Matt Oremland, PhD

Director of Data Science

mattoremland@gmail.com • (518) 560-1473 • Saratoga Springs, NY 12866
https://www.linkedin.com/in/matthew-oremland-31a312101

PROFESSIONAL SUMMARY

Data science and machine learning leader with PhD in Mathematics and 10+ years of experience building and deploying production ML systems in healthcare and pharmaceutical organizations. Proven track record architecting ML systems from problem framing through deployment, monitoring, and iteration with end-to-end ownership. Deep expertise in machine learning fundamentals including ensemble methods, deep learning, and LLM-based systems deployed in regulated healthcare environments. Strong people leadership experience building and mentoring high-performing teams with focus on responsible AI, model quality standards, and real-world deployment. Exceptional stakeholder management and communication skills translating complex technical concepts for executive audiences and driving measurable business impact.

PROFESSIONAL EXPERIENCE

Redica Systems
2024 – Present

Director, Data Strategy & Analytics — *Remote*

- Owned and evolved data science strategy for pharmaceutical clients (Eli Lilly, Merck) translating company objectives into clear, prioritized roadmap with measurable business impact—built production ML systems driving regulatory intelligence and quality optimization
- Architected and deployed production LLM classifier (OpenAI API) with end-to-end ownership from problem framing and data strategy through deployment, monitoring, and iteration—system processes unstructured healthcare regulatory text in production serving multiple client use cases
- Set technical direction and standards for production ML systems including validation frameworks (confusion matrices), bias testing, performance monitoring, and responsible AI practices ensuring models were appropriate for healthcare regulatory use and real-world deployment
- Partnered closely with client stakeholders including executives, quality leaders, and data engineering teams to embed data science into operational workflows and decision-making—represented data science in strategic planning forums communicating tradeoffs, risks, and impact
- Connected technical work to business outcomes by building risk-scoring models that enabled clients to optimize resource allocation, reduce compliance risks, and improve operational efficiency—demonstrated tangible ROI through cost savings and risk mitigation
- Championed responsible and interpretable AI by implementing comprehensive documentation of model behavior, limitations, explainability approaches, and audit trails ensuring transparency and regulatory compliance in healthcare applications

Gilead Sciences
2023 – 2024

Director, Data & Analytics (Medical Affairs) — *Foster City, CA*

- Led and scaled Data Science and Analytics organization of ~5 people including analysts and dashboard developers with clear expectations, ownership, and career development—owned hiring, onboarding, and talent development building high-performing team
- Owned data science and analytics strategy across Medical Affairs organization translating business objectives into prioritized roadmap balancing near-term delivery (dashboard releases, data quality improvements) with longer-term platform investments (data lake architecture, governance frameworks)
- Architected enterprise data platform and ML-ready infrastructure on cloud supporting analytics and machine learning workloads—built scalable data pipelines, implemented data quality controls, and established MLOps-adjacent best practices for analytics delivery

- Partnered with Product, IT, Medical Affairs leadership, and cross-functional stakeholders to embed analytics into strategic decision-making and operational workflows—influenced executive stakeholders on data strategy and capability investments
- Demonstrated ability to connect technical work to business outcomes by building analytics capabilities that improved HCP/HCO engagement tracking, enabled data-driven medical strategy, and optimized resource allocation across therapeutic areas
- Established model evaluation and monitoring practices for analytics systems including data quality scorecards, performance dashboards, and feedback loops ensuring reliable systems that stakeholders depended on

Takeda Pharmaceuticals 2021 – 2022

Associate Director, Digital & Data Science *Lexington, MA*

- Led data science team driving ML model development and deployment for manufacturing and quality operations—owned strategy, prioritization, and delivery across predictive analytics, real-time monitoring, and optimization use cases
- Set technical direction and standards for production ML systems including ensemble methods, PLS, and PCA deployed for quality risk detection, process optimization, and automated limit calculations—established model validation, version control, and monitoring practices
- Architected ML systems with end-to-end ownership from problem framing (identifying quality risks, optimizing processes) through data strategy (pipeline design, feature engineering) to deployment (production integration) and monitoring (performance tracking, drift detection)
- Deep understanding of model evaluation and real-world feedback loops—implemented continuous monitoring of model performance, established retraining triggers, and iterated based on manufacturing outcomes and quality metrics
- Partnered with Manufacturing, Quality, Supply Chain, and IT leaders embedding data science into operational decision-making—represented data science in strategic forums communicating technical tradeoffs and business impact to senior stakeholders
- Connected ML systems directly to business outcomes including reduced batch failures, improved yields, optimized resource allocation, and cost savings—demonstrated measurable ROI and margin impact from predictive analytics

