GARGI KALE

gargik@uchicago.edu | +1(312)-221-5719 | LinkedIn | Kaggle | Github

EDUCATION

University of Chicago, Physical Sciences Division — **Chicago, IL**
Master of Science in Data Science — **Expected Dec 2026**

- Relevant Courses: Advanced Adversarial Machine Learning, Time Series Analysis and Forecasting, AI Agents for Science, Foundations of Machine Learning and Neural Networks
- Research assistant at The Rustandy Center for Social Sector Innovaton

MKSSS's Cummins College of Engineering for Women, Pune — **Pune, IND**
Bachelor of Technology in Computer Engineering — **June 2024**

- Relevant Courses: Cloud Computing, Reinforcement Learning , Introduction to Natural Language Processing, Statistics for Computer Science, Big Data & Analytics

Indian Institute of Technology, Madras — **Online Program**
Bachelor of Science in Programming and Data Science — **May 2024**

- Relevant Courses: Introduction to Large Language Models, Deep Learning, Machine Learning Techniques and Practices, Advanced Algorithms, Tools in Data Science

TECHNICAL SKILLS

- **Programming Languages:** Java, Python (PyTorch, TensorFlow, spaCy, LangChain), SQL, R
- **Database & Analytics:** MySQL, MongoDB, Redis, Snowflake, Databricks, Stata, Spark
- **ML Concepts:** Hypothesis Testing, ANOVA, Regression Models, Decision Trees, Neural Networks, Time series forecasting
- **Cloud Technologies:** AWS, Azure, Docker, Kubernetes
- **DevOps & CI/CD Tools:** Jenkins, Terraform, GitHub, GitLab, Bitbucket, AWS CodePipeline

WORK EXPERIENCE

JPMorganChase — **Hyderabad, IND**
Software Engineer — **Aug 2024 - Sept 2025**

- Implemented ETL workflows across multiple datasets, enabling scalable and efficient data processing.
- Developed in-built mechanisms for data reconciliation to ensure accuracy, consistency, and completeness for end users.
- Built and deployed an end-to-end AWS-based data warehousing solution, including infrastructure provisioning of Lambdas, ECS tasks and Step Functions for secure orchestration of confidential data pipelines.

Oneirix Labs — **Pune, IND**
Natural Language Processing Intern — **Aug 2023 - Dec 2023**

- Compared performance of multiple Large Language Models (LLMs), including DeBERTa, Longformer, to improve accuracy in extracting information from natural language queries.
- Developed customised spaCy pipelines for preprocessing natural language queries, enabling cleaner and more efficient downstream processing.
- Designed algorithms to generate optimised SQL queries from extracted data, enhancing query efficiency and performance.

JPMorganChase — **Bengaluru, IND**
Software Engineer Intern — **June 2023 - July 2023**

- Designed and developed an internal application using Spring Boot, ReactJS, and Java, with APIs and a UI layer that enabled reading all repository properties and automatically comparing them against default configurations to ensure consistency and compliance.

SELECT PROJECTS

Multi-Agent Framework for Generating Executable Research Papers

- Built a domain-agnostic autonomous multi-agent system using OpenAI API (LLMs) orchestrating planning, code generation, execution to transform research questions into fully executable scientific papers.
- Implemented an automated environment construction, sandboxed execution, logging, and reproducibility fault-injection pipeline, generating and evaluating 30 synthetic papers across domains and outperforming a single-LLM baseline (24.1 vs 14.1 avg. score). *github* , Dec 2025

Marathi Paraphrase Generation

- Analysed 4 LLMs for the generation of Marathi paraphrases. Created a dataset of Marathi phrases and their corresponding paraphrases to evaluate the performance of the LLMS using performance metrics like BLEU scores, METEOR scores, cosine similarity, and Levenshtein distance. *kaggle* , May 2024

Spotify Playlist Recommendation

- As part of the Microsoft Engage Mentorship program, created a machine learning model using KNN algorithm which predicted most suitable playlist for a song. Used Django to create a frontend integrated with the model. *github* , July 2022