

Anuraag Khare

anuraagkhare2121@outlook.com | +919663248149 | [LinkedIn](#) | [Github](#)

Skills

Languages: Python, Go, TypeScript
Databases: PostgreSQL, MongoDB, Redis

Web: Django, FastAPI, Express.js, HTML5/CSS
Tools and Platforms: Linux/Bash, Docker, K8s, GenAI
Cloud: GCP, AWS

Work Experience

HASHEDIN BY DELOITTE | SOFTWARE ENGINEER 2

Bengaluru | Jul 2024 – Present

- Working on backend systems for internal tools used by the sales and support teams of a Leading Global Tech Company.
- Developed a no-code dynamic multi agent creation platform using Google ADK, FastAPI, Redis and PostgreSQL with support for MCP servers.
- Engineered scalable RESTful APIs using Django (ASGI) and FastAPI, leveraging async python capabilities to integrate Generative AI solutions, enhancing one key service with Server-Sent Events (SSE) to significantly reduce response time.
- Optimized backend deployment with a multi-stage Docker CI/CD pipeline, cutting server deployment time by 33% (4.5min to 3min) and build time by 45% (1.5min to 50s).
- Architected and developed scalable backend systems using GCP Cloud Functions, GKE, Pub/Sub, Redis (Memorystore), and PostgreSQL, with WebSocket-based live audio/video chat service supporting large number of concurrent users with auto session recovery.
- Delivered multiple organisation wide tech talks on latest AI technologies like A2A (Agent-to-Agent Protocol) and MCP (Model Context Protocol).

KUTTL | SOFTWARE DEVELOPMENT ENGINEER

Bengaluru | Feb 2024 – Jul 2024

- **Founding Engineer:** Engineered a scalable backend system integrating a RAG-based data pipeline with GenAI-powered invoice processing automation.
- Improved LLM contextual accuracy while minimizing hallucinations, manual intervention, and operational overhead. Tech stack: Python, LangChain, FastAPI, Temporal, AWS.
- Designed and implemented scalable MongoDB schemas and developed RESTful APIs (Express.js, Django) to support complex front-end data operations like fuzzy search, sorting, and filtering over large data.

COMMSCOPE | ENGINEER 1, R&D

Bengaluru | Jul 2022 – Dec 2023

- Contributed to the development of application layer of the 5G stack, for efficient network communication and service orchestration. Tech stack: C++, Python, django, gRPC, AWS
- Engineered a channel selection algorithm for the 5G CBRS Channel Selection Module, enhancing network efficiency and minimizing downtime in comparison to the 4G CBRS system.
- Optimized the Commercial Mobile Alert System (CMAS) feature for 5G systems, enabling support for concurrent alerts.
- Reduced deployment time and increased reliability by nearly 90% through the development of a Django utility app for remote management of k8s pods across multiple on-prem clusters, featuring shell execution and system health monitoring.

PHONEPE | SRE INTERN

Bengaluru | Jan 2022 – Jul 2022

- Cut troubleshooting time by 50% by designing and implementing an interactive on-premise Elasticsearch cluster health monitoring and diagnostics tool using Python, ELK REST APIs, RabbitMQ, remote shell, and a Tkinter-based GUI, enabling faster issue detection and resolution.
- Established a 3 node ELK cluster, integrating Prometheus and Grafana services for enhanced monitoring and email alert using RabbitMQ.

ABB INDIA LTD. | PROJECT INTERN

Remote | Jul 2021 – Aug 2021

- Conducted data analytics research on raw data from assembly belt robotic automatons, leveraging Python, pandas and Matplotlib to provide insights on irregular power consumption and assembly delays.

Publications

PEST IDENTIFICATION AND CONTROL USING DEEP LEARNING AND AUGMENTED REALITY

2022 INTERNATIONAL CONFERENCE ON EDGE COMPUTING AND APPLICATIONS (ICECAA) | Nov 2022

- Developed a **Flutter**-based Android app to help rural farmers detect pests using live camera input, displaying results in **AR** and recommending targeted **control measures**.
- Trained a custom **ML model** via **transfer learning** on a publicly available [pest dataset](#) featuring **9 common species**.

Projects

HELLOPDF

PYTHON, STREAMLIT, LANGCHAIN, FAISS | FEB 2024

- A **RAG interface** to chat with your PDF data and ask queries without using paid OpenAI APIs. Built using **opensource LLMs** from HuggingFace.
- Uses local embeddings for data privacy and low cost. Support for conversation chain history.

Education

B.E in Computer Science and Engineering

Bengaluru, KA | Aug 2018 - Jul 2022

DR. AMBEDKAR INSTITUTE OF TECHNOLOGY