



HAFEEZ ULLAH

Chairperson & Lecturer, Department of Computer Science |
Director Artificial Intelligence and Data Innovation Lab (AIDIRL), University
of Gwadar |
PhD Candidate in Computer Science (Artificial Intelligence), FAST-NUCES
Karachi, Pakistan

Nationality: Pakistani

Date of birth: 07/07/1990

Phone: (+92) 336 358 5618

WhatsApp: (+92) 336 358 5618

Email: hafeezullah@ug.edu.pk

Website: hafeezug.github.io

LinkedIn: linkedin.com/in/hafeez-ullah-7a5328353

GitHub: github.com/hafeezug

Hugging Face: huggingface.co/hafeez007

Research lab: aidirl.github.io

Work address: Department of Computer Science, University of Gwadar, Airport Road, 91200 Gwadar, Balochistan, Pakistan

ABOUT ME

Academic leader and artificial-intelligence researcher with more than nine years of experience in university teaching, departmental leadership, and IT governance. Currently Chairperson of the Department of Computer Science at the University of Gwadar and a PhD candidate in Computer Science with an Artificial Intelligence specialisation at FAST-National University of Computer and Emerging Sciences (NUCES), Karachi. Research centres on Generative AI, Natural Language Processing, and low-resource languages, with a sustained focus on Balochi. Built BNer, the first gold-standard named entity recognition dataset for Balochi (58,893 tokens across five domains), and contributed the Balochi language to Global PIQA, a participatory multilingual commonsense-reasoning benchmark. Established the Artificial Intelligence and Data Innovation Research Lab (AIDIRL) and launched both the BS Computer Science (AI Specialisation) and BS Data Science programmes at the University of Gwadar.

RESEARCH INTERESTS

- **Low-resource NLP** — corpus construction, subword tokenisation, contextual embeddings, and named entity recognition for Balochi and other Perso-Arabic script languages.
- **Generative and agentic AI** — fine-tuning and instruction-tuning large language models, LLM evaluation, and agentic AI systems.
- **Multilingual evaluation** — culturally grounded benchmarks and cross-lingual performance gaps in state-of-the-art LLMs.

WORK EXPERIENCE

Director, Artificial Intelligence and Data Innovation Research Lab (AIDIRL)

University of Gwadar [03/2026 – Current]

City: Gwadar

Country: Pakistan

- Founded the laboratory in March 2026 and direct its research agenda across fifteen thematic areas spanning low-resource NLP, generative AI, and applied data science.
- Built the lab's research portfolio around local Balochistan and Gwadar challenges, with priority in coastal and fisheries applications, precision agriculture, and education technology.

- Launched and oversee the Gwadar AI Literacy Initiative (GAILI), the lab's public outreach programme for educators, students, and public-sector professionals.
- Supervise the lab's research internship programme, directing 30 interns across six teams towards publishable AI and NLP outputs and open dataset releases.
- Lead the lab's publication pipeline, dataset and model release standards, and collaboration with national and international research groups.
- Maintain the lab's public presence and research dissemination at <https://aidirl.github.io>

Chairperson, Department of Computer Science

University of Gwadar [01/2026 – Current]

City: Gwadar

Country: Pakistan

- Lead and oversee all academic and administrative operations of the Department of Computer Science.
- Drive strategic planning, curriculum enhancement, and quality-assurance initiatives.
- Launched the BS Computer Science (Artificial Intelligence Specialisation) programme and developed its complete curriculum.
- Established the Artificial Intelligence and Data Innovation Research Lab (AIDIRL) in March 2026 and launched the Gwadar AI Literacy Initiative (GAILI) in April 2026.
- Chaired the 6th Board of Studies of the department (May 2026).
- Mentor faculty, manage departmental resources, and liaise with university leadership.

Lecturer, Department of Computer Science

University of Gwadar [16/01/2017 – Current]

City: Gwadar

Country: Pakistan

- Deliver lectures and laboratories in Programming Fundamentals, Database Systems, Machine Learning, Deep Learning, Natural Language Processing, Generative AI, and Big Data Analytics.
- Develop comprehensive lesson plans, lab manuals, and assessment rubrics.
- Supervise Final Year Projects and undergraduate research students, and provide academic guidance.
- Design and evaluate examinations, assignments, quizzes, and project deliverables.

Provost, Boys Hostel

University of Gwadar [20/12/2022 – 08/2023]

City: Gwadar

Country: Pakistan

- Oversaw daily hostel operations, including student housing and facility management.
- Managed hostel staff and responded to student concerns and complaints.
- Maintained discipline in coordination with the warden.

