6 month ML Engineering Internship @ Broad Institute

The life sciences are in the midst of a data revolution. Cheap and accurate genome sequencing is here, advanced imaging is routine, and clinical data is increasingly stored in electronic formats. These innovations — and the massive data sets they produce — have brought us to the threshold of a new era in medicine, one where the data sciences hold the potential to propel our understanding and treatment of human disease.

As a result, The Broad Institute has created a machine learning initiative to drive innovation beginning with the world's leading killer, cardiovascular disease. We are building open-source software to derive insights from massive amounts of multimodal clinical and genomic (ECGs, MRIs, EHRs) data using deep learning architectures. Our aims include more than just predictive accuracy for the onset of disease -- we will use our models to uncover new biology itself that can drive new therapeutics and treatment.

The role:

- You will work on a small team of machine learning scientists, engineers and cardiovascular disease experts.
- You will identify the highest impact opportunities by whiteboarding solutions with the top genetic and clinical researchers in the world.
- You will conduct theoretical ML research with collaborators from MIT and Harvard and implement cutting edge methods for learning representations of cardiovascular health.
- You will collaborate regularly with MD/PhDs from across the life sciences to explore how to embed cardiac physiology into learning systems.
- You will hunt for new ways to accelerate therapeutic development with industry collaborators.
- You will deploy cloud-based software to a diverse audience of researchers across the world.
- You will work with real patient genomic and clinical data to train and validate impact.

Qualifications

- BS/MS/PhD in Computer Science or technical field (Physics, Math, etc.) with significant machine learning coursework from computer science departments. Strong undergraduate or graduate students who are still pursuing their degree are encouraged to apply. We also welcome non-traditional candidates with demonstrable technical proficiency and a track record of working on teams.
- Very well versed in Deep Learning approaches to sequence, image, or graphical network analysis.
- Fluent in Python, Tensorflow or PyTorch
- Familiar with cloud-based data pipelines (experience with Redshift/BigQuery, Apache Beam, Spark)