Venkata Siva Naga Sai Aravind Kollipara
Seattle, WA | +17164269215 | aravindkollipara123@gmail.com | [LinkedIn](#) | [GitHub](#)

SUMMARY

Software Engineer with 5+ years of experience in scalable backend systems, cloud infrastructure, and data optimization using Java, Python, and AWS. Recently focused on AI-driven applications, including LLM-based assistants, RAG with ChromaDB, and multi-modal tools using LangGraph, Cohere, and Gemini. Skilled in CI/CD, TDD, and serverless architectures, with a track record of building robust, intelligent solutions and user-friendly interfaces in fast-paced Agile environments.

EXPERIENCE

Software Development Engineer, Amazon | Seattle, United States of America   **10/2022 – Present**
- Designed and implemented scalable cloud-based backend scheduling services using **Java**, while leveraging **CI/CD** pipelines.
- Spearheaded the scaling of the back-end scheduling service, improving the system capacity from **500** to **12000** agents. Optimized **API** response time by **50%** and ensured **99.99%** system uptime through performance tuning and system monitoring.
- Implemented a **back-filling** mechanism to restore missing shift data in **data lake** by reprocessing affected records, ensuring data integrity and preventing future data loss in the scheduling system.
- Led the development of the **Shift Exchange feature**, allowing agents to **swap shifts dynamically**, improving operational flexibility and reducing manual scheduling interventions by **85%**.
- Engineered and optimized **REST API's** with **Node.js** and **Type-scrip**t, by reducing the latency, and enhancing the micro-services.
- Integrated **AWS Kinesis** for **real-time data streaming, enabling** seamless communication between microservices.
- Optimized **DynamoDB** and **RDS** by implementing efficient indexing strategies (GSI and LSCI), and optimizing partition keys.
- Championed **test-driven development (TDD)** and **clean coding practices**, reducing production bugs and improving code quality.
- Drove migration of **Java Development Kit** (JDK) from **JDK8** to **JDK17** and upgraded Cloud Deployment Kit (**CDK**) from **CDK V1** to **CDK V2**, ensuring compatibility and utilizing new features to enhance infrastructure-as-code practices.
- Collaborated with cross-functional teams to prioritize and implement Santander-specific requirements.

Application Development Associate, Accenture | Hyderabad, India   **01/2021 – 06/2021**
- Developed a hotel recommendation application integrated with ratings, awards, proximity, and user preferences, improving recommendation accuracy by **25%**, using **Python, .NET, and GraphQL** on **Google Cloud Platform (GCP)**.
- Implemented real-time data updates and personalized features, collaborating with clients to meet requirements while developing executing test cases to improve test quality and ensure code reliability.
- Achieved a **30%** increase in user satisfaction and a **20%** rise in booking rates through optimized functionality.

Database Developer, BNP Paribas | Chennai, India   **01/2020 – 12/2020**
- Resolved data processing issues, optimized **SQL** queries and database schema for better performance and scalability which resulted in improving system uptime by **9%**.
- Tested and developed scalable backend solutions to efficiently support a larger number of agents.
- Created **Tableau** dashboards to visualize sales and compensation data, boosting reporting efficiency by **20%** and project completion rates by **15%**.

SKILLS

Programming Languages: Python, Java, JavaScript, Golang, C++, .NET, HTML, CSS, C#, .NET
Databases: SQL Server, MySQL, PostgreSQL, Snowflake, Databricks, Redshift, MongoDB, CosmosDB
Frameworks and Tools: ReactJS, Django, Solr, AWS, GCP, RESTful API, JIRA, Git, Angular, Kafka, Kubernetes, Spring Boot, Spring, Docker, Apache Flink, Spark, Google Dataflow, Tableau, Linux, GraphQL, CI/CD Codepipelines, Agent AI, Cohere, Bedrock, Lang chain, Langgraph, LLM, ChromaDB, Ngrok, LangSmith, Gorq, Chatbot, Agents, GEMINI, LLaMA, OpenAI

KEY PROJECTS

AI Assistant Hub on Ngrok – Multi-modal Chat bot with Image Generation
- Built a dual-mode AI assistant using **LangGraph**, **Groq LLaMA 3**, and **Google Gemini**, enabling real-time text-based conversation and prompt-to-image generation with tool-based reasoning via Tavily Search.
- Designed a **Streamlit** interface with **ChromaDB-based RAG** for PDF Q&A, multimodal output, and **Ngrok** for public domain.

ChatGPT-Style Assistant Built on Serverless Architecture of AWS
- Built a **server-less chat-bot** using **AWS Bedrock** and **Cohere LLM** for real-time language generation.
- Assembled Python-based Lambda integrated with **API Gateway** for scalable **RESTful AI** interaction.
- Utilized **Bedrock APIs** to simplify LLM integration and reduce infrastructure complexity.

Neural Networks for Bird Species Classification
- Worked on a **Deep Learning** project using convolutional Neural Networks Methodology for bird species identification, achieving 90% accuracy. Preprocessed audio data through Clustering, Re-Sampling, segment identification, and spectrogram computation.
- Leveraged dropout layers in the model design to prevent over-fitting, enhanced model performance to **93%**.

EDUCATION

Master of Science in Computer Science, University at Buffalo, NY.   August 2021 – August 2022
Bachelor of Technology in Information Technology, SRMIST, India.   August 2016 – May 2020