

Abdul Rahman Shaikh

Data Scientist • AI/ML Engineer

+1 224 848 2725 iamsabdurahman@gmail.com Chicago, IL
sabdulrahman.github.io linkedin.com/in/iamsabdurahman github.com/sabdulrahman

SUMMARY

Data Scientist (Ph.D.) with 7+ years in ML/LLM systems end-to-end, from problem framing and feature engineering to A/B testing, deployment, and drift monitoring. Python, SQL, PyTorch, Spark, AWS; MLOps via MLflow/Kubeflow, Docker, Kubernetes. Translate state-of-the-art research into reliable, cost-aware products with measurable impact.

EDUCATION

- Northern Illinois University** Illinois, USA
Ph.D. & M.S. in Computer Science (Dec 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, 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:** PostgreSQL, MySQL, MongoDB, Tableau, Power BI

EXPERIENCE

- Northern Illinois University (DATA Lab, VA Lab & WASTE Lab)** DeKalb, IL
Data Scientist Jan 2018 - Present
 - Architected and deployed scalable ML pipelines processing 1M+ daily transactions, reducing model inference latency by 65% through distributed computing frameworks like Apache Spark and Kubernetes orchestration.
 - Led end-to-end development of recommendation systems using deep learning architectures (transformers, neural collaborative filtering) that increased user engagement by 26%.
 - Enhanced model efficiency by implementing quantization, pruning, and distillation techniques, achieving 4x throughput on edge devices without sacrificing accuracy.
 - Implemented MLOps best practices with CI/CD pipelines, A/B testing frameworks, and monitoring systems using tools like MLflow, Kubeflow, and DataDog, reducing model deployment time from weeks to hours.
 - Built multi-agent LLM systems (LangChain, Flowise, AutoGen) for exploration and summarization in complex domains.
- Amazon** India
Data Analyst Aug 2016 - Dec 2017
 - Accelerated e-commerce database performance by 30% through SQL and MongoDB query optimization.
 - 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: Scholarly Document Intelligence System** Built an advanced RAG pipeline using LLMs with vector databases (Pinecone/Weaviate), achieving 65% reading time reduction across 100+ academic papers.
- GenHealth: Multimodal Diagnostic Intelligence System** Developed a multimodal AI system combining medical text, imaging, and structured data to boost diagnostic accuracy. Utilized NLP and CNNs to extract, align, and predict outcomes.
- MixArt: Advanced AI Art Generation Platform** Engineered a production-ready art generation system using FLUX and SDXL-Turbo with ControlNet integration, featuring real-time style transfer and deployment via Gradio.
- ChatFit: Personalized AI Fitness Coach** Built an end-to-end AI chatbot using LLMs for personalized workout and diet advice tailored to user profiles. Integrated React frontend and FastAPI backend with session logic and prompt tuning.

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
Led AI-driven visualization projects, significantly advancing multi-view exploratory analytics. Aug 2021 - Present
- Teaching Assistant at NIU** DeKalb, IL, USA
Supported undergraduate courses in C++, Java, and Databases. Aug 2018 - Aug 2023