Ross Everett Altman

+1-516-375-4505 | ross.e.altman@gmail.com | github.com/knowbodynos
linkedin.com/in/ross-altman | Publications | Patents

Summary

Principal Machine Learning Engineer and Data Scientist bridging research depth and startup execution. 10+ years experience building, training, and deploying scalable data and ML systems across physics and biotech. Technical lead for generative design, transfer learning, and foundation model initiatives. Passionate about building interdisciplinary knowledge to deliver efficient and impactful solutions.

Experience

Inari Agriculture, Cambridge, MA Jan 2019 – Present
Machine Learning Engineer

- Trained a biologically-informed foundation LLM for plant genomes and fine-tuned LoRA adapters, improving gene expression prediction by **25% Spearman correlation** and automating predictive design pipelines across **five research programs**.
- Architected and deployed a secure, in-house protein design platform using Airflow, AlphaFold, and generative models enabling high-throughput generation and efficiency screening of synthetic CRISPR enzymes while saving **$50K annually** in vendor licensing costs.
- Built a high-throughput AlphaFold protein-protein interaction screening workflow that cut experimental resource requirements by **97%**, accelerating novel target discovery.
- Developed a cross-species transfer learning method leveraging multimodal VAEs to augment sparse soybean scRNA-seq datasets by **300%**, improving gene network inference accuracy.
- Launched high-throughput ML screen for phenotypic impact of CRISPR edits, expanding model-assisted product advancement pipelines to **46%** more plants and reducing the burden on greenhouse capacity.
- Automated quantitative root phenotyping using computer vision and graph analysis, **transforming 1,000+ noisy images** into interpretable feature sets for trait scoring.

Insight Data Science, Boston, MA Sept 2018 – Dec 2018
Data Science Fellow

- Developed a Word2Vec-based pipeline to boost content relevance of Wikipedia page previews, validated by a **17% improvement** in clickstream engagement.
- Delivered a Chrome extension using Flask and AWS for real-time user-side data retrieval and ranking.

Northeastern University, Boston, MA Sept 2011 – May 2017
Graduate Research Assistant, PhD Physics

- Designed distributed, high-performance C++/Python algorithms to construct a **database of 100k vacuum state solutions** compatible with Type-IIB string theory, bridging geometry with the Standard Model extensions.
- Implemented deep-learning "equation learner" networks projecting this out to a landscape of over 10^{10k} total vacuum states.
- Built and hosted a MongoDB-backed database search engine for global collaborators and co-founded the **String Data** initiative at the intersection of physics and ML.

Education

PhD, Physics, Northeastern University — 2017

- **Focus:** String theory, topology, machine learning.

MEng, Applied Physics, Cornell University — 2011
BSc, Applied & Engineering Physics, Cornell University — 2009

Skills

Programming:	Python, C++, Shell, SQL, PyTorch, TensorFlow, HuggingFace, FastAPI
Core ML:	Foundation Models, LLMs, LoRA, Transformers, Diffusion Models, VAE, Geometric Deep Learning, Decision Trees, Statistical Modeling, Representation Learning, Causal Inference
MLOps:	MLFlow, Airflow, Docker, Kubernetes, AWS, Terraform, CI/CD, Model Deployment, Data Pipelines
Bioinformatics:	Protein Modeling, RNA-seq Analysis, Network Biology, Phenotyping Pipelines, Gene Regulation Modeling