Contact

glunetta02@gmail.com

www.linkedin.com/in/giovanni-lunetta (LinkedIn)

Top Skills

Data Science
Prompt Engineering
LangChain

Certifications

Databricks Fundamentals
LangChain Chat with Your Data
LangChain for LLM Application Development
Johnson & Johnson Design Thinking Challenge
GenAIBIZ

Honors-Awards

2021 New England Scholar
2022 Babbidge Scolar
ASA Datafest Best Statistical Insight

Giovanni Lunetta

Data Scientist | Healthcare AI | TLDP @ J&J
Pompton, New Jersey, United States

Summary

I build production AI and data systems in healthcare — GenAI platforms, ML pipelines, and cloud automation that drive real business outcomes.

At J&J (Technology Leadership Development Program, Data & Decision Science), my work spans GenAI document intelligence, unsupervised preference modeling for surgical workflows, enterprise AWS cost optimization, and ML pipeline validation at scale. I care about shipping things that get adopted, not just built.

Outside of work, I build too: giovannilunetta.com runs a self-evaluating dual-LLM system I architected, and MemSavr is a macOS app I shipped with Apple code signing. M.S. in Data Science from UConn.

Always happy to connect with people working on applied AI — let's talk.

Experience

Johnson & Johnson

Technology LDP Analyst
June 2024 - Present (1 year 9 months)
Raritan, New Jersey, United States

GenAI Document Automation: Built a document intelligence platform using hybrid retrieval (pgvector + BM25), LLM agents, and schema-based extraction — achieving 50% token reduction and enabling faster, more reliable regulatory document processing.

Pre-Surgical Configuration Analytics: Designed unsupervised preference models (Fuzzy C-Means) achieving 67% surgical classification (+17 over baseline). Delivered Tableau dashboards adopted for clinical decision support.

AWS Cost Optimization: Architected a 3-Lambda event-driven pipeline eliminating 200+ hours/year of manual processing, with multi-million dollar cost attribution at 85%+ coverage.

ML Pipeline Validation: Validated a 14-node Kedro ML pipeline (1.3M+ records, XGBoost, ROC-AUC 0.717) with scalability assessment across multiple therapeutic indications.

HCP Segmentation: Built provider segmentation models (90+ features, K-Means/Fuzzy C-Means) and treatment-pattern analytics to inform commercial targeting strategy.

Microsoft 365 Copilot Program: Managed enterprise pilot for 877 users with automated lifecycle workflows (Python + Excel macros), reducing processing time by 50%.

Johnson & Johnson
Technology Intern
May 2023 - August 2023 (4 months)
Titusville, New Jersey, United States

Developed a machine learning model using healthcare claims data to predict patient drug eligibility 6 months in advance, optimizing doctor outreach targeting and pharmaceutical processes.

Education

University of Connecticut
Master's degree, Data Science · (August 2023 - June 2024)

University of Connecticut
Bachelor of Arts - BA, Statistics & Political Science · (September 2020 - May 2023)