

Sarthak Kumar

+91-9582243710 | sarthakkumar3110@gmail.com | sarthak kumar | sarthakkumar100903

Education

Motilal Nehru National Institute of Technology, Allahabad

2021 - 2025

B.Tech. in Electronics And communication Engineering

CPI: 8.2

Experience

Member Of Technical Staff Intern

Bengaluru, India

Nutanix

January 2025 - June 2025

- **Tech Stack:** Python, React, Redux, Go-Lang, Calm, Epsilon
- Led development of a Proof of Concept Kubernetes service provider on Prism Central, enabling Helm chart deployment across 5+ multi-cluster cloud environments, significantly reducing manual provisioning overhead and enhancing scalability for C++ cloud infrastructure operations.
- Designed and implemented Nested Runbook support, allowing Blueprints to invoke other Runbooks; this improved workflow modularity by 40
- Integrated Macros, Dynamic, and Global Variables into the React UI and backend to standardize configuration across distributed systems, reducing manual input errors and setup time by 25

Data Scientist Intern

Hyderabad, India

Teliolabs Communication Private limited

May 2024 - August 2024

- **Tech Stack:** Python, NLTK, Faiss DB, Bert, Tensorflow, React, Django.
- Formulated an ML model to predict solutions for Jira Tickets by processing and summarizing large-scale ticket data from Confluence, storing embeddings in a vector database; achieved 85

Projects

Runalytix – Intelligent Runbook Validator and Optimizer Video | Design Doc

May 2025

- **Tools & Technologies:** Python, FastAPI, React, Redux, LangChain, RAG, OpenAI, Nutanix Calm, Epsilon, IDF
- Devised **Runalytix**, an AI assistant for validating and optimizing **Calm Runbook scripts** in real-time, significantly reducing debugging time by **50%** and enhancing the reliability and performance of C++ cloud infrastructure automation.
- Implemented **LLM-based analysis** to detect syntax, logic, and security issues in Shell, PowerShell, and HTTP tasks, achieving **90% issue detection accuracy** critical for maintaining secure, high-performance, and robust cloud services.
- Integrated a context-aware suggestion engine using **LangChain + RAG**, improving fix recommendation accuracy by **35%** and reducing Runbook execution errors by **40%**, contributing to more stable and performant cloud operations and efficient resource management.

Sustainable Traffic Management system Github | Video

Jan. - Apr. 2024

- **Tools & Technologies:** JavaScript, Python, HTML, CSS, Gsheet, SheetDB, P5.js, YoloV4
- Used **Braess's Paradox** to optimize traffic flow, increasing simulated throughput by **30%** in dense areas, demonstrating principles of performance optimization for complex distributed systems, relevant to C++ cloud infrastructure.
- Developed a **custom algorithm** to model congestion across **50+ city layouts**, analyzing flow and bottlenecks, applicable to understanding resource contention and data flow in large-scale distributed cloud infrastructure.
- Designed a **city planning tool** to simulate and evaluate the impact of new road construction in urban layouts, improving traffic efficiency predictions by **35%** and aiding data-driven infrastructure decisions for optimal resource utilization and scalability.
- Leveraged **YOLOv4** for real-time traffic detection from video and trained a routing model to optimize path selection, improving system response speed by **25%** and demonstrating low-latency processing and efficient resource allocation.

Intelligent Chatbot for User Queries Github | Video

Sep. - Dec. 2023

- **Tools & Technologies:** RNN, LSTM, Python, OpenAI (v3.5), NLTK, RASA, Spacy, MySQL, Ngrok, AWS EC2
- Built a context-aware chatbot designed for **MNNIT** and **Botrush** using intent recognition and dialogue management, achieving **75%+ intent classification accuracy** and handling concurrent user interactions as a reliable, scalable service.
- Embedded **OpenAI API (v3.5)** for generating robust fallback responses when user queries fell outside predefined intents, enhancing system resilience and user experience in a distributed service context.
- Deployed the bot on **Telegram**, successfully handling **100+ user queries** via **AWS EC2 t2.micro** using **Ngrok** tunneling, demonstrating experience with cloud deployment, networking fundamentals, and managing services in a Linux environment.

Technical Skills

Programming Languages: C++ (C++14/17/20, Multi-threading, STL, Memory Management), Go, Python, C

Cloud Containerization: Docker, Kubernetes, AWS EC2

Systems Infrastructure: Linux, Systems Programming, Distributed Systems, Computer Networking (TCP/IP, Sockets)

Core CS Performance: Data Structures, Algorithms, Performance Analysis (gdb, Valgrind)

Databases: Cassandra, MySQL, MongoDB

Version Control: Git

Publications

Advance Forecasting of Multivariate Time Series using Block Sampling Technique *Nov. 2024*
Research Paper / IEEE Reference / Presented at IEEE UPCON 2024, SRMCEM, Lucknow, India

- Developed **Advance Block Sampling Method** that enhances time series model accuracy by **over 24%** on multivariate time series data.

Achievements

- One of the winners in **Nutanix Hackathon 2025** in Hybrid Cloud Management
- Secured **415 rank** in **Unstop Amazon ML Challenge 2024**.
- 2nd place in **AtomQuest** event of **TechFest 2023** held in **IIT Bombay**.
- Achieved **LeetCode Knight** status with a contest rating of **2109**, ranking in the top **1.8%** globally..