

# MARUTI KAMESHWAR MODALI

College Park, MD (Ready to relocate)  
+1 240-501-7790  
[marutikameshwari@gmail.com](mailto:marutikameshwari@gmail.com)  
[linkedin.com/in/maruti-kameshwari-modali](https://linkedin.com/in/maruti-kameshwari-modali)  
[github.com/marutikameshwari](https://github.com/marutikameshwari)

## SUMMARY

Software Engineer with over 1.5 years of experience architecting scalable enterprise applications and end-to-end AI/ML solutions. Expert in bridging the gap between backend development and Machine Learning, with a focus on Large Language Models (LLMs), RAG pipelines, and MLOps. Proven track record of deploying high-impact intelligent systems via CI/CD and AWS, delivering production-ready code in fast-paced Agile environments.

## EXPERIENCE

- Bajaj Finserv Direct Ltd** Pune, India  
*Software Engineer* Jul 2023 - Aug 2024
- Re-engineered a normalized **PostgreSQL** database architecture and indexing strategy for a high-volume fintech platform to optimize data retrieval, achieving a **40% improvement in query performance**.
  - Engineered a real-time **Java analytics dashboard** with integrated JPA/Hibernate to automate tracking of application approval trends, **reducing manual stakeholder reporting effort by 30%**.
  - Conducted **A/B testing** on competing user onboarding flows to identify and eliminate friction points in the registration process, resulting in a **15% increase in user conversion**.
- Bajaj Finserv Direct Ltd** Pune, India  
*Software Engineer Intern* Jan 2023 - Jun 2023
- Built an asynchronous rapid onboarding pipeline using **Java Spring Boot** and **AWS SQS/SNS** to handle high-velocity customer data ingestion, successfully **reducing total onboarding latency by 60%**.
  - Partnered with Product Owners in an **Agile sprint** to architect a **Bulk Upload feature via AWS S3**, streamlining multi-user data processing and **eliminating 95% of manual data entry errors**.
  - Streamlined the deployment lifecycle by resolving 450+ technical debt issues and maintaining 90% JUnit coverage, leveraging **CI/CD pipelines** to ensure **99.9% system uptime** for data-critical services.

## PROJECTS

- Phishing Website Detection Pipeline** – *Python, Airflow, MLflow, AWS, Docker* [GitHub ↗](#)
- Architected an automated **MLOps pipeline** using **Airflow** and **MLflow** to evaluate **SVM, Logistic Regression, and Gradient Boosting**, streamlining feature selection to achieve an **0.98 F1-score** and eliminate training bottlenecks.
  - Engineered a robust **CI/CD workflow** via **GitHub Actions** and **Docker** for deployment to **AWS EC2**, enabling **real-time inference** and ensuring seamless production updates with minimal deployment overhead.
- Multi-Agent RAG Pipeline with Dynamic Routing** – *LangGraph, Python, Hugging Face, ChromaDB, Groq* [GitHub ↗](#)
- Orchestrated a **multi-agent framework** using **LangGraph** to route queries between private data and public search, overcoming static LLM limits via specialized agents to ensure **high response accuracy**.
  - Developed a retrieval engine using **Hugging Face** and **ChromaDB** for persistent memory, achieving **high-speed retrieval** and improved contextual awareness for complex private data queries.
- Natural Language SQL (Text-to-SQL) Agent** – *LangChain, LangSmith, Llama-3, Groq* [GitHub ↗](#)
- Designed a **RAG-based SQL agent** via **LangChain** for a **Student Database**, optimizing query accuracy through systematic **LangSmith** tracing and reasoning analysis for non-technical users.
  - Integrated **Llama-3 (Groq)** for low-latency streaming inference to create an "instant-reply" experience, **significantly reducing query latency** and allowing for real-time analytics without manual SQL intervention.

## EDUCATION

- University of Maryland** College Park, MD  
*May 2026 Master of Science in Data Science; GPA: 3.92/4.0*
- Vellore Institute of Technology** Andhra Pradesh, India  
*Aug 2019 - Jul 2023 Bachelor of Technology in Computer Science; GPA: 8.65/10.0*

## TECHNICAL SKILLS

<b>AI &amp; Generative AI</b>	LangChain, LangGraph, LangSmith, RAG, AWS Bedrock, MCP, Neo4j, Mistral-7B, Groq, Ollama, ChatGpt, LoRA & QLoRA Fine-tuning, ChatGPT
<b>ML &amp; Data Science</b>	PyTorch, TensorFlow, Scikit-Learn, Hugging Face, NLTK, NumPy, Pandas, Matlab, R, Tableau, Matplotlib, Seaborn
<b>MLOps &amp; Cloud</b>	MLflow, Airflow, Docker, GitHub Actions, AWS (EC2, S3, Lambda), Linux, Jenkins
<b>Software Engineering</b>	Java, Spring Boot, Flask, JPA Hibernate, JUnit, SQL (MySQL), MongoDB, HTML/CSS, Git