Director of Information Technology

University of Gwadar [12/09/2022 – 08/2023]

City: Gwadar

Country: Pakistan

- Oversaw and evaluated ICT operations of the Directorate of IT.
- Analysed institutional ICT requirements and devised corresponding policies.
- Managed procurement of IT hardware and software.
- Supervised IT staff, including network administrator, web developer, and database administrator.

Chairperson, Department of Computer Science

University of Gwadar [09/2022 – 08/2023]

City: Gwadar

Country: Pakistan

- Oversaw departmental operations, faculty evaluation, and policy development.
- Spearheaded curriculum development and course scheduling across all programme levels.

- Launched the BS Data Science programme and developed its complete curriculum.
- Guided and motivated students, ensuring alignment with academic standards.

Deputy Campus Coordinator

University of Turbat — Gwadar Campus [26/09/2018 – 30/09/2020]

City: Gwadar

Country: Pakistan

- Oversaw day-to-day campus operations in the absence of the Director of Campus.
- Supervised secretariat and administrative staff of the campus.
- Provided administrative support to academic departments.
- Oversaw procurement processes, HR management, and record keeping.

Head of Department, Computer Science

University of Turbat — Gwadar Campus [01/2017 – 09/2020]

City: Gwadar

Country: Pakistan

- Directed departmental operations, staff evaluation, and academic policy implementation.
- Led curriculum development, class scheduling, and faculty workload planning.
- Provided academic counselling and motivational support to students.

EDUCATION AND TRAINING

Doctor of Philosophy (PhD), Computer Science — Artificial Intelligence

FAST–National University of Computer and Emerging Sciences (NUCES) [08/2024 – Ongoing]

City: Karachi

Country: Pakistan

Website: <https://nu.edu.pk>

- Specialisation: Generative AI · Natural Language Processing · Low-Resource Languages.
- **Thesis:** *Corpus Construction, Contextual Embedding Generation, Named Entity Recognition, and Fine-Tuning LLMs for a Low-Resource Language, Balochi.*

MS Computer Science

Mohammad Ali Jinnah University [09/2020 – 12/2022]

City: Karachi

Country: Pakistan

Website: <https://jinnah.edu>

- **Thesis:** *Balochi Named Entity Recognition Using Machine Learning Approach.*
- Merit Certificate — awarded for being among the highest CGPA achievers.

BS Information Technology

Balochistan University of Information Technology, Engineering and Management Sciences (BUIITEMS) [11/2009 – 02/2013]

City: Quetta

Country: Pakistan

Website: <https://buiitems.edu.pk>

TEACHING

Courses delivered at the University of Gwadar (lecture and laboratory components):

- Deep Learning with Laboratory (CS-454)
- Generative AI with Laboratory
- Natural Language Processing with Laboratory
- Machine Learning · Big Data Analytics
- Python Programming · Programming Fundamentals · Database Systems

Supervision: Final Year Projects and undergraduate research students; design of lab manuals, midterm and terminal laboratory examinations, and assessment rubrics.

LANGUAGE SKILLS

Mother tongue: Balochi

LANGUAGE	LISTENING	READING	WRITING	SPOKEN PRODUCTION	SPOKEN INTERACTION
Urdu	C1	C1	C1	C1	C1
English	C2	C1	C1	C2	C2
Brahui	A1	A1	A1	A1	A1

DIGITAL SKILLS

- **Programming and environments:** Python; Jupyter Notebook; Google Colab; Kaggle Notebooks; Git and GitHub.
- **NLP and machine learning:** Hugging Face Transformers, Scikit Learn, Pytorch, Tensorflow
- **AI Agents and Gen AI:** LangGraph, Crew AI, Autogen, Google ADK, LangChain, Llama Index
- **Data Analysis:** Pandas, Numpy, Matplotlib, Seaborn, Tableau, Excel
- **Data annotation and release:** Label Studio; doccano; inter-annotator agreement (Cohen's κ); Hugging Face Hub and Zenodo dataset publishing.

PUBLICATIONS

Peer-reviewed and preprint

- **Global PIQA: Evaluating Physical Commonsense Reasoning Across 250+ Languages and Cultures.** Co-author — contributed the Balochi language portion. Preprint: arXiv:2510.24081; submitted to NeurIPS 2026. A participatory benchmark developed by 350+ researchers spanning 100+ language varieties; verified translations and integrated culturally specific items for Balochi to assess cross-lingual performance gaps.

Under submission

- **Balochi Named Entity Recognition Using a Machine Learning Approach.** Introduces BNer, the first gold-standard NER dataset for Balochi, comprising 58,893 tokens across five domains. Benchmarks CRFs, SVMs, and k-NNs using POS and BILOU tagging schemes, with SVM achieving a 96% F1-score.

