Andrew Ganse, PhD

research.ganse.org | github.com/aganse | linkedin.com/in/andrew-ganse | andrew@ganse.org

Summary

Hands-on senior data scientist/ML engineer and tech lead: 8+ years shipping production ML/AI products and 12+ years of statistical modeling, uncertainty quantification, and model evaluation on large, noisy, real-world data. I write the code and run the analyses end to end, from data discovery and pipelines through validation, training, deployment, and monitoring. Startup veteran who delivers customer pilots and works closely with engineering and product.

Professional Experience

Senior Data Scientist — ThruWave Inc., Seattle WA · Jul 2020 – Mar 2026 (company closed)
Data science + machine learning for automated final QC of packaged products in warehouses.

- Built production return-fraud detection: per-product models flagging returned boxes whose contents deviated from the expected population, catching adversarial substitutions (e.g. a brick for a mini-PC matched to the original on size and weight), under severe class imbalance and false-positive cost constraints.
- Delivered 10+ production pilots at major corporations and government agencies, shipping 16+ models stable under data drift, converting multiple deployments into paying customers.
- Designed and built the company's ML and evaluation platforms - experiment tracking, model training, evaluation, deployment, and monitoring - for anomaly-detection and object-detection models in production.
- Owned data engineering end to end: built and maintained a 170TB image corpus backing training-set curation; reprocessed and recalibrated the full corpus on a temporary parallel EC2 fleet.
- Led and mentored a 4-member data science team and set roadmaps with product/engineering while remaining an individual contributor; improved delivery velocity ~50% via AI coding agents, CI/CD, and standardized ML workflows.
- Applied transformer-based vision-language models (OpenCLIP) for zero-shot and centroid-based classification and evaluation, building the embedding and similarity-search tooling behind them.
- Lead inventor on granted patent for high-throughput anomaly detection in 3D millimeter-wave imaging systems.
- Tech: Python, SQL/PostgreSQL, dbt, Scikit-learn, MLflow, AWS, Docker, PyTorch/TensorFlow, OpenCLIP, Git.

Data Scientist — Echodyne Corporation, Kirkland WA · Dec 2017 – Jul 2020
Data science + machine learning on meta-materials tracking radar.

- Built and scaled the company's field-test data, labeling, and model performance evaluation platform, automating time-series processing, dashboards, and reporting across 100+ live tests, raising evaluation cadence from monthly to weekly.
- Designed controlled, matched-condition field experiments to measure model performance changes across releases.
- Built the company's first ML-based object classification capability for a production system, driving increased customer adoption, expanded implementation in follow-on versions, and influencing 10+ feature releases.
- Autonomous automotive perception (CNNs) and drone tracking/classification using radar, GPS, RGB camera, lidar.
- Tech: Python/Pandas, SQL/PostgreSQL, Scikit-learn, TensorFlow/Keras, Airflow, MLflow, Docker, Matlab.

Principal Scientist — Anseres Research & Technology LLC · Sep 2016 – Dec 2017
Federal scientific consulting in space/defense; led SBIR proposal on deep learning for clutter reduction in sonar.

Data Scientist — Spare5 · Jan 2016 – June 2016.
Crowdsourced data labeling with machine-learning-driven quality control.

- Produced scoring model that resulted in 8% lift in search-traffic to a customer's data in first few months.

Earlier career: Senior Research Physicist, Applied Physics Lab, Univ. of WA · 1999 – 2015
Applied research and statistical modeling in ocean-acoustic, electromagnetic, and gravimetric sensing domains.

Education

Ph.D. Geophysics, University of Washington: statistical inference, parameter estimation, inverse problems.
B.S. Electrical Engineering, University of Washington: nonlinear optimization and signal processing.

Selected Side Projects Built (see GitHub and research.ganse.org)

Flow_models: Flow-based invertible neural networks for distribution mapping and generative image modeling.
RAG_MLflow_eval: Langchain-based RAG evaluation with LLM-as-judge, logging datasets and traces to MLflow.