Machine Learning for Healthcare and Biology Zarif Azher - 4/20/22

TJ Machine Learning Club Guest Lecture



SIGNUP AT: TINYURL.COM/BIOCODE-2022





HOSTED BY TJ BIOINFORMATICS SOCIETY

BICCODE

DON'T MISS OUT!

SATURDAY APRIL 23, 9 AM - 3PM **AT TJ**

COMPETITION FOR ALL LEVELS - OVER \$1000 IN PRIZES!!!

Overview
Where are we going??



The ML Big Idea Ok ok we already know..but let's just review

- Computer algorithms taking data as input
- Generating some type of output based on the input
 - For a given task
- Doing the task...
 - Better, faster, cheaper than existing methods



ML For Healthcare What is he even talking about??

Healthcare

Relies on + generates a LOT of data



Aaah there is so much!!! Let me help







ML For Healthcare (cont) Sooo here it is...

• Healthcare processes are data-hungry

- But also data-fat (original term⁽²⁾)
- Given all of this data about a person, what does it mean?
 - What **data** is **important**, and what is **junk**??
- (My take on the) **Big Goal** of **Healthcare-ML**: Making sense of complicated data to help patients and clinicians understand patient health
 - Other takes may include..."replace doctors", etc.









ð







Rahmani et al, 2021

Electronic Health Records Let's get specific

- Patient interactions generates:
 - Diagnostic notes, progress, demographics, etc —> health record
- Since the late 1990s, health records have become electronic
- documents...a whole lot of unstructured data
- Clinicians: estimated 16+ minutes per patient looking over health record
 - Added up, huge waste of time and productivity!!!

Electronic Health Records (cont) Discussion time :)

Original Investigation | Health Informatics

July 23, 2021

Review of Patient Records

Ethan Andrew Chi¹; Gordon Chi¹; Cheuk To Tsui, BS¹; <u>et al</u>

Article Information » Author Affiliations

JAMA Netw Open. 2021;4(7):e2117391. doi:10.1001/jamanetworkopen.2021.17391

shorturl.at/mxMP8

5 minute read + 5 minute discuss

This is just one example - there are many similar studies!!!

Development and Validation of an Artificial Intelligence System to Optimize Clinician

ML For Biology Yes, there is a subtle difference



We barely understand what biological phenomenon does and how it works

But what if computers get it better??



ML For Biology (cont) There's so much :0

- **Biology** can be **represented** as **numbers**
 - Metabolic rate, population size, white blood cell count
- And sequences
 - DNA, RNA, amino acids/proteins
- And categories
 - Genes, proteins, pathways, organisms
- organisms and structures

ML models can learn from these to make intelligent predictions about

ML For Biology (cont) Let's think about a few examples

- Predicting properties of proteins
- Protein folding (AlphaFold!)
- Risk of developing certain diseases often applied to cancer

What other ideas/ applications can you think of?

What types of ML models can be used? (Ex; DNN, CNN, RNN, etc)

A Cancer Subtype Prediction Study Discussion time again :)

Research Open Access Published: 03 April 2020

Convolutional neural network models for cancer type prediction based on gene expression

Milad Mostavi, Yu-Chiao Chiu, Yufei Huang 🗠 & Yidong Chen 🗠

BMC Medical Genomics 13, Article number: 44 (2020) Cite this article

10k Accesses 32 Citations 8 Altmetric Metrics

shorturl.at/qM078

5 minute read + 5 minute discuss

Remember, this is just one example - there are many other studies!!!

Graph Convolutional Networks (if extra time)

If we're here, that means we have a lot of extra time left:)

Let's move over to the whiteboard to graphs, how they apply to biology, GCNs, and what the heck I'm talking about



Thanks! (PS: Don't forget about BioCode!) tinyurl.com/biocode-2022

