing with Generative AI, CS*6303* Large Language Models, *CS331* Artificial Intelligence, *CS487* Cloud Development, *CS437* Deep Learning, *CS331* Principles and Techniques of Data Science

ACADEMIC EXPERIENCE

Research Assistant – *German Research Center for Artificial Intelligence (DFKI) - RPTU* *July 2025– Present*

- Research Assistant at RPTU, in affiliation with DKFI - Germany, focusing on prompt engineering and contrastive learning strategies using LLMs for property prediction of molecular compounds.
- Conducted an extensive literature review to develop a multimodal, multi-agent contrastive learning framework leveraging large language models (LLMs) for molecular property prediction.

Teaching Assistant – *CS100 Computational Problem Solving* *Sept 2023 – Dec 2023*

- Assisted undergraduate students in Python programming, algorithm design, and debugging foundational CS problems

Research Assistant – *Center for Speech and Language Technologies (CSaLT) – LUMS* *Sept 2023 – May 2025*

- Researched **10+** Large Language Models (LLMs) for low-resource languages, focusing on vulnerabilities to bias and jail breaking, while optimising speech processing models for Urdu NLP tasks
- Collaborated on language processing models, conducted literature reviews, and contributed to **4** research papers

Research Assistant – *AI in Healthcare Initiative (AIHI) – LUMS* *May 2024 – May 2025*

- Led backend integration for Sehat Sanjha, an AI speech to text assistant supporting real time physician interactions and patient triage, integrated RL based ML models for decision support and dynamic response generation
- Collaborated as web developer on an AI healthcare startup to develop Sehat Sanjha, an AI-powered physician assistant app aimed at making healthcare more accessible
- Facilitated backend development using React, Redux, and Lambda functions, and integrated **100+** refined test cases
- Revamped the database schema to support reinforcement learning based model refinement, enabling more efficient feedback loops from physician patient interactions.

Teaching Assistant – *CS535 Machine Learning* *Sept 2024 – Dec 2024*

- Supported instruction and grading for core ML concepts including supervised learning, model evaluation, and optimization

Teaching Assistant – *CS5302 Foundations of Generative AI* *Jan 2025 – May 2025*

- Facilitated sessions on transformer models, diffusion techniques, and prompt engineering in GenAI systems

Teaching Assistant – *AI600 Machine Learning (MS in AI)* *Jan 2025 – May 2025*

- Taught and mentored graduate students in advanced machine learning topics, including theory, implementation, and research applications

RESEARCH PROJECTS

Confidence Aware Multi-Teacher Distillation for OOD Generalization Link *Sept 2024 – Mar 2025*

- **Role:** Primary Researcher Supervisor Dr. Muhammad Tahir (LUMS)
- **Objective:** Improve out of distribution (OOD) generalization by mitigating cue specific biases (texture/shape) in image classification through a knowledge distillation framework.
- **Methodology:**
  - Employed entropy-based confidence weighting to ensemble logits from shape- and texture-biased teacher models (ViT, VGG16).
  - Used ViT-Small as a student, learning via KL divergence loss against ensembled logits weighted by teacher confidence.
  - Propagated teacher biases (shape-texture) using a weighted dynamic method to produce a much balanced student model which is OOD generalizable.
- **Results:** Evaluated shape-texture biases of different models across various architectures, specifically, VGG16/11, ViT-Base/Small, ResNet50, DenseNet. Compared single teacher, equal weighted multi teacher, and proposed confidence aware distillation frameworks on ImageNet-1k and Animals10.
  - Achieved balanced bias (Shape: 0.49, Texture: 0.51) using confidence awake KD), improving performance on Stylized ImageNet. and Canny Edge datasets, outperforming single-teacher and naive ensemble baselines.
  - In base pretrained (ImageNet-1k) models VGG-16 showed a strong texture bias (81%), while ViT showed strong shape bias (59%).
  - Confidence aware KD outperformed all single teacher and multi-teacher KD frameworks in terms of bias mitigation. However, ViT to ViT KD achieved higher accuracy (67.45%) compared to confidence aware KD (64.09%), highlighting the need for potential improvements in the loss function, such as incorporating structural knowledge distillation.

PakBBQ: A Culturally Adapted Bias Benchmark for QA Link (Pending EMNLP Decision) *Sept 2024 – May 2025*

- **Role:** Primary Researcher. Supervisor Dr. Agha Ali Raza (LUMS)
- **Objective:** To design and evaluate a culturally contextualized benchmark (PakBBQ) for measuring social bias in LLMs within the Pakistani socio linguistic and regional landscape, addressing the limitations of Western centric datasets like BBQ for evaluating biases in LLMs.
- **Methodology:**
  - Constructed 17,180 QA pairs across 214 templates in English and Urdu, covering 8 sociocultural bias categories (Age, Disability Status, Language Formality, Gender Identity, Physical Appearance, Regional, Religion, and

