Achyuth Madabhushi

achyuth.madabhushi@pm.me
www.achyuthmadabhushi.com
github.com/amadabhu
Rockville, MD

Skills

Python · SQL · scikit-learn · REST API · Linux · Git · Docker · Kubernetes · Terraform · AWS · Databricks · Snowflake · Trino · Iceberg · HDFS

Selected Projects

2026

WIFX — Applied ML Ranking & Match Prediction
huggingface.co/spaces/AMadabhushi/WIFX

- Built an end-to-end system ranking women's international football players, national teams and clubs from match event data, deployed as a public interactive application.
- **Benchmarked 33 classification models** with temporal cross-validation and a hard date cutoff; selected on log-loss and calibration rather than accuracy, rejecting gradient-boosted candidates that overfit — XGBoost reached 0.99 train accuracy but ranked near-worst on cross-validated log-loss.
- **Evaluated 17 regression models** (OLS, Ridge, Lasso, ElasticNet, SVR, Gaussian Process, PCR/PLS, Random Forest, GBM, XGBoost, stacked ensemble), tracking train-vs-CV R^2 gap as an explicit overfitting control.
- **Published a held-out validation report** on UEFA Women's Euro 2025 showing the model matched a majority-class baseline (0.0 lift), with confusion matrix and confidence-stratified accuracy — documenting where model outputs should not be trusted.

Experience

2023–Present

Senior Machine Learning Engineer
Steampunk

- **Built AI observability, evaluation and monitoring tooling** with DeepEval and Arize Phoenix for a production RAG system — bias, explainability, fairness, toxicity and summarization measurement wired into the deployment pipeline.
- Prototyped, architected and shipped the first enterprise multimodal retrieval-augmented generation application at the US Coast Guard using Python, FastAPI, LangChain, AWS and Chainlit; implemented keyword, semantic and hybrid retrieval with BM25 indexing.
- Hardened the prototype for IL4/IL5 secure production deployment with Helm charts, Docker and Git alongside the internal R&D team; led testing strategy and iterated on API design.
- **Scaled the Databricks backend from 12 to 40,000 users** across 11 enterprise organizations; designed Terraform modules, implemented monitoring controls for data pipeline infrastructure, and refactored pipelines to improve data search, lineage and processing.
- Implemented NIST SP 800 security and privacy controls enterprise-wide, including technical attribute-based access control automating the data governance strategy.
- Served as AI/ML subject matter expert to the Coast Guard CTO: ran an end-to-end AI risk assessment for attack-tree generators supporting disaster planning, and maintained the AI/ML inventory for the Office of Data and AI.

- Set design patterns and reusable containerized tooling for Steampunk's AI Lab, mentoring several teams on software engineering practice.

2022–2023

Senior Staff Data Architect
Johns Hopkins Applied Physics Laboratory

- **Applied supervised classification to video feeds to detect anomalous events** at sensitive naval installations (work under NDA); took the prototype to a containerized production deployment enabling rapid reporting.
- Refactored data search and processing infrastructure.

2020–2022

Lead Data Engineer
Thomson Reuters Special Services

- **Led data engineering for a new product, Data Platform, ingesting up to 500 GB/day from 15 vendors** — built on Snowflake with Apache NiFi orchestration and AWS S3 / Iceberg storage.
- Established design patterns that increased code reuse between the Data Engineering and Data Science teams; grew the team from myself to 3 developers in 6 months.
- Product was awarded a $15M multi-year contract.
- Developed a scalable search application over large-scale data supporting cyber crime investigations.

2018–2020

Research Software Engineer
RAND Corporation

- **Built a support vector machine classifier** (scikit-learn, Python, Django, MongoDB) to assist budget development for the government of Puerto Rico, reaching 84.2% classification accuracy.
- Maintained an enterprise HDFS cluster serving 250 researchers running machine learning workloads; performed upgrades and migrations independently.
- Developed technical recommendations to Pentagon leadership on ethical and responsible AI; conducted risk assessments for AI use cases in humanitarian aid scenarios.
- Selected for the RAND summer associate program (under 10% acceptance rate).

2013–2018

Earlier
Engineers Without Borders · NASA · Locus Technologies · Heinz College

- Project manager (Engineers Without Borders); NASA research grant recipient; engineer on the Locus EIM SaaS application; graduate teaching assistant in geographic information systems.

Education

Carnegie Mellon University
M.S., Civil and Environmental Engineering — Advanced Infrastructure Systems

- **Graduate research applying machine learning to structural health monitoring**: classified accelerometer time-series to quantify bridge damage in a laboratory setting.

University of Delaware
B.Eng., Environmental Engineering — Computational Fluid Dynamics