Sanket Navnath Janger

+1 (609) 302-1913 | sjanger@binghamton.edu | LinkedIn | GitHub | Binghamton, NY

Education

State University of New York at Binghamton *Binghamton, NY*
Master of Science in Computer Science (AI Track) *Expected Graduation - May 2026*

Savitribai Phule Pune University *Pune, India*
B.E. in Computer Engineering (Honors: Data Science) *Aug 2020 – Jul 2024*

Technical Skills

Languages: Python(Proficient), SQL, Java, C#
Data and Analytics: NumPy, Pandas, Scikit-learn, PyTorch, Matplotlib, Seaborn
Databases and Cloud MySQL, PostgreSQL, AWS (S3, Athena, IAM, EC2, Lambda), GCP, Github Actions, Docker, Kubernetes
Visualization: Power BI, MS Excel(Pivot Tables, VLOOKUP, SUMIFS/COUNTIFS, Power Query)
Foundations and Operating System: Data Structures and Algorithms (DSA), Object Oriented Programming (OOP), Agile, Windows, MacOS, Linux,

Experience

Data Science Intern — Global Health Impact Project *Binghamton, NY May 2025 – Aug 2025*

- Engineered end-to-end Python ETL pipelines using Pandas to ingest, reconcile, and standardize data from disparate sources, reducing manual processing and improving reproducibility.
- Implemented modular QA checks for data validation and deployed time-series/ML forecasting models to estimate missing values and project short-term health index trends.
- Developed automated "entity charts" and supported front-end tasks using TypeScript/JS to surface key status indicators to stakeholders via Git-based collaboration.

Software Engineering Intern — Emerging Technologies *India Feb 2023 – Mar 2023*

- Developed a suite of full-stack applications using C# and .NET, focusing on backend logic and real-time data processing.
- Engineered a dynamic ID Card Generator featuring a drag-and-drop UI and a backend system for automated data mapping.
- Implemented Firebase for cloud data storage and user authentication across multiple mini-projects.
- Built a Voice-to-Text/Text-to-Voice Translator utilizing external APIs and server-side scripting to handle audio data processing.

Academic Projects

NYC Yellow Taxi Real-Time Analytics Platform

- Built a scalable analytics pipeline using Python + AWS (S3, Athena) to process millions of trip records, enabling fast SQL exploration and time-series forecasting.
- Developed a FastAPI backend and a Next.js frontend using the Vercel AI SDK to enable natural-language queries and deliver real-time insights on demand, fare, and zone performance.
- Created Power BI dashboards and automated deployments with GitHub Actions CI/CD to streamline reporting and ensure repeatable releases.

Ensemble Learning for Medical Risk Prediction

- Engineered an ensemble that fuses CNN-derived representations with gradient boosting for risk prediction on medical data, improving robustness across patient groups.
- Designed experiments with cross-validation, hyperparameter tuning, and benchmarking against single-model baselines; reported 0.72 AUC-ROC.
- Packaged the inference pipeline for fast scoring (preprocessing + model loading + prediction), delivering sub-second local inference per request.

Retail Optimization Dashboard

- Built SQL datasets (CTEs/joins/aggregations) for sales KPIs, stockouts, and reorder signals.
- Developed Power BI dashboards with trends and drilldowns by store/product/time.
- Added demand forecasting views to support inventory planning.

Customer Segmentation via Clustering

- Implemented K-Means and unsupervised learning workflows
- Automated pattern discovery in transactional data
- Applied clustering evaluation metrics and interpretability methods