Kian Maleki — Data Scientist | Machine Learning | MLOps | Physics PhD

+1 (737) 600 9803 | kian@drkianmaleki.com | drkianmaleki.com | linkedin.com/in/kian-maleki-phd | github.com/drkianmaleki
University of Iowa (2026) PhD in Physics (GPA: 4.03) | Creighton University (2017) MS in Physics (GPA: 4.00)

Technical Experience

TripleTen | AI & Machine Learning Program 2026

- Supervised and unsupervised learning with Scikit-learn for real-world applied machine-learning tasks.
- Deep learning, computer vision, NLP, transformer-based workflows, model evaluation, SQL, and Git.
- Completed 15+ end-to-end ML projects covering classification, regression, RAG, MLOps, and deployment workflows.

Extern, Pfizer | AI-Powered Document Intelligence Externship 2026

- Built an AI-powered document intelligence pipeline to automate processing of healthcare documents using native PDF text extraction, LLMs, and RAG architectures.
- Developed retrieval-augmented pipelines to improve accuracy and reduce hallucinations in document-based question-answering systems.
- Delivered a functional prototype demonstrating how AI-driven automation improves efficiency in healthcare supply-chain operations.

Technical Expertise

- **Modeling:** Gen AI/RAG, experiment tracking (MLflow), data versioning (DVC), drift monitoring (Evidently, PSI), HPC, Monte Carlo simulation, Bayesian statistics.
- **Engineering:** Python, SQL, FastAPI, Pydantic, Docker, Linux/Bash, Git, GitHub Actions, Jenkins, CI/CD.
- **Numeric Stack:** Pandas, NumPy, Polars, Numba, Scikit-Learn, XGBoost, Transformers, TensorFlow, PyTorch.
- **DevOps:** pytest, Pydantic validation, AWS S3/EC2, system health monitoring, reproducible pipelines.

Data-Driven Projects

Sequence Acceleration Benchmark for Gradient Boosting 2026

Tech: Python, XGBoost, Scikit-Learn, NumPy, Statistical Testing, Reproducible Benchmarking
link: https://github.com/drkianmaleki/sequence-acceleration-benchmark

- Developed a reproducible benchmark for finite-horizon learning-curve prediction in gradient boosting, evaluating 51 sequence-acceleration methods from 13 families across 18 synthetic convergence regimes.
- Designed failure-detection, perturbation-diagnostic, and conservative extrapolation rules for identifying when learning-curve forecasts are reliable enough for operational use.
- Submitted the associated paper, "Finite-Horizon Learning-Curve Prediction for Gradient Boosting: Regime Dependence, Failure Detection, and Conservative Extrapolation Rules," to *Machine Learning*.

Ambiguity Range Framework 2026

Tech: Python, Scikit-Learn, XGBoost, PyTest, Statistical Testing, Model Evaluation
link: https://github.com/drkianmaleki/ambiguity-framework

- Developed a model-agnostic diagnostic toolkit for evaluating binary classifiers beyond global metrics such as AUC-ROC.
- Implemented ambiguity metrics to quantify prediction concentration near decision boundaries and threshold-sensitive degradation in precision, recall, and F1.
- Built reproducible experiments across real-world datasets with calibration analysis, bootstrap confidence intervals, paired statistical tests, and automated test coverage.

Heart Disease Prediction: End-to-End MLOps Pipeline 2026

Tech: DVC, Git, MLflow, PyTest, GitHub Actions, Evidently
link: https://github.com/drkianmaleki/Heart-Disease-Prediction

- Built an end-to-end MLOps pipeline for heart disease prediction with reproducible data versioning, experiment tracking, model registry, and CI/CD validation.
- Automated testing with GitHub Actions and pytest, enforcing model-performance thresholds before deployment.
- Implemented data and concept drift monitoring with Evidently to support proactive retraining triggers.

PDF Document Intelligence Extractor 2026

Tech: Python, NLP, Regex, Pandas, PDF Text Extraction, RAG
link: https://github.com/drkianmaleki/PDF-Document-Intelligence-Extractor

- Built a document intelligence pipeline to extract key regulatory information from PDF documents using native text-layer parsing

rather than OCR.
- Developed rule-based and NLP extraction methods to identify, normalize, and structure information from semi-structured formatted text.

RAG-Powered Knowledge Assistant 2026

Tech: Python, PyTorch, Transformers, NLP, Docker

link: https://github.com/drkianmaleki/RAG-Powered-Knowledge-Assistant

- Built a local retrieval-augmented generation system for document question answering with context-aware responses using Transformers and PyTorch.
- Designed a sliding-window chunking strategy with overlap to preserve semantic continuity across long documents.

Professional Experience

Creighton University | Assistant Professor of Physics Aug 2025 – Present

- Teach undergraduate physics courses (40–180 students), delivering lectures on optics and electronics with a focus on applied problem-solving.
- Designed and implemented course curricula, syllabi, assignments, and lab materials aligned with departmental standards.
- Supervised and mentored 1–5 teaching assistants, improving instructional quality and consistency in grading.
- Developed interdisciplinary lab curriculum integrating AI tools and chatbots into scientific workflows.

University of Iowa | Graduate Research Assistant, Physics Aug 2018 – Aug 2025

- Optimized large-scale numerical simulations, reducing computational runtime from days to under one hour through algorithmic improvements.
- Developed high-dimensional modeling frameworks for complex physical systems using statistical and numerical methods.
- Analyzed high-dimensional experimental data to extract signal from noise using statistical inference and optimization.

Creighton University | Graduate Research Assistant, Physics Aug 2015 – Aug 2018

- Developed simulation-based models for physical systems using Python and Fortran, including N-body simulations and molecular dynamics.
- Engineered parallelized computation pipelines for large-scale simulations using HPC resources.

Selected Publications

- Maleki, K. (2026). "Finite-Horizon Learning-Curve Prediction for Gradient Boosting: Regime Dependence, Failure Detection, and Conservative Extrapolation Rules." *Machine Learning*, submitted.
- Maleki, K. (2026). "The Ambiguity Range Framework: A Diagnostic Toolkit for Operational Evaluation of Binary Classifiers." Manuscript under review.
- Maleki, K. et al. "Crystal fields, exchange and dipolar interactions..." *Physical Review B*, 2025.
- Maleki, K. et al. "A General and Modular Approach to Solid-State Integration..." *Nano Letters*, 2025. Related to a patented quantum device.