Manuscripts in preparation

- **BalToken: Subword Tokenizer Optimization for a Low-Resource Language, Balochi.** Cross-lingual benchmarking of fifteen tokenisers using fertility, Min-Entropy (H_∞), and Rényi entropy metrics.
- **Large Language Models for Perso-Arabic Script Languages: A Survey.** PRISMA-based survey covering eighteen Perso-Arabic script languages across encoder-only, encoder-decoder, and decoder-only architectures
- **Large Language Models, Datasets, and Benchmarks for Riddle Generation and Evaluation: A Survey.** Monolingual and multilingual survey identifying the riddle-reasoning coverage gap as the basis for a planned multilingual riddle benchmark.

PROJECTS

- **Balochi Corpus Construction (80–100M tokens).** Design and implementation of a synthetic-data and web-scraping pipeline using NLLB-200-distilled-1.3B via CTranslate2, with LaBSE round-trip

filtering and a quality stack of GlotLID, KenLM, and MinHash-LSH deduplication; production notebooks orchestrated across Colab and Kaggle GPU allocations.

- **Balochi Language Models and Embeddings.** Training of a monolingual Balochi BERT-base model, development of cross-lingual embeddings from Persian and Urdu to Balochi, instruction tuning of Gemma, and fine-tuning for named entity recognition — the six-objective programme of the doctoral thesis.
- **AIDIRL Research Portfolio for Balochistan and Gwadar.** Research-project portfolio mapped to the lab's fifteen themes and directed at local challenges, with priority in coastal and fisheries applications, precision agriculture, and education technology; includes an EdTech dropout early-warning pipeline for Gwadar.
- **Gwadar AI Literacy Initiative (GAILI).** AIDIRL outreach programme delivering AI literacy training to educators, students, and public-sector professionals in Gwadar.
- **Undergraduate Research Internship Programme.** Direction of 30 interns across six teams towards publishable AI and NLP research, a benchmark coverage matrix for Pakistani languages, and annotation and release standards.

CONFERENCES, WORKSHOPS AND SEMINARS

- **Effective Use of Artificial Intelligence in Teaching and Learning.** One-day hybrid workshop delivered voluntarily under the Gwadar AI Literacy Initiative (GAILI), AIDIRL, Department of Computer Science, University of Gwadar — April 2026.
- **Promoting Literacy in the Digital Era.** Volunteer panellist, RISE Balochistan Programme, Syed Hashmi Auditorium, RCDC, Gwadar — April 2026.
- **How to Use AI Tools in Office Management, Finance and Engineering.** One-day voluntary workshop delivered at the Gwadar Development Authority — September 2025.

ACADEMIC SERVICE, NETWORKS AND AFFILIATIONS

- **Founder, Artificial Intelligence and Data Innovation Research Lab (AIDIRL),** Department of Computer Science, University of Gwadar — established March 2026.
- **External evaluator, MS thesis examinations,** FAST–National University of Computer and Emerging Sciences (NUCES), Karachi — May 2026.
- **Chair, 6th Board of Studies,** Department of Computer Science, University of Gwadar — May 2026.
- **Contributor, Global PIQA multilingual benchmark consortium** — Balochi language lead within a 350+ researcher collaboration.

CERTIFICATIONS AND TRAINING

- **AI Agent Developer** — Vanderbilt University [2025]
- **Generative AI Engineering with LLMs** — IBM [2025]
- **Generative AI for Educators** — IBM [2025]
- **Fundamentals of Building AI Agents** — IBM [2025]
- **NVIDIA: Prompt Engineering and Data Analysis** — Whizlabs [2025]
- **OpenAI GPTs: Creating Your Own Custom AI Assistants** — Vanderbilt University [2025]
- **Generative AI for Code Generation for Python** — Edurik [2025]
- **Deep Learning Specialization** — DeepLearning.AI [2024]
- **Generative AI Concepts** — DataCamp [2024]

HONOURS AND AWARDS

Gold Medal

Syed Hashmi Reference Library

- Awarded for achieving the highest score in the Balochi Language Certificate Course.

Merit Certificate

Mohammad Ali Jinnah University

- Awarded for being among the highest CGPA achievers.

MANAGEMENT AND LEADERSHIP SKILLS

- **Academic leadership:** departmental headship across two universities; strategic planning, quality assurance, faculty evaluation, and mentoring.
- **Curriculum development:** authored complete curricula for BS Computer Science (AI Specialisation) and BS Data Science; Board of Studies chairmanship.
- **IT governance:** ICT policy formulation, hardware and software procurement, and supervision of technical staff.
- **Research leadership:** establishment and direction of a research laboratory; supervision of interns, final-year projects, and research students.
- **Community engagement:** design and delivery of voluntary AI literacy programmes for educators, students, and public-sector institutions in Gwadar.