TOM MARSH

Boston, MA | [LinkedIn](#) | [GitHub](#)| [Personal Website](#)

SUMMARY   Senior Machine Learning Engineer with 8 years building production ML systems at TB/PB scale. Track record delivering measurable business impact: foundational ML infrastructure generating €15M+ annually, platform migrations saving $200K+ yearly, and 0-1 startup ML system contributing to $10M+ valuation. Physics background (MSc) with end-to-end ownership from architecture to deployment.

EXPERIENCE

Ströer Labs NZ - Senior Machine Learning Engineer

Christchurch, New Zealand   **Feb 2018 - Nov 2025**

- Reduced ad impression render time by 50% through ML inference optimisation, generating €15M+ additional annual revenue across 3B+ daily ad requests (1T+ annually)
- Led complete migration of ML data pipelines from Spark to DBT, eliminating Spark cluster costs ($10K+/month), reducing pipeline runtime 85% (20 min to 3 min), and simplifying maintenance for successor team
- Architected migration from AWS SageMaker to custom Kubernetes infrastructure, enabling GPU training for LightGBM models (previously impossible), unifying ML platform with company infrastructure
- Cut cloud egress costs by 35% ($200K+ annually) for product vertical through ML optimisation; maintenance team expanded savings to $30K/month post-transition
- Led 0-1 development of multiple ML products (classification, time-series forecasting, optimization) from concept to production on PB-scale datasets, establishing technical architecture before transitioning to dedicated teams
- Drove ML platform modernization during company's bare-metal-to-AWS migration, collaborating with DevOps and Data Engineering to redesign datalakes, account structures, and deployment practices
- Designed technical interview process for ML Engineering roles: created programming assessments, architecture evaluation framework, and systems design questions used across MLE/SWE/DevOps hiring
- Conducted 80+ technical interviews over 6 years, contributing to 15 hires as NZ office grew from 20 to 30 engineers
- Hosted interns each summer and mentored juniors through intermediate

Pyper Vision - Founding Machine Learning Engineer (Contracts, part-time)

Contract 1 | Christchurch, New Zealand   **Jul 2024 - Feb 2025**

- Established complete ML infrastructure as sole technical hire, enabling aerospace startup pivot from hardware to ML-based fog forecasting (company expanded 5 to 20+ airports, valued $10M+ post-funding)
- Architected end-to-end system on GCP from scratch: automated data collection, preprocessing pipelines, model training infrastructure, and production inference API with <100ms p99 latency
- Evaluated and deployed model architectures (LightGBM GBDT, TensorFlow CNNs) achieving commercially competitive forecast accuracy, delivering production system in 8 months
- System scaled to UK market expansion (advised on integration strategy post-contract)

- Designed physics-validated fog classification system using pseudo-labeling pipeline derived from meteorological research papers
- Implemented clustering analysis by airport fog patterns and model error modes, creating systematic debugging framework for operations team to diagnose regional forecast failures

---

EDUCATION

MSc, Physics - University of Canterbury **Jul 2022 - Jul 2024**

- Thesis: Extreme precipitation events under climate change using 4000+ years of simulation data at multiple resolutions
- Collaboration with National Institute of Water and Atmospheric Research (NIWA)
- Presented research at Meteorological Society of New Zealand Annual Conference (2023)

BSc, Physics & Computer Science - University of Canterbury **Feb 2015 - Nov 2017**

- Capstone: Y2SiO5 thin film growth for quantum information devices using Pulsed Laser Deposition

---

TECHNICAL SKILLS

ML/Data: PyTorch, TensorFlow, LightGBM, XGBoost, Scikit-learn, Pandas, NumPy, Spark, MLFlow, DBT
Infrastructure: AWS (SageMaker, EMR, S3, Glue, Athena), GCP (Cloud Functions, GCS), Docker, Kubernetes
Data Engineering: Apache Airflow, Kafka, SQL (Postgres, Presto, Trino), time-series databases
Languages: Python, SQL, Java, Scala, Kotlin, Go
Miscellaneous: Unix, Source Control (Git)