

Michael Paglia

Troy NY, 12180, United States | 518-852-7364 | michaelpagliadev@gmail.com
mpaglia.com | linkedin.com/in/michaeljpaglia | github.com/michaelpaglia

EDUCATION

University at Albany

Albany, NY

Bachelor of Science in Computer Science

May 2024

- Graduated *summa cum laude* (GPA: 3.85/4.0)
- Member of Phi Beta Kappa

EXPERIENCE

Lead Data Analyst, AI Agents

Oct. 2024 – Present

Meta

New York, NY

- Manage multiple data annotation and labeling projects end-to-end, collaborating with researchers and cross-functional stakeholders to enhance AI systems powering core Meta products
- Deploy specialized knowledge experts to optimize data quality for AI agents powered by Llama 4, driving improvement in synthetic data inputs and achieving 90% research team approval rating across seven-figure weekly budget operations
- Lead and mentor 30+ analysts on scalable data methodologies while owning KPIs and timelines across Product Data Operations' generative AI capabilities

Founder

Mar. 2025 – Present

518 Cares (Nonprofit)

Troy, NY

- Founded healthcare access nonprofit serving 50K+ population in Troy, NY community with web-based resource database and eligibility screening tool at 518cares.org
- Developed platform aggregating free and low-cost healthcare services with privacy-first architecture

Founder

Nov. 2020 – Present

M.P. Mercantile, LLC

Troy, NY

- Founded a start-up e-commerce business hosted on 3 major online marketplaces with 1K+ unique customers
- Achieved over six figures annually in gross sales revenue and 25% growth year-over-year

Data Analyst, Generative AI

May 2024 – Oct. 2024

Meta

Remote

- On assignment via Magnit, conducted over 1K+ quality assurance tests weekly on code generated by Llama 3
- Collaborated with project managers to improve vendor performance to yield a 15% increase in labeling accuracy

National Science Foundation Research Fellow

Jan. 2024 – May 2024

The Research Foundation for SUNY (Data Management and Mining Lab)

Albany, NY

- Implemented and tested 10+ machine learning algorithms for temporal graph and multi-dictionary signal decomposition, achieving a fit of over 98% to datasets with 20M+ points for data reconstruction
- Integrated custom graph and time-series dictionary generation to reduce data retrieval time by nearly 90%

Software Engineering Intern

May 2023 – Aug. 2023

Regeneron Pharmaceuticals, Inc.

Rensselaer, NY

- Developed and deployed a CRUD application and REST API used by IT team members, improving operational efficiency and saving 2 hours of manual effort each week
- Provided smart contract-based access management using Python, Flask, JavaScript, and Dremio Data Lake, achieving a 20% improvement in secure data access verification times across an organization of 4K+ employees

PROJECTS

PySpady | *Python, NumPy, TensorLy, Terraform, AWS S3, Git* | <https://bit.ly/pyspady>

Sep. 2023 – May 2024

- Features data compression, missing value imputation, future value prediction, and autocompletion
- Machine learning determines the best combination of dictionaries, models, and optimizers for a given dataset

TECHNICAL SKILLS

Languages: Python, Java, JavaScript, SQL, C++, C#, C

ML/Data: pandas, NumPy, TensorLy, Matplotlib, Seaborn

Web & APIs: Flask, Django, Node.js, Express, React, PostgreSQL, HTML/CSS

Tools: Git, Docker, GCP, Dremio, Railway, Vercel, VS Code