

NAKUL KUMAR

+91 7899426942 | rsnakul2021@gmail.com | github.com/rsnakul2021 | [linkedin.com/in/nakulrav](https://www.linkedin.com/in/nakulrav) | [nakul/website](https://nakul.website) | Bengaluru, KA

Education

Boston University, *Master of Science in Applied Data Analytics* | Boston, USA **GPA: 3.96/4.00** **Jan 2025**

- Teaching Assistant for graduate courses: Foundations of Machine Learning, Data Mining
- Coursework: Data Analytics and Visualization, Big Data Analysis, Data Science with Python, Advanced Deep Learning

R. V. College of Engineering, *Bachelor of Engineering in Computer Science* | India **CGPA: 9.04/10.00** **Aug 2022**

- Coursework: Data Structures, OS, Algorithms, OOP, DBMS, Data Science, Software Engineering, Parallel Computing

Technical Skills

Programming: Python, R, JavaScript, C, C++, SQL, XML

Machine Learning & AI: LLMs, TensorFlow, PyTorch, LangChain, LangGraph, NLP, Transformers, Gen AI, RAG

Cloud & DevOps: Node.js, AWS, GCP, Snowflake, Databricks, Azure, Docker, CI/CD, MLOps, Git

Data Engineering: Apache Spark, Kafka, ETL, React, SQL Optimization, Data Warehousing, Schema Design

Visualization & Analytics: Power BI, Tableau, RStudio, Data Mining, A/B Testing

Experience

Data Engineer

Aug 2022 – Mar 2023

Mavenir Systems

Bangalore, India

- Developed scalable pipelines and architectures using Apache Kafka for real time transformation of data collected from client's (Airtel and Verizon) network test point access units.
- Boosted network performance by 35%, optimizing real-time data flows with Prometheus and Apache Kafka.
- Optimized ETL processes, ensuring seamless data integration from network sites into structured format for analytics.
- Designed and deployed scalable ETL pipelines using Apache NiFi, enhancing data ingestion and transformation efficiency for real-time network performance analytics, resulting in a 20% reduction in processing latency

Research Intern

Feb 2022 – Aug 2022

Purdue University

West Lafayette, IN (Remote)

- Collaborated with a team of three to develop a SaaS-based natural language processing algorithm, transforming text data to detect vulnerabilities in GitHub workflows and enhance the CICD pipeline efficiency. [\[Research Paper\]](#)
- Researched and engineered AI-driven solutions to detect GitHub vulnerabilities, reducing security threats by 63%.
- Designed and implemented a Spark pipeline to secure GitHub workflows prior to committing changes to user repositories.
- Conducted research on brain tumor detection using pretrained InceptionV3 and V4 convolutional neural networks to enhance diagnostic accuracy. [\[Research Paper\]](#)

Software Engineer Intern

Jul 2020 – Dec 2020

Yokogawa Technology Solutions

Bangalore, India

- Developed a satellite vehicle density detection system in collaboration with the department of Bangalore traffic police, utilizing Scikit-learn and TensorFlow to optimize traffic flow at bottleneck junctions in city.
- Reduced city-wide traffic congestion by 12% by deploying a machine learning-based vehicle density detection system.
- Contributed to the development of a mask detection software during the COVID-19 pandemic by fine-tuning a VGG-16 convolutional network, leveraging IBM's 12-core POWER8 microprocessor for accelerated model training and execution.

Projects

Android Event Log Anomaly Detection Pipeline | *PySpark, Google Cloud Console, Android Studio*

Aug 2024

- Engineered a robust Spark pipeline to transform Android Event Log data into numeric vectors for anomaly detection, ensuring efficient data processing.
- Trained and deployed machine learning models on Google Cloud Platform's Dataproc, leveraging scalable cloud resources for seamless performance.

User-Specific Text-to-Image Generation Model | *Python, Amazon Web Services, Android Studio*

Apr 2024

- Developed and trained a stack generative adversarial network on user-specific datasets to generate images from text input, achieving an accuracy of 89% by experimenting with optimizers like ADAM, AdaGrad, and RMSProp.
- Deployed the model as a fully functional application on an Android platform, integrating a user-friendly graphical user interface.

Healthcare AI Chatbot with Llama and Transformers | *Flask, Llama, Hugging Face*

Sep 2023

- Developed a conversational AI chatbot using Llama for medical query understanding and Hugging Face's Transformers for accurate, empathetic responses. Deployed via Flask to assist patients with appointment scheduling and FAQ's.

Certifications

- Developing LLM-powered applications (LangChain)** - Feb 2025
- Version Control and Code Deployment (Git and GitHub)** - Aug 2024
- Machine Learning Master Course (Python and R)** - Jul 2023
- TensorFlow and Keras for Deep Learning** - Aug 2021