





# Atmika Pai

aap253@cornell.edu | New York, NY | +1 (510) 365 – 0290 |     @atmikapai13

## EDUCATION

**Cornell University (Cornell Tech)**, New York, NY

Aug 2024 – May 2026

**Master of Science in Information Systems, Conc. in Urban Technology** | **GPA: 3.94**

*Fellowships:* Related Companies Fellow (Merit), Emerging Markets Institute Fellow (Johnson College of Business)

**University of California, Berkeley**, Berkeley, CA

Aug 2018 – May 2022

**Bachelor of Arts in Data Science and Economics**

## TECHNICAL SKILLS

**Coding Languages:** Python, SQL, JavaScript, R, STATA, HTML, CSS

**Technical Tools:** Jupyter, SAS, MySQL, Git, Snowflake, Microsoft Suite, Oracle, Figma, Tableau, Unity, GravitySketch

**GIS Tools:** ArcGIS Suite (Online, Pro, StoryMaps), QGIS, GeoPandas, Google Maps API, Folium

**Data Science:** Data Wrangling & Preprocessing, ETL, Data Analysis, Data Visualization, Database Management & Design, Supervised/Unsupervised Learning, Time-Series Forecasting, Natural Language Processing, Computer Vision

## PROFESSIONAL EXPERIENCE

**Tata Consultancy Services**, Edison, NJ

Summer 2025

AI & Data Graduate Intern – TCS Ai.Cloud Business Unit (North America)

- Developing AWS QuickSight app integrating ML analytics to drive real-time, data-driven decision-making for insurtech client

**CSAA Insurance Group, a AAA Insurer**, Walnut Creek, CA

Jun 2022 – Jun 2024

Product Analyst II – Personal Lines Auto & Property Insurance (Arizona, Nevada, Utah)

- Developed AZ and UT's 3-year financial budget using microeconomic factors and historical KPI trends in Oracle Essbase
- Fine-tuned geographic risk segmentation in pricing algorithms using SQL, Snowflake, and InsurQuote competitive intelligence tool
- Built Tableau dashboard to visualize zip code-level changes, helping insurance agents effectively target low-risk profiles in all 3 states

Product Analyst I – Personal Lines Auto & Property Insurance (Arizona)

- Led monthly product performance meetings to inform senior PMs, RVPs, and actuaries of KPI results compared to financial plan
- Unified data retrieval logic across two disparate data warehouses in SQL, enhancing data integrity and delivering accurate reporting
- Leveraged unified logic to create novel PowerBI report on consumer behaviors in digital sales, and shared monthly with senior PMs

Product Management Intern

Summer 2021 – 22

- Summer 2022:* Developed business strategy for specialized Electric Vehicle insurance coverage using Porter's Five Forces framework; partnered with actuaries and product specialists, earning leadership recognition for rigorous roadmap and cross-functional execution
- Summer 2021:* Identified +5.5% sales growth opportunity by analyzing competitiveness and market share geo-correlations using SQL

## PROJECTS

**NYC AlgoMatcher** (React/Next.js, Tailwind CSS, Framer Motion, Vercel)

- Built an interactive survey tool and website to help NYC residents identify which city agency algorithms (as reported under Local Law 35) may process their data; mapped survey logic to 46 algorithms used by 14 agencies and deployed via Vercel

**Network Analysis of NYC Subway Accessibility** (ArcGIS Pro, ArcGIS Story Maps, ESRI GDB)

- Visualized NYC Subway network architecture (network, nodes, flows, and agents) using GTFS data and ArcGIS Pro's Network Analyst Toolkit; published dead transit zones heat maps and isochrone interactive maps via ArcGIS StoryMaps

**Predictive Maintenance of NYC Subway Disruptions** (Supervised/Unsupervised Machine Learning, Python, QGIS)

- Designed predictive maintenance framework using linear models, tree-based algorithms and K-Means clustering to quantify disruption probabilities by disruption type, passenger delays by subway line, and deriving high-impact subway lines
- Presented white paper at Open Data Week to NYC Open Data Team for prioritized MTA intervention

**Disaster Tweet Classification using Natural Language Processing** (NLTK, Sci-Kit Learn)

- Engineered NLP model to classify tweets as disaster-related or not, achieving 89.2% test accuracy on test set
- Developed preprocessing pipeline (stopword removal, vectorization, lemmatization) and applied feature extraction (Bag of Words, N-Grams) to prepare large-scale tweet datasets for training Logistic Regression and Bernoulli Naive Bayes ML models

**IntelliBART Database and Ridership Analytics** (RDBMS, SQL, MySQL, Time Series Forecasting)

- Database Creation:* Independently built *IntelliBART*, a relational database management system (RDBMS) integrating Bay Area Rapid Transit (BART) ridership data with GTFS geospatial files; engineered efficient ETL pipeline and robust database schema
- Analytics:* Forecasted ridership to gauge potential return to pre-pandemic levels using SARIMA & Prophet time-series methodologies

## PROFESSIONAL DEVELOPMENT

**Student Assistant, Esri Developer Summit & User Conference**, California, CA

- Supporting event logistics for Esri Developer Summit (Mar 2025, Palm Springs) and User Conference (Jul 2025, San Diego)
- Gained exposure to emerging GIS technologies and geospatial applications of AI and LLMs

**Graduate Research Assistant, Urban Tech Hub and Cornell AAP**, New York, NY

- Pioneered data-driven sustainable tree-planting strategies using QGIS, directly influencing decisions made by New Yorkers for Parks

## LANGUAGES & INTERESTS

**Languages:** Konkani, English, Hindi | **Interests:** Marathoner, film aficionado