

KARTHIK YELURIPATI

sy493@njit.edu | Newark, New Jersey | Ph no: 8505673687 | [Linkedin](#) | [GitHub](#) | [Portfolio](#) | [X](#) | [Google Scholar](#)

EDUCATION

New Jersey Institute of Technology

Newark, New Jersey

Master of Science in Data Science

Chaitanya Bharathi Institute of Technology

Hyderabad, India

Bachelor of Engineering in Electrical and Electronics Engineering

EXPERIENCE

AI Engineer

Berkeley, California

Floz

July 2024- Present

- Developed a RAG system using **Pinecone** for embeddings, enabling recommendations for 20,000+ construction items.
- Built a multimodal RAG system integrating **computer vision** to analyze architecture layouts and forecast project cost.
- Architected a high-performance, large scale **distributed processing** backend using **Typescript, Cloudflare D1**.

Software Engineer Intern

Englewood Cliffs, New Jersey

NBCUniversal

Sep 2023- Apr 2024

- Implemented test suites for NBC News websites across critical user flows using **Selenium WebDriver & TestNG**
- Conducted API and security testing using **Postman & Java**, enhancing API interfaces and overall application security.
- Analyzed and performed 200+ test cases in **TestRail**, ensuring 100% alignment and resolving security vulnerabilities.

Data Scientist Intern

River Vale, New Jersey

STEERus

Jun 2023-Sep 2023

- Developed a scalable web application for personality traits and talent assessments using **JavaScript** and **AWS**.
- Implemented a normalized database schema on **Amazon RDS** with **MySQL** optimizing concurrent user assessments.
- Performed statistical analysis to develop a novel trait-based scoring system and generated reports using **Tableau**.

Software Developer

Newark, New Jersey

New Jersey Institute of Technology

Feb 2023-Jun 2023

- Collaborated with **Prudential Financial** to build a monitoring system with an internal data pipeline using **React.js**.
- Leveraged **Ansible API** to automate system processes using **Python** and reduced configuration time by 60%.
- Designed Splunk dashboards and enabled real-time monitoring of key performance indicators for internal workflows.

Data Science Research Assistant

Newark, New Jersey

New Jersey Institute of Technology

Nov 2022-Jan 2023

- Applied advanced **NLP** techniques to analyze sentiment and extract themes from large-scale Twitter datasets.
- Implemented data preprocessing and tokenization using **NLTK** and **spaCy**, optimizing training on extensive datasets.
- Achieved 92% accuracy in topic classification using a fine-tuned BERT model and surpassed previous benchmarks.

Programmer Analyst

Hyderabad, India

Cognizant Technology Solutions

Dec 2021- Jul 2022

- Analyzed medical datasets using **PySpark & Hadoop** MapReduce, processing over 200GB of data to derive insights.
- Engineered real-time ETL pipelines using **Apache Airflow** and **Kafka**, reducing data processing latency by 30%.
- Performed large-scale data scraping of **EMR** patient records, life sciences, and hospital data, optimizing ingestions.

RESEARCH PUBLICATIONS

Forecasting electrical demand for the residential sector at the national level using deep learning. *IEEE AIMV '21*. [[Paper](#)]

Speech Enhancement using Deep Learning. *IJTRD '21* [[Paper](#)]

SKILLS & TOOLS

Programming Languages: Python, Java, Go, MySQL, R, Bash | **Deep Learning:** OpenCV, LLMs, Transformers, BERT, GPT, Langchain, NLP, Time Series Analysis, RLHF | **MLOps:** MLflow, Kubeflow, TensorFlow Extended (TFX), Weights & Biases | **AI Frameworks:** JAX, TensorFlow, Keras, PyTorch, Sci-Kit Learn | **Web:** System Design, JavaScript, React, Rest APIs, Selenium, GraphQL | **Big Data:** Hadoop, Hive, Spark, Pig, Kafka, Scala, Airflow | **Cloud Computing:** AWS (Bedrock, EC2, S3, RDS, EMR, SageMaker), Cloudflare | **Misc:** Docker, Kubernetes, Jenkins, OOPs, A/B Testing.