

Geet Purohit

☎ 973-814-0238 github.com/geetpurohit [linkedin.com/in/geetpurohit](https://www.linkedin.com/in/geetpurohit) ✉ geetp40@gmail.com

EDUCATION

Rutgers University

B.S. Computer Science

B.S. Astrophysics

Dec 2022

GPA: 3.7/4.0

TECHNICAL SKILLS

Languages: Python, Java, JS/TS, HTML/CSS, L^AT_EX, SQL

Developer Tools: Docker, MySQL, PostgreSQL, MongoDB, KNIME, DataRobot

Technologies/Frameworks: Git, Apache Spark, TensorFlow, Tableau, QlikSense

EXPERIENCE

Verizon | *Data Scientist*

Jan 2023 - Present

- Architecting a **credit risk** based fullstack app, **synthetically generating** a customer database to monitor credit policy changes and their impact to **bad debt writeoffs & involuntary churn**
- Engineered **9 Python & SQL anomaly detection microservices** with Verizon EDW integration, ensuring **CPNI data security** through IAM with Gitlab CI/CD impacting **1000s of users across 15 teams**, saving **>50 hours** every month
- Led and architected **2 ML Python XAI frameworks** with **Streamlit/Qlik** frontend to monitor isolated real-time contributions to Verizon Gross Adds, reducing hypothesis generation time by **73%** for shareholder calls
- Created an **automated serverless forecasting of 9 major value segment brands** like **Walmart, Tracfone, Straight Talk, FiOS**, pioneering a new **12h email report** system impacting **3 organizations**
- Hosted **Verizon 2023 Analytics Summit** highlighting multiple breakthroughs in the DS industry and recent developments with AI, among **prominent executives and shareholders** from Verizon, Bayer, BNY Mellon, NY Life Insurance, Thoughtspot, etc.

Verizon | *Data Science Intern*

Jun 2022 - Aug 2022

- Established a centralized **data science pipeline framework/API** for different Corporate Finance teams using Gitlab, Domino Data Lab, and GCP, making the collaboration process **63% more efficient**
- Developed a real-time pipeline using above framework to automate **AT&T and TMobile portins** EDA through D-Tale, train **DataRobot ML models**, and forecast competitor market share trajectories

Human Genetics Institute of New Jersey | *Data Engineering Intern*

Sep 2021 - May 2022

- Spearheaded training of DeepLabCut Machine Learning models to **extract pose estimation** and coordinate interpolation data from **>1000 wet-lab mice experiments** using Python, Tensorflow, and NVIDIA CUDA toolkit
- Developed **4 Oxford IMARIS Python extensions** to analyze spinal cord based injury segmentation and automate the 3D reconstruction of lesion sites making **post-microscopy analysis 5x faster**, aiding in discovering cures for spinal cord injuries

PROJECTS

D.A.P.S Link | *Python, End-to-End DS Pipeline (ETL, EDA, Fcst, ML, Client)*

Nov. 2022

- An automated pipeline that can establish an RDBMS connection, analyze data-drift, visualize EDA/ETL insights, engineer time-series features, train an SKLearn machine learning model, and forecast data

Are You Blinking? Link | *Python, Computer Vision*

May 2022

- Trained an Oriented Gradient + Linear SVM based Eye Aspect Ratio model to detect blinks, categorize them into Morse Code, and translate into English in real-time through a live camera feed
- Inspired from Jeremiah Denton, a Vietnam PoW who used blinking to translate key information to U.S. Naval Intelligence

A * is Born Link | *Python, Graph Search, School Project*

Dec. 2021

- Implemented 4 graph traversal algorithms such as BFS, DFS, and Repeated A* for path finding in random 50 x 50 maze settings, analyzed their behaviors under several configurations, and presented their run-time efficiencies