Ross Everett Altman

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

Summary

Experienced data scientist, physicist, and engineer. Self-starter, lifelong learner, bad pun teller. Driven to understand how things work. Solving puzzles through data, algorithms, hardware, and research fundamentals.

Work Experience

Inari Agriculture *Cambridge, MA*

Machine Learning Scientist *Jan 2019 - Present*

- Co-led a four-person team developing a BERT-based protein language model to discover novel stability-enhancing mutations for improving efficiency of CRISPR-Cas proteins for genome editing in plants.
- Developed epistasis-aware DNA and protein language models to identify and prioritize high-impact deleterious mutations in sequenced corn varieties for correction via genome editing.
- Developed a method for discovering causal interventions for trait improvement in soybean by integrating gene regulatory network analysis, AlphaFold2, and protein language models.
- Mentored 8+ junior colleagues and interns, attended multiple conferences, and maintained regular communication with external scientific advisors.

Insight Data Science *Boston, MA*

Fellow *Sept 2018 - Dec 2018*

- Built a https://www.wikontext.us/ to enhance Wikipedia page previews by intelligently displaying more relevant information in contrast to the baseline naive page summary.
- Implemented a Word2Vec-based sentence embedding model, and externally validated it using knowledge of bi-directional clickstream information with up to 17% more accuracy than baseline.
- Served the model 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 modular, distributed algorithm in C++/Python to systematically compute type-IIB string theory vacuum states on an HPC cluster.
- Computed 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 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 industry collaborators.

Education

Northeastern University *Boston, MA*

PhD in Physics (Dissertation) *Sept 2011 - May 2017*

- **Focus:** String theory/phenomenology, high-dimensional geometry, topology, computational methods, 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	PyTorch, TensorFlow, Scikit-learn, Huggingface, Lightning, MLFlow.
Engineering Stack	AWS, Docker, Kubernetes, Airflow, Flask/FastAPI, Terraform, MongoDB.
Technical Skills	NLP, Computer Vision, Network Analysis, Statistics, Algebraic Geometry/Topology, Bioinformatics, CICD.
Soft Skills	Team/Project Leadership, Cross-functional Collaboration, Literature Review, Rapid Prototyping, Self-Motivation, Scientific Communication, Technical Documentation.

References available upon request.