

Isha Kookanapalli Math

ik2688@nyu.edu

+1 (917) 417-3430

LinkedIn : Isha

GitHub : isha962

Education

NYU Tandon School Of Engineering

Sept 2024 - May 2026

MS in Computer Engineering

New York, USA

- Relevant Coursework: Machine Learning, Deep Learning, Data Structures & Algorithms, Big Data.

JSS Institute Of Technical Education

Jan 2020 - May 2024

BE in Computer Science

Bengaluru, India

Skills

Programming: Python, SQL, Java, Kotlin, C/C++

Data Engineering: ETL Pipelines, Data Modeling, Airflow, BigQuery, Redshift, PostgreSQL, MySQL

Cloud Platforms: GCP (Cloud Functions, Cloud Run, BigQuery), AWS (S3, Glue, Redshift, EKS)

Tools & Frameworks: Pandas, NumPy, TensorFlow, PyTorch, Keras, Docker, Kubernetes

Analytics: Data Cleaning, Feature Engineering, Dashboarding (Power BI, Looker Studio)

Experience

Software Engineering Intern

June 2025 – August 2025

Assure AI

Boston, MA, USA

- Architected Neo4j analytics pipelines with Python ETL and graph transformations, improving query performance by 28% and reducing dashboard load latency.
- Created and optimized Cypher queries and integrated NLP techniques for entity recognition and relationship inference, reducing average search time by 30%.
- Reduced failed pipeline runs by 15% by standardizing Neo4j ingestion and validation checks.

Software Engineer Intern

May 2024 – Aug 2024

Donyati

Bengaluru, India

- Developed scalable backend services using Google Cloud Functions and Firebase for real-time processing, improving system performance by 15% and reducing navigation complexity.
- Revamped backend services using Python and Django, increasing throughput by 20% by optimizing API calls and reducing server load.
- Implemented monitoring and alerting for ETL jobs across 5+ scheduled workflows, cutting troubleshooting time by 25%.

AI/ML Intern

Aug 2023 – Sept 2023

Tequed Labs

Bengaluru, India

- Created an image identification system using Python with TensorFlow, Keras, and PyTorch, achieving 95% accuracy on diverse test datasets.
- Streamlined data exploration and performance tracking via custom visualization pipelines, reducing analysis time by 30% and enabling faster model iteration.

Projects

Criminal Identification System Using Facial Recognition

- Implemented an MTCNN + FaceNet recognition pipeline achieving 98% accuracy for real-time identification from video inputs and live camera streams.
- Optimized a MySQL schema for 10,000+ profiles and embeddings, reducing lookup time by 40% and improving matching throughput.
- Improved preprocessing and inference with OpenCV, Keras, and TensorFlow, sustaining 0.5 seconds per frame throughput during real-time execution.

BioLLM Data Preparation Pipeline (Airflow & Dataproc)

- Designed a Dataproc PySpark ETL to clean, LSH deduplicate, and tokenize 500,000+ biomedical abstracts into training-ready LLM datasets.
- Orchestrated a 4-stage Spark workflow using Airflow over SSH, reducing manual runs by 80% and standardizing reruns across stages.
- Validated runs via YARN history and added a Streamlit dashboard with 10+ data-quality checks to verify outputs.

Women Safety App (Android)

- Shipped one-tap SOS alerts that send live GPS to 3+ trusted contacts and trigger emergency notifications.
- Enabled background SOS delivery and continuous tracking via Firebase, achieving 99% alert success in device tests under low-connectivity scenarios.
- Integrated audio and video evidence capture with cloud upload and secure retrieval, supporting recordings up to 60 seconds per incident for verification.