Ross Everett Altman

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

bit.ly/ross-altman-publications | bit.ly/ross-altman-patents

Summary

Data scientist and full-stack ML engineer with 6+ years of industry experience. PhD computational and mathematical physicist by training. Self-starter, lifelong learner, and teller of bad puns. Passionate about building interdisciplinary knowledge with people of diverse experience to deliver impactful results. Solving hard problems through data, algorithms, teamwork, and strong research fundamentals.

Work Experience

Inari Agriculture
Cambridge, MA

Machine Learning Engineer — *Jan 2024 - Present*

- Worked with stakeholders to deploy and scale a self-service AlphaFold pipeline in an isolated environment to accelerate design of proprietary proteins.
- Developed a high-throughput protein-protein interaction AlphaFold screen to discover causal interventions for trait improvement in soybean.
- Inferred enriched gene regulatory networks for target discovery in soybean by integrating single-cell RNA-seq data from multiple related species using a transfer learning approach.
- Developed a novel and robust method extracting physiological features from hundreds of noisy soybean root images using a combination of computer vision and network analysis.

Data Scientist — *Jan 2019 - Jan 2024*

- Developed a high-throughput *in silico* screen for estimating phenotypic impact of CRISPR-Cas edited alleles in plants using a mixture of protein language models and bioinformatic methods.
- Trained a protein language model to discover novel stability-enhancing mutations for improving CRISPR-Cas nuclease efficiency for genome editing in plants.
- Supervised a team of co-op students to train a transformer-based genomic foundation model in plants to predict tissue-specific *cis*-regulatory activity.
- Pioneered the application of deep learning for sequence-to-function modeling and network perturbation-response modeling, leading to the establishment of a full AI department.
- Mentored 10+ junior colleagues, interns, and data science Masters co-op students.

Insight Data Science
Boston, MA

Fellow — *Sept 2018 - Dec 2018*

- Implemented a Word2Vec model to enhance Wikipedia page previews by displaying the most relevant information to the current page.
- Used Wikipedia's bi-directional clickstream data to validate a 17% increase in relevance over default page previews.
- Built a Chrome browser extension served using a jQuery/Flask/AWS stack to achieve high performance on real-time, client-side data.

Northeastern University
Boston, MA

Graduate Research Assistant — *Sept 2011 - May 2017*

- Designed a distributed algorithm in C++/Python to systematically compute type-IIB string theory vacuum states on an HPC cluster.
- Compiled the world's largest database of 10^5 string theory vacua, resulting in discovery of candidate universes extending the Standard Model.
- Estimated the number of vacua ($> 10^{(10^4)}$) in the string landscape using deep neural networks using specialized "equation learner" layers.
- Built and maintained a queryable web UI with MongoDB backend based on feedback from stakeholders in the string theory community.
- Co-founded the String Data group for ML applied to string theory, leading to international conferences and collaboration with industry.

Education

Northeastern University
Boston, MA

PhD in Physics — *Sept 2011 - May 2017*

- **Focus:** String theory/phenomenology, high-dimensional geometry, topology, high-performance computing, machine learning.

Cornell University
Ithaca, NY

MEng in Applied Physics — *Sept 2010 - May 2011*

Cornell University
Ithaca, NY

BSc in Applied & Engineering Physics — *Sept 2005 - May 2009*

Skills

Programming Languages	Python, C/C++, Shell, HTML/CSS, JavaScript.
Machine Learning Stack	PyTorch, TensorFlow/Keras, Scikit-learn, OpenCV, Huggingface, Lightning AI, MLFlow.
Engineering Stack	AWS, Docker, Kubernetes, Airflow, Flask/FastAPI, Terraform, MongoDB.
Technical Skills	NLP, Computer Vision, Network Analysis, Statistics, Geometry/Topology, Bioinformatics, CICD.
Soft Skills	Self-driven, Cross-functional collaboration, Scientific communication, Interpersonal skills, Rapid prototyping.