Parth Naik

New York, NY
+1 (646) 732-4576
parthnaik@nyu.edu
linkedin.com/in/naikparth

Education

New York University (NYU) *New York, NY*
M.S. Computer Engineering *Expected May 2026*
Coursework: Big Data, MLOps
GPA: 3.5

Vishwakarma Institute of Information Technology *Pune, India*
B.Tech Artificial Intelligence and Data Science *May 2024*
GPA: 3.7
Coursework: AI, Data Science, Business Intelligence, Visualization

Skills

Analytics: Excel (Pivot Tables, VLOOKUP), Tableau, Power BI
Programming: Python (Pandas, NumPy, scikit-learn), SQL (MySQL, BigQuery), R
Tools: Google Workspace, Microsoft Office, Looker Studio, Miro, Jira

Experience

Research Assistant — NLP & Knowledge Systems *Aug 2023 – Dec 2023*
Vishwakarma Institute of Information Technology, Pune, India

- Conducted data collection, preprocessing, and entity extraction from Sanskrit and English texts related to Indian mythology.
- Developed a question answering (QA) system using NLP techniques including named entity recognition and document similarity scoring.
- Integrated a domain-specific knowledge base and achieved over 80% accuracy on custom evaluation queries.

Projects

US-101 Vehicle Trajectories – Traffic Flow Analysis (NGSIM) *Summer 2025*

- Analyzed over 2.2 million vehicle trajectory records (1.4 GB) from the NGSIM US-101 dataset to uncover traffic bottlenecks and driver behavior trends.
- Optimized BigQuery storage using partitioned tables, reducing data query latency by 85%.
- Discovered 20% higher congestion in center lanes during peak hours by analyzing velocity and vehicle density patterns.
- Detected 35,000+ lane changes, which revealed a 42% increase in aggressive driving during rush hour.

NYC Vision Zero – Real-Time Risk Scoring Platform *Spring 2025*

- Built FastAPI-based REST API for real-time traffic risk classification (Safe to Dangerous), deployable on Raspberry Pi 5 edge devices.
- Containerized services and deployed to Kubernetes (KVM@TACC); implemented staging, canary, and production pipelines using Argo Workflows.
- Integrated Prometheus and Grafana for live monitoring of latency, drift, and classification metrics, reducing false positives by 15%.
- Enabled weekly retraining via Ray and MLflow, leveraging MinIO for model artifact tracking and automated evaluation.

Certifications

- **Google Data Analytics Professional Certificate** – Coursera *March 2024*