Abdul Rahman Shaikh

DeKalb, IL ● +1-224-848-2725 ● iamsabdurahman@gmail.com

linkedin.com/in/iamsabdurahman ● sabdulrahman.github.io ● Google Scholar

SUMMARY

Ph.D. candidate in Computer Science with 7+ years of research experience in applied machine learning, statistical modeling, and AI system development. Proven track record of extracting insights from complex datasets and building scalable solutions across NLP, generative AI, and multimodal systems. Strong foundation in translating research into production-ready applications using Python, PyTorch, HuggingFace, LangChain and modern ML tools.

EDUCATION

Northern Illinois University — DeKalb, Illinois
Ph.D. & M.S. in Computer Science (Expected Aug 2025) — GPA: 3.9

TECHNICAL SKILLS

- **AI/ML Frameworks:** LangChain, PyTorch, HuggingFace, Keras, OpenCV, Scikit-learn, Flowise, N8N
- **LLM & GenAI:** Llama, LangChain, Fine-tuning (LoRA, QLoRA), RAG, Flash Attention
- **Programming:** Python, SQL, JavaScript, C, C++, Java, Node.js, D3.js, R, LaTex, TypeScript
- **Data Science:** Pandas, NumPy, Matplotlib, Seaborn, Plotly, Statistical Analysis, A/B Testing
- **Cloud & MLOps:** AWS, Azure, Docker, Kubernetes, MLFlow, Git, Weights & Biases
- **Databases & BI:** MySQL, MongoDB, PostgreSQL, PowerBI, Tableau

EXPERIENCE

Northern Illinois University (DATA Lab, VA Lab & WASTE Lab) — DeKalb, IL
Researcher — Jan 2018 - Present

- Designed and deployed end-to-end machine learning systems leveraging PyTorch, MLflow, and Docker, reproducible training, versioned model tracking, and scalable batch inference for over 1M data records achieving 91% accuracy.
- Engineered a retrieval-augmented generation (RAG) pipeline using LangChain, FAISS, and GPT-4 to support dynamic question answering over large academic and clinical corpora, reducing manual review time by 65% in applications.
- Designed a LLM-powered visualization framework improving insight discovery by 45% and created Power BI/Tableau dashboards to visualize KPIs and pricing trends supporting strategic decisions and cost reduction.
- Deployed machine learning models to production environments using FastAPI, Docker, and Kubernetes, with complete CI/CD integration and experiment/version tracking via MLflow and DVC that improved model performance by 23%.
- Built multi-agent LLM systems (LangChain, Flowise, AutoGen) for exploration and summarization in complex domains.

Amazon — India
Quality Specialist — Aug 2017 - Dec 2017

- Optimized SQL and MongoDB query performance for e-commerce data pipelines, reducing database latency by 30%.
- Implemented automated data quality assurance workflows using Python (Pandas, NumPy, Scikit-learn) to validate 500K+ daily records, achieving 95%+ accuracy and enhancing operational reliability by 22% across critical Amazon systems.

PROJECTS

- **LLMFlow: Summarization of Scholarly Documents (Python, LangChain)** Integrated GPT-4 and RAG pipelines with LangChain to summarize large academic documents, reducing reading time by 60% across 200+ documents. (Jan - July 2024)
- **MixArt: Generative Artwork with Stable Diffusion (Python, Stable Diffusion, LoRA)** Designed an art-generation pipeline using Stable Diffusion fine-tuned with LoRA, allowing customized outputs based on user input. (July - Dec 2024)
- **EcoScan: AI-Powered Ocean Pollution Mapping (Python, TensorFlow, CNNs)** Developed an AI system utilizing CNNs to analyze satellite imagery for detecting oceanic microplastic pollution, enhancing efficiency by 90%. (Jan - Apr 2024)
- **GenHealth: Multimodal Medical Report Analysis (Python, GPT-4, BLIP-2)** Created a multimodal analysis pipeline combining GPT-4 and BLIP-2 for medical reports, boosting diagnostic extraction accuracy by 30%. (Sep - Dec 2024)

PUBLICATIONS

- **A. R. Shaikh**, H. Alhoori, and M. Sun, "YouTube and Science: Models for Research Impact," *Journal of Scientometrics*, 2022. [*doi: 10.1007/s11192-022-04574-5*]
- **A. R. Shaikh**, M. Sun, and H. Alhoori, "Toward systematic design considerations of organizing multiple views," IEEE Visualization and Visual Analytics (VIS), 2022. [*doi: 10.1109/VIS54862.2022.00030*]
- M. Sun, **A. R. Shaikh**, H. Alhoori, and J. Zhao, "SightBi: Exploring Cross-View Data Relationships with Biclusters," *IEEE TVCG 2021*, [*doi: 10.1109/TVCG.2021.3114801*] **Best Paper Honorable Mention.**
  For a complete list of publications, please visit - **Scholar profile** or **Website**

TEACHING/LEADERSHIP

Mentor at NIU WASTE LAB — DeKalb, IL, USA
Mentored and supervised 6 graduate students resulting in 4 co-authored research papers. — Jan 2024 - Present

Lab Head at NIU VA LAB — DeKalb, IL, USA
Designed and Led multiple projects related to exploration using multiple view visualizations. — Aug 2021 - Present

Teaching Assistant at NIU — DeKalb, IL, USA
Supported undergraduate courses in C++, Java, and Databases courses. — Aug 2018 - May 2025