

AYUSH KHAMRUI

Bangalore, Karnataka | khamruiasok@gmail.com | Ph- 7001477696 | <https://www.linkedin.com/in/ayush-khamrui/>

Project - <https://github.com/Ayush-Khamrui/GenFlow-AI-Powered-Document-Intelligence>

WORK EXPERIENCE

Senior Software Engineer | Capgemini

Apr 2024 - Present

- **Led a 5-member team in designing and architecting complex LLM systems (RAG & Fine-Tuning)** across AWS Bedrock, Azure, and Hugging Face using Langchain and LlamaIndex; co-owned the team's technology vision, contributing significantly to its success.
- **Optimized LLM retrieval by employing various retrievers, prompt compression techniques, and guardrails**; crafted efficient prompts using context and zero-shot, one-shot and few-shot techniques; and applied advanced algorithms and data structures with a keen understanding of space and time complexities.
- **Enhanced code quality and maintainability** by applying object-oriented programming principles and design patterns to refactor, optimize, and debug code.
- **Developed efficient, maintainable code in Python and React**, adhering to industry coding standards to solve complex business problems.
- **Integrated Neo4j and ChromaDB as knowledge bases**, working extensively on Linux/Unix platforms for deployment and application management.
- **Created multiple APIs using Python Flask and FastAPI** integrating all services (Query, UML diagram generation, RAGAS Score, Carbon footprint calculation) to deliver robust backend solutions.
- **Integrated Azure Cache for Redis for caching** LLM responses and using it for faster retrieval enhancing application performance.
- **Followed SDLC in an agile environment**, collaborating with cross-functional teams to ensure on-time deliveries.
- **Developed and evaluated LLMs using RAGAS scores and MLflow**, demonstrating strong technical aptitude and solid computer science fundamentals.

Software Engineer | Capgemini

Sep 2022 – Mar 2024

- **Led frontend development of a GenAI Chatbot using React and Python**, integrating AWS Bedrock custom LLM with the RAG approach.
- **Designed data models and low-level classes** for efficient chatbot data handling and processing.
- **Utilized object-oriented programming and design patterns** to develop scalable, maintainable codebase.
- **Integrated 5+ REST APIs** into the chatbot for user authentication, bot interactions, feedback, analytics, and export features.
- **Applied SOLID and DRY principles to refactor code**, improving readability and increasing efficiency by 12%.
- **Drove best practices** by actively participating in code reviews, design reviews, and architecture discussions.
- **Experimented with LangChain and Azure OpenAI**, promoting their adoption and measuring impact on project outcomes.
- **Led a 3-member team on an AWS GenAI pilot project**, using React and Python, achieving 30% GenAI adoption.

Technical SKILLS

- *Programming Languages:* JavaScript, Python
- *Front-end Technologies:* React
- *Back-end Technologies:* Flask, FastAPI
- *Generative AI / AI Tools:* AWS Bedrock, AWS CodeWhisperer, Tabnine, Github Copilot, Neo4j
- *AI/ML Libraries and Frameworks:* Numpy, Pandas, PyTorch, scikit-learn, TensorFlow
- *AWS services:* S3, AWS Lambda, AWS Textract, **AWS Bedrock**, DynamoDB, Route53, Cognito, AWS Kendra, AWS Step Function
- *Azure services:* **Azure OpenAI**, Storage accounts, Document intelligences, AI Search
- *Version control:* Git, Github, AWS CodeCommit

EDUCATION

Electronics and Communication Engineering | JIS College of Engineering

Jun 2018 - Jun 2022

GPA: 9.17

- Published **3 Research Papers (1 in the Chemical domain and 2 in the IoT Domain)**
- Patent publication 'A CAR IGNITION SYSTEM BASED ON APPLICATION CONTROLLED BIOMETRIC SENSOR' (**PATENT APPLICATION NO: 202031031645 A**)

AWARDS

- Awarded as the **"Star"** employee by Capgemini in July - September 2023.
- Awarded as the **"Rising Star"** employee by Capgemini in April - June 2023.
- Ranked **top-10** in AWS GenAI hackathon conducted by AWS across **35+ premium partners** by developing a Taxation UI in **React** within a record **2 hours** from scratch using **Amazon CodeWhisperer**.

LICENCES AND CERTIFICATIONS

- Certified **AWS Cloud Practitioner** (License No: HDZQ7MWCJRREQVW7, Issued by AWS)
- Capgemini Certified **Full Stack Developer - Level 1 & 2** (Issued by Capgemini)