Dr. Rajkrishna Mondal

📱 7843909820   ✉ rajkrishna.math@gmail.com   📍 West Bengal, India

in linkedin.com/in/rajkrishna92   🐙 github.com/rajkrishna92   🌐 rajkrishna92.github.io

PROFESSIONAL SUMMARY

Ph.D. in Mathematics and Senior Data Scientist with extensive experience in Generative AI, Agentic Workflows, and Multi RAG architectures. Proven track record of spearheading scalable AI systems, optimizing Large Language Model (LLM) performance, and leading cross-functional technical teams. Expert in transforming raw LLM capabilities into robust, enterprise-grade tools using Python, LangGraph, ADK, MCP and MLOps best practices.

WORK EXPERIENCE

Accrete AI   **Mumbai, India**
Senior Prompt Engineer   *June 2025 – Present*

- Spearheading the architecture and production deployment of scalable Generative AI systems, reducing hallucination rates through optimized RAG pipelines using Vector Stores (Chroma) and semantic search.

- Designing autonomous agentic workflows using Google's Agent Development Kit (ADK) and LangGraph, enabling multi-step reasoning and complex tool execution.

- Engineering high-impact prompt strategies (Few-Shot, Chain-of-Thought) for diverse LLMs (GPT-4, Gemini, Ollama), ensuring deterministic outputs in production environments.

- Developing high-performance backend microservices using Python and FastAPI to expose GenAI capabilities as scalable APIs.

- Managing end-to-end MLOps lifecycles, utilizing Docker for containerization and Kubernetes for orchestration to ensure high availability.

- Establishing organizational prompt governance and mentoring junior engineers on advanced Python patterns and prompt engineering techniques.

Infinity Analytics Pvt. Ltd.   **Mumbai, India**
Data Scientist (Project Lead Developer)   *Sept 2021 – May 2025*

- Led the design and deployment of an AI-powered cognitive search engine, transforming data interaction through advanced NLP and agent-based workflows.

- Defined system architecture integrating LLMs (OpenAI, Ollama), Vector Stores, and LangChain for automated SQL generation and document ingestion.

- Streamlined internal processes by automating data extraction, significantly reducing manual effort and operational costs.

- Managed cross-functional teams, overseeing timelines and CI/CD pipelines to ensure scalable, secure deployments.

- Acted as technical liaison, translating complex AI concepts into actionable business insights for stakeholders, directly contributing to revenue growth and improved customer satisfaction.

TECHNICAL SKILLS

Generative AI & LLM:
LLMs (GPT-4, Gemini, Llama, ...), RAG, ADK, MCP, LangChain, LangGraph, CrewAI, Llama Index, Prompt Engineering (CoT, Few-Shot), Agentic Workflows, Semantic Search.

Languages & Core:
Python, SQL, R, C++, MATLAB, LaTeX.

Libraries & Frameworks:
PyTorch, TensorFlow, Keras, Scikit-Learn, Pandas, NumPy, FastAPI, Spacy, NLTK.

MLOps & Tools:
Docker, Kubernetes, Git, CI/CD, ChromaDB, Vector Stores, PySpark, Tableau, Power BI.

EDUCATION

Motilal Nehru National Institute of Technology Allahabad (MNNITA)   **Prayagraj, India**
Ph.D. in Mathematics (Soft Computing & Medical Diagnostics)   *Aug 2022*
Research Topic: "Design and Development of Soft Computing Diagnostic information system in Medical Science"

Motilal Nehru National Institute of Technology Allahabad (MNNITA)   **Prayagraj, India**
Master of Science (M.Sc) in Mathematics and Scientific Computing   *June 2016*

Calcutta University   **Kolkata, India**
Bachelor of Science (B.Sc) in Mathematics Honours   *July 2014*

KEY CERTIFICATIONS

Generative AI & LLMs: Generative AI with LLMs (DeepLearning.AI), LangChain for LLM Application Development, Advanced RAG with Chroma, Multi AI Agent Systems with crewAI.

Deep Learning & NLP: Deep Learning Specialization (Coursera), Natural Language Processing Specialization (DeepLearning.AI), Deep Neural Networks with PyTorch (IBM).

Data Engineering: Big Data Specialization (UCSD), Complete SQL Mastery.

PUBLICATIONS

Author of multiple peer-reviewed journal and conference papers. Full list available on Google Scholar.

Mondal, R., & Srivastava, P. (2022). *Fuzzy Utility Matrix-Based Intelligent Decision-Making Model and Its Application to Diet Recommendation System for Metabolic Disorder Patients.* International Journal of Fuzzy System Applications (IJFSA), 11(1), 1–22. DOI: 10.4018/IJFSA.303563

Srivastava, P., & **Mondal, R.** (2021). *Design and Development of Intelligent Information System using Hesitant Fuzzy Weighting Linguistic Term Sets for Computing with Words.* Lecture Notes in Networks and Systems (Springer), Vol. 214, 2022. DOI: 10.1007/978-981-16-3807-7

Srivastava, P., & **Mondal, R.** (2021). *QUALIFLEX Based Ranking System by Using Interval-Valued Hesitant Fuzzy Set and Its Application to Rank Diabetic Patients.* Proceedings of the 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), pp. 78–85 (IEEE). DOI: 10.1109/ICCCIS51004.2021.9397179

Mondal, R., & Srivastava, P. (2021). *Design and Development of an Intelligent System to Assess Kidney Performances of Persons Suffering from Diabetes.* International Transaction Journal of Engineering, Management & Applied Sciences & Technologies, 12(4), 1–14. DOI: 10.14456/ITJEMAST.2021.64

Srivastava, P., & **Mondal, R.** (2020). *A Hesitant Fuzzy Envelope Based Expert System in Human Decision Making.* Nepal Journal of Mathematical Sciences, 2(1), 1–6. DOI: 10.3126/njmathsci.v2i1.36503

Srivastava, P., & **Mondal, R.** (2020). *Diabetes Diagnostic Intelligent Information System.* TEST Engineering & Management, 82 (Jan–Feb), 14455–14467. URL: http://www.testmagzine.biz/index.php/testmagzine/article/view/3136

Mondal, R., Verma, A., & Gupta, P. K. (2020). *A Medical Diagnostic Information System with Computing with Words Using Hesitant Fuzzy Sets.* Advances in VLSI, Communication, and Signal Processing (Springer), pp. 971–980. DOI: 10.1007/978-981-32-9775-3_86

Verma, A., **Mondal, R.**, Gupta, P., & Kumar, A. (2018). *Neural Based Energy-Efficient Stable Clustering for Multilevel Heterogeneous WSNs.* Proceedings of the 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC), pp. 208–212 (IEEE). DOI: 10.1109/ICSCCC.2018.8703353