

# GARGI KALE

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

## EDUCATION

University of Chicago, Physical Sciences Division  
Master of Science in Data Science

Chicago, IL  
Expected Dec 2026

- Relevant Courses: Data Interaction, Responsible Use of Data and Algorithms, 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  
Bachelor of Technology in Computer Engineering

Pune, IND  
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  
Bachelor of Science in Programming and Data Science

Online Program  
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), JavaScript, 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  
Software Engineer

Hyderabad, IND  
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  
Natural Language Processing Intern

Pune, IND  
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  
Software Engineer Intern

Bengaluru, IND  
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

Marathi Paraphrase Generation

- Analysed 4 LLMs (MultiIndicParaphraseGeneration, T5, BART, PEGASUS) 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

Statistical analysis of stock prices

- Analysis of stock prices using interactive animated visualisation techniques. Forecasted stock prices using Auto-Regressive Integrated Moving Average (ARIMA) model in python. [kaggle](#) , May 2023

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

## ADDITIONAL INFORMATION

- Active participant of JPMorganChase GoodWorks
- Winner of Intercollegiate BMC Hackathon

2024 - 2025  
2024