CS273B: Deep learning for genomics and bio-medicine

Instructors: Anshul Kundaje, James Zou, Serafim Batzoglou

TAs: Anna Shcherbina, Nadine Hussami, Naveen Arivazhagan, Alon Devorah

Course overview

- Deep learning + computational biology = exciting!
- This class is projects focused research class:
  best way to learn and state-of-the-art very limited
- Prerequisites: first course in machine learning and statistics. We will cover the relevant biology and neural networks.

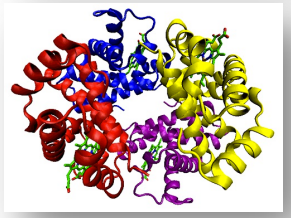
Course overview

Module 1: crash course of deep learning and TensorFlow/Keras (7 classes).

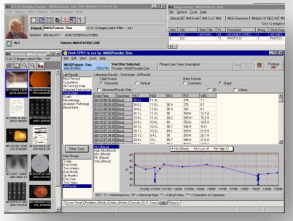
Paper presentations 2/class

Module 2: genetics and regulatory genomics (4 classes).

Module 3: protein structure and pharmacogenomics (3 classes).



Module 4: electronic health records and medical imaging (4 classes).



Inside the black box!

Course overview

Check Canvas for latest and resources.

9/26	Intro to neural networks, backprop + intro to medical imaging problems	James	Stegle Review	http://msb.embopress.org/content/12/7/878 , http://neuralnetworksanddeeplearning.com/
9/28	Convolutional neural network + intro to func genomics	Anshul	DeepBind DeepCpG	http://www.nature.com/nbt/journal/v33/n8/full/nbt.3300.html, http://biorxiv.org/content/early/2016/05/27/055715,
10/3	Recurrent neural network + intro to protein structure problems	Anshul, Serafim	DeepNano: Deep Recurrent Neural Networks for Base Calling in MinION Nanopore Reads	http://arxiv.org/abs/1603.09195
10/5	Unsupervised deep learning + Deep generative models	James		
10/10	Training deep neural networks + intro to EHR/clinical data	James		
10/12	Demo on Azure/Keras	Anshul + TAs, Serafim		
10/17	Demo on Azure/TensorFlow	James + TAs		
10/19	Func Genomics/Variant/PopGen	James	DeepSEA; Basset	http://www.nature.com/nmeth/journal/v12/n10/full/nmeth.3547.html http://genome.cshlp.org/content/26/7/990
10/24	Func Genomics/Variant/PopGen	James	DanQ; The human splicing code reveals new insights into the genetic determinants of disease	http://nar.oxfordjournals.org/content/44/11/e107; http://sites.utoronto.ca/intron/xiong2015.pdf

Assessments

Project (50 %): team of 4-6 students. Each team will be provided with Azure GPU compute resources. We'll have suggested projects + datasets. Final paper and talk.

Paper presentation (20%): each team selects and presents a paper in detail (35mins).

Paper review (20%): each team provides written reviews of 2 other papers.

Quizzes (10%): in class.

Logistics

Section: Fridays 10:30 – 11:20am (Hewlett 102).

Office hours:

James: Wednesdays 5 – 7pm (Packard 2nd floor)
Anna and Nadine: Mondays 5 – 7pm (Lane L339)
One more TBA.

Start thinking about teams in the next two weeks.

Ask questions on Piazza!

Lecture 1: feedforward neural network

- Activation and architecture
- Training: backpropagation
- Expressivity and efficiency (demo)