Mithil Baria

Mumbai, India 421202 +919920925281 | mithilbaria98@gmail.com | LinkedIn: https://www.linkedin.com/in/mithil-baria-8873 47173 |

WWW: https://github.com/Mithil21

AI/LLM Engineer and Full-Stack Software Developer with 5+ years of experience building secure, scalable applications and Al-driven systems in finance, healthcare, and enterprise domains. Expert in backend architecture (Java, Spring Boot) and deploying ML solutions using PyTorch, Transformers, and RAG frameworks. Built intelligent tools like a RAG-based code reuse engine and GPT-2-powered positivity generator. Demonstrated success in delivering 25–50% improvements in performance, security, and delivery timelines. Published researcher and hackathon winner, passionate about solving real-world problems through AI and secure system design. Open to relocation and global opportunities.

Skills

Languages: Java, Python, TypeScript, C++, JavaScript

AI/ML: PyTorch, TensorFlow, Transformers, Scikit-learn, Hugging Face, LangChain, FAISS

DevOps & Cloud: AWS (EC2, S3, IAM), Azure AD, Jenkins, Docker, GitHub Actions

Tools: Git, Jira, Postman, Streamlit, Cucumber, SVN

Frameworks: Spring Boot, Hibernate, Flask, Angular, React, Node.js

LLM/RAG: Prompt Engineering, Embedding-based Search, Retrieval-Augmented Generation, GPT Fine-tuning

Security: Spring Security, OAuth2, JWT, SSO

Integration

Databases: MongoDB, PostgreSQL, Oracle,

MySQL

Work History

Software Engineer

Wolters Kluwer, Pune

August 2022 - July 2025

- Architected and developed enterprise-grade applications (Spring Boot + Angular), boosting performance by 25%
- Directed and mentored 8-member team, improving delivery timelines by 30%
- Designed 5+ custom backend modules and schedulers for automating batch processes, task orchestration, and activity logging, improving responsiveness by 40%
- Integrated Azure AD SSO + Spring Security, strengthening authentication workflows by 50%
- Led POC on JWT token vulnerability in Azure AD architecture; proposed secure token-refresh mechanisms
- Winner, Code Games Hackathon: Built RAG-based reusable code generation tool reducing testing overheads and increasing delivery efficiency

Software Developer Intern

Phillips, Bengaluru

September 2021 - July 2022

- Built and tested backend APIs using Spring Boot; implemented 30+ BDD test cases to enhance QA coverage
- Automated deployment with AWS and Azure DevOps, improving CI/CD efficiency by 15%
- Worked on Spring Security and IAM workflows to tighten authentication in healthcare systems

Research Intern

- Developed ML models for early cancer diagnosis, improving predictive accuracy by 25%
- Contributed to algorithm development aligned with real-world medical workflows

Associate Consultant

Infrasoft Technologies, Mumbai

July 2019 - January 2021

- Engineered and optimized financial applications using Spring and Hibernate
- Reduced MongoDB query time by 30%, improving real-time system reliability
- Built UI components using JavaScript and AJAX for responsive interfaces

Education

Master of Science in Computer Science January 2023 Somaiya College of Engineering, Mumbai, India

- Thesis: Fake Review Detection using RoBERTa achieved 75%+ accuracy
- Relevant Courses: Advanced Machine Learning, Secure Software Development, Cybersecurity, Data Privacy

Bachelor of Science in Information Technology June 2019 University of Mumbai, Mumbai

GPA: 7.21 / 10

Projects

Code Reusability with RAG: Developed RAG system retrieving code snippets from structured corpus based on user input. Integrated with Streamlit UI; boosted software delivery and reduced redundant coding.

https://github.com/Mithil21/MemoryMatrix

Fake Review Detection (Published): Built web app using RoBERTa to flag suspicious reviews on Amazon. Hosted full-stack system (Flask + Angular).

URL Shortener with Docker: Created a containerized microservice for URL shortening using Java and Spring Boot. Deployed using Docker for isolated and scalable environments.

https://github.com/Mithil21/URLShortnerOnDocker

Positivity Generator (GPT-2 LoRA): Fine-tuned GPT-2 using LoRA/QLoRA for delivering personalized motivational messages.

Gujarati OCR Engine: Built ML model for handwritten Gujarati script recognition, achieving 82% accuracy for regional digitization.

Certifications & Awards

- Winner, Code Games Hackathon, 2024: Developed & Deployed AI tools using RAG, Transformers, and custom fine-tuning techniques
- Certified in Java, Python, JavaScript (Grasshopper)
- Paper Published: Fake Review Detection using RoBERTa, IJMTST

Leadership & Extras

- National Level Cricket Player
- Trekking Leader
- Book Author
- Enthusiast in Swimming and Boxing