Socioeconomic Status (SES)).
    - Adapted BBQ using four strategies: Direct Transfer (DT), Target Modification (TM), Newly Added (NA) and Sample Removed (SR) templates rooted in Pakistani contexts
    - Benchmarked 6 multilingual LLMs (GPT 4.1, GPT 4.1mini GPT 4.1nano, Gemini 2.0 flash, Gemini 2.0 flash - lite, DeepSeek-V3) on 17,180 QA pairs across Urdu and English using three template types: Directly Transferred (DT), Target Modified (TM), and Newly Added (NA), under zero-shot conditions with cyclic prompting and majority voting across ambiguous and disambiguated contexts.
- **Results:**
    - Gemini-2.0 Flash Lite achieved the highest accuracy in English (88%), while Gemini-2.0 Flash led in Urdu (81%).
    - Models consistently performed worse on NA templates (Urdu accuracy range 50–68%) due to the challenge of capturing Pakistan specific socio-cultural biases.
    - Disambiguated prompts improved model accuracy by ~12 percentage points on average (e.g., GPT4.1-Mini improved from 64% to 87% in Urdu).
    - Negatively framed questions reduced stereotypical responses, outperforming nonnegative formulations across both languages (e.g., GPT4.1-Nano English: 83% vs 78%).
    - Bias score analysis revealed a) Stronger counter-bias behavior in Urdu vs. English, particularly for categories like Religion and Language Formality, b) Gemini models scored -1 (strong counter-bias) in all disambiguated categories, indicating effective bias mitigation.
    - Identified cross linguistic disparities, Urdu accuracy lagged English by 7–17 percentage points, attributed to lower training data representation and higher prompt sensitivity (Urdu $\sigma \approx 0.11$ vs. English $\sigma \approx 0.07$)

PERSONAL/COURSE PROJECTS

AI-Driven Stock Insights Platform with AWS & Claude Chatbot Link *Sept 2024 – Dec 2024*

AI-powered stock insights platform with AWS backend and Claude chatbot for real time stock market analysis

- Developed and deployed an AWS-based stock data portal integrating S3, Lambda, ECS, and Neon Postgres for automated stock report processing and real-time trend analysis
- Implemented an AI powered chatbot using Claude (LLM) and custom prompt engineering to enable natural language queries, providing users with contextualized stock trends, summaries, and investment insights
- Engineered a cloud native architecture capable of scaling to millions of users, leveraging AWS auto scaling and serverless functions to maintain sub **200ms** response latency for real time stock insights

AI for All – Evaluating Transformer and LLM Models in Low Resourced Settings Link *Sept 2024 – Dec 2024*

Evaluated LLM robustness and safety in low-resource languages via QA and adversarial testing

- Benchmarked XLM-RoBERTa and LLaMA 3B models on QA tasks across English, Urdu, Sindhi, and Pashto, revealing strong performance for Urdu **(EM: 0.72, F1: 0.72)** but significant drops in Sindhi and Pashto (e.g., **Pashto EM: 0.28, F1: 0.46**)
- Demonstrated model fragility by applying adversarial perturbations, leading to near zero performance in Sindhi and Pashto, highlighting poor robustness in low resource settings
- Conducted jailbreaking attacks on LLaMA 3B, achieving pass rates up to **68.5%** (English) and **40%+** in low resource languages, with higher severity (avg 4.3) and lower safety consistency in Pashto and Sindhi

End-End Data Science Chatbot using Generative AI Link *Jan 2024 – May 2024*

AI chatbot for EDA, data cleaning, and multi-format analysis

- Built an AI Data Science chatbot, aimed to perform intuitive statistical analysis, EDA and data cleaning, enhancing accessibility for users with limited Data Science knowledge
- Fined tuned GPT 3.5 to analyse data in **10+** supported formats, deployed on Hugging Face using Gradio
- Implemented custom prompt templates and validation logic to ensure accurate, context-aware responses, reducing hallucination in data interpretation tasks by over **30%** during testing

GenVidea – Ai Generated Video Software *Jan 2024 – May 2024*

End-to-end AI video creation and transcription tool

- Collaborated with an AI startup to develop software for auto video and image generation from text inputs
- Designed the UI/UX with Figma, integrated open-source AI APIs, used PostgreSQL on Google Cloud for back end, and developed a responsive front-end using TypeScript and Astro JS, increasing responsiveness by **15%**

PROFESSIONAL EXPERIENCE

Delivery Hero (FoodPanda) – *Sales Performance Intern* *Jun 2024 – Aug 2024*

- Streamlined self-sign up funnel by identifying bottlenecks and refining the onboarding process, reducing QC rejection rate by **15%**
- Integrated Google APIs to automatically sort and upload data from Drive to Sheets, enabling systemised analysis and reducing manual data handling, increasing hourly throughput by **25%**
- Automated invoice distribution via WhatsApp using Python, reducing process time by **20%**

Pakistan Television Network – *Data Science Intern* *Jun 2023 – Jul 2023*

- Revamped and enhanced PTV World's ticker headline system using machine learning models for text similarity, resulting in **20%** reduction in headline overheads
- Developed a predictive analysis model at Pakistan Television Network using Data Science techniques, enhancing viewer engagement metrics through data-driven content strategies by **10%**

TECHNICAL SKILLS & INTERESTS

Languages: English, Urdu
Technical Skills: Python, MERN, Java, C++/C, MATLAB, Haskell, React, LaTeX, Pytorch, AWS
Interests: Formula 1, Golf, Horse Riding, Applied Machine Learning, Data Science

LEADERSHIP EXPERIENCE AND CO-CURRICULAR ACTIVITIES

TechStars Startup Weekend, Lahore - *Convenor* *Jul 2022 – Mar2023*

- Executed Startup Weekend Lahore 2023 in collaboration with National Incubation Center Lahore and Google for Startups, led workshops and partnered with top industry leaders for startup guidance and judging
- Managed a 54-hour event for over **40** startups and **150+** attendees across Pakistan, overlooking a team of 60 and securing **800k+** PKR in sponsorships

SPADES - *Assistant Director (Socials and Operations)* *Sept 2022 – May 2023*

- Led the organisation of PSIFI's flagship event, coordinating social function and concerts for over **1,400** attendees
- Devised creative event ideas, managed itineraries, bookings, and secured necessary clearances and approvals