Tidal Wave Analytics LLC 2023 – Present

Consulting Data Scientist *Saratoga Springs, NY*

- Built and deployed production random forest model on Databricks (AWS SageMaker-adjacent) for Paradigm—owned end-to-end ML system from problem definition through production deployment with nightly predictions serving operational workflows
- Architected RAG system using PyTorch and HuggingFace processing unstructured text data—implemented data pipelines, model inference, and monitoring demonstrating experience with NLP and production ML systems
- Demonstrated ability to work in high-growth startup environments delivering ML systems rapidly while maintaining quality standards—partnered directly with VP of Product and engineering teams balancing speed with rigor

Regeneron Pharmaceuticals 2016 – 2021

Senior Process Data Scientist *Rensselaer, NY*

- Developed and deployed production ML systems (ensemble methods, PLS, PCA) for biopharmaceutical manufacturing with demonstrated business impact—automated limit calculations reduced manual effort by 80%, feed optimization improved yields, quality prediction reduced batch failures
- Architected automated data pipelines and ML infrastructure processing large-scale manufacturing data—implemented ETL processes, feature engineering, model training workflows, and deployment pipelines establishing foundation for scalable ML operations
- Deep expertise in model evaluation, monitoring, and experimentation—designed validation frameworks, conducted A/B testing on process parameters, implemented performance monitoring, and iterated models based on manufacturing outcomes and quality feedback
- Exceptional communication skills translating complex ML concepts for non-technical manufacturing and quality stakeholders—presented to senior leadership, influenced strategic decisions, and drove adoption of predictive analytics across organization

EDUCATION

Doctor of Philosophy (PhD) in Mathematics 2014

Virginia Polytechnic Institute & State University

Dissertation: Optimization and control methods for agent-based models in biological systems. Research involved machine learning algorithms, optimization methods, and computational modeling.

SKILLS & TECHNOLOGIES

Machine Learning & AI: Production ML Systems • Deep Learning • NLP • LLM/GenAI • Ensemble Methods • Random Forest • PLS/PCA • PyTorch • TensorFlow • Scikit-Learn • HuggingFace • OpenAI API

MLOps & Infrastructure: ML System Architecture • Model Deployment • Model Monitoring • Performance Tracking • Version Control • AWS SageMaker (via Databricks) • Cloud Infrastructure • Data Pipelines • Feature Engineering

Data Platforms & Tools: Python (Advanced) • SQL • Snowflake • Databricks • AWS • Pandas • Data Warehouses • ETL/ELT • Big Data Processing

Healthcare & Regulated Environments: Healthcare Data • Pharmaceutical Operations • Regulatory Compliance (FDA, GxP) • Responsible AI • Model Validation • Audit Trails • Explainable AI • Data Governance

Model Development & Evaluation: Model Evaluation Frameworks • Experimentation & A/B Testing • Real-world Feedback Loops • Bias Testing • Performance Monitoring • Drift Detection • Validation Protocols

Leadership & Strategy: Team Leadership & Development • Hiring & Onboarding • Technical Direction • Strategy Development • Portfolio Prioritization • Executive Communication • Stakeholder Management • Cross-functional Partnership

PRESENTATIONS & PUBLICATIONS

- CHPA Regulatory, Scientific & Quality Conference (2025) — Solo Presenter: The Impact of AI on FDA-Regulated Industries
- Redica Systems Webinar (2025) — Solo Presenter: From Reactive to Proactive: Leveraging AI/ML to Improve the Quality Unit

PUBLICATIONS

- "Optimization and control of agent-based models in biology: a perspective" — Bulletin of Mathematical Biology, 2017
- "A computational model of invasive aspergillosis in the lung and the role of iron" — BMC Systems Biology, 2016