deep
genomics

(Senior) Machine Learning Scientist, AI Foundation Model Specialist

North America

Machine Learning /   Full-Time /   Hybrid

About Us

Deep Genomics is at the forefront of using artificial intelligence to transform drug discovery. Our proprietary AI platform decodes the complexity of genome biology to identify novel drug targets, mechanisms, and genetic medicines inaccessible through traditional methods. We co-develop drug programs and AI models with partners and internally, and pursue major technology builds with pharmaceutical partners. With expertise spanning machine learning, bioinformatics, data science, engineering, and drug development, our multidisciplinary team located in Toronto, Cambridge, MA, and select other sites is revolutionizing how new medicines are created.

Where You Fit In

We are seeking a passionate and highly skilled Machine Learning Scientist with expertise in Foundation Models to join our core AI research team, focusing on developing and scaling our next-generation foundation models for biology. You will work with domain area experts to explore fundamental research questions at the intersection of deep learning and biology, and apply the latest in AI research to unique, large-scale biological datasets (many terabytes) to develop models that push the state-of-the-art. If you are excited by using your strong ML knowledge to work with others and to solve frontier problems in drug discovery, this is a unique opportunity.

Key Responsibilities

- Collaborate closely with computational biologists to integrate domain knowledge, define scientifically meaningful tasks, and translate biological challenges into ML frameworks. You will work together to formulate key scientific questions in biology that can be addressed through innovative ML/AI approaches and design computational experiments to test hypotheses.

- Contribute to, and potentially lead (depending on level) research into novel deep learning architectures, training paradigms (e.g., self-supervised, multi-modal), and algorithms tailored for large-scale biological sequence data and related modalities.

- Rigorously implement, train, debug, and evaluate models to demonstrate scientific validity and potential for downstream application.

- Stay current with advancements in both machine learning and computational biology literature, identifying cross-disciplinary opportunities to solve real-world challenges.

- Contribute to team knowledge sharing and code quality through documentation and code reviews.

- Share research findings through internal presentations, and potentially external publications or conference presentations.

Basic Qualifications

- PhD or MSc with a strong research focus in Machine Learning, Computer Science, Statistics, Physics, or a related quantitative field.

- Deep understanding of the theoretical underpinnings and practical application of modern deep learning, including architectures like Transformers and related sequence models (e.g. state-space models).

- Proven ability to implement, train, and debug highly-performant deep learning models using frameworks like PyTorch or JAX.

- Experience working with large datasets and understanding the challenges associated with scale (even if not directly managing infrastructure).

- Strong analytical and problem-solving skills, with the ability to translate ambiguous scientific problems into tractable ML formulations.

- Excellent communication skills, capable of discussing complex ideas with both technical and scientific audiences.

Preferred Qualifications

- 2+ years of relevant post-graduate experience in an industrial R&D or applied science setting, applying advanced ML to solve complex scientific or technical problems.

- Experience technically leading projects or mentoring junior researchers/engineers.

- Demonstrated potential through strong PhD/MSc research, impactful projects, relevant internships, or open-source contributions.

- Track record of impactful research demonstrated through first or senior author publications in top-tier ML or relevant scientific journals/conferences(e.g., NeurIPS, ICML, ICLR).

- Proficiency with cloud computing platforms (e.g., AWS, GCP) for large-scale model training and experimentation.

What you'll gain

- Your ML models will directly and immediately impact the creation of new genetic medicines for patients with unmet needs. This is not a disjoint R&D division - your models will be front-and-center in collaborations with drug developers and with our established pharmaceutical partners.

- Discovery of truly causal relationships in complex biological systems. These can even predict the impact of ultra-rare events (like genetic variants only ever seen in a single patient) that break typical correlative ML paradigms.

- Immerse yourself in genomics and drug discovery through daily collaboration with world-class bioinformaticians and biologists, quickly gaining the domain knowledge to maximize your ML innovation and impact (no prior biology expertise required).

- An opportunity to publish and present work at the forefront of foundation models for genome biology, with applications in medicine and drug discovery.

What we offer

- A collaborative and innovative environment at the frontier of computational biology, machine learning, and drug discovery.

- Highly competitive compensation, including meaningful stock ownership.

- Comprehensive benefits - including health, vision, and dental coverage for employees and families, employee and family assistance program.

- Flexible work environment - including flexible hours, extended long weekends, holiday shutdown, unlimited personal days.

- Maternity and parental leave top-up coverage, as well as new parent paid time off.

- Focus on learning and growth for all employees - learning and development budget & lunch and learns.

- Facilities located in the heart of Toronto - the epicenter of machine learning and AI research and development, and in Kendall Square, Cambridge, Mass. - a global center of biotechnology and life sciences.

Deep Genomics welcomes and encourages applications from people with disabilities. Accommodations are available on request for candidates taking part in all aspects of the selection process.

Deep Genomics thanks all applicants, however only those selected for an interview will be contacted.

[Deep Genomics Home Page](#)

Jobs powered by LEVER