Machine Learning for Healthcare: Introduction

David Sontag
Clinical Machine Learning Group
MIT
Instructors

• David Sontag
  – Associate Professor in EECS (course 6) and part of CSAIL and IMES
  – PhD ‘10, then 5 years as professor at NYU
  – Leads clinical machine learning research group

• Steven Horng, MD (tomorrow)
  – Board certified in emergency medicine and clinical informatics
  – Research on AI/ML in emergency medicine
Students

- ~20 students from as far as New Zealand and South Korea
- Introduce yourself!
  - Name?
  - Where do you live?
  - What is your profession?
  - Why did you register for this class?
Outline for morning

1. Brief history of AI and ML in healthcare
2. Why now?
3. Examples of how ML will transform healthcare
4. What is unique about ML in healthcare?
1970's: MYCIN expert system

- 1970's (Stanford): MYCIN expert system for identifying bacteria causing severe infections
- Proposed a good therapy in ~69% of cases. Better than infectious disease experts

![Diagram of expert system]

**FIGURE 1-1** Major parts of an expert system. Arrows indicate information flow.

**FIGURE 33-1** Short sample dialogue. The physician's inputs appear in capital letters after the double asterisks.

**Dialogue interface**

- I am ready
- **THIS IS A 26 YEAR OLD MALE PATIENT**
- My understanding is:
  - The age of the patient is 26
  - The sex of the patient is male
- **FIVE DAYS AGO, HE HAD RESPIRATORY-TRACT SYMPTOMS**
- What is his name?
- **JO**
- My understanding is:
  - The name of the patient is Jo
  - Respiratory-tract is one of the symptoms that the patient had
- **A COUPLE OF DAYS BEFORE THE ADMISSION, HE HAD A MALAISE**
- Please give me the date of admission
- **MARCH 12, 1979**
- My understanding is:
  - The patient was admitted at the hospital 3 days ago
  - Malaise is one of the symptoms that the patient had 5 days ago
1980’s: INTERNIST-1/QMR model

- 1980’s (Univ. of Pittsburgh): INTERNIST-1/Quick Medical Reference
- Diagnosis for internal medicine

Probabilistic model relating:
- 570 binary disease variables
- 4,075 binary symptom variables
- 45,470 directed edges

Elicited from doctors:
- 15 person-years of work

Led to advances in ML & AI
(Bayesian networks, approximate inference)

Problems: 1. Clinicians entered symptoms manually
2. Difficult to maintain, difficult to generalize

[Miller et al., ‘86, Shwe et al., ‘91]
1980’s: automating medical discovery

1990’s: neural networks in medicine

- Neural networks with clinical data took off in 1990, with 88 new studies that year
- Small number of features (inputs)
- Data often collected by chart review

**Problems:**
1. Did not fit well into clinical workflow
2. Hard to get enough training data
3. Poor generalization to new places

[Figure 2. A multilayer perceptron. This is a two-layer perceptron with four inputs, four hidden units, and one output unit.]

<table>
<thead>
<tr>
<th>Subject</th>
<th>No. of Examples</th>
<th>Network</th>
<th>D‡</th>
<th>Accuracy$</th>
<th>Neural</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast cancer</td>
<td>57</td>
<td>9-15-2</td>
<td>0.6</td>
<td>80</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>Vasculitis</td>
<td>404</td>
<td>8-5-1</td>
<td>8.0</td>
<td>94</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>351</td>
<td>20-10-10-1</td>
<td>1.1</td>
<td>97</td>
<td>84</td>
<td></td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>356</td>
<td>20-10-10-1</td>
<td>1.1</td>
<td>97</td>
<td>94</td>
<td></td>
</tr>
<tr>
<td>Low back pain</td>
<td>100</td>
<td>50-48-2</td>
<td>0.2</td>
<td>90</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>Cancer outcome</td>
<td>5,169</td>
<td>54-40-1</td>
<td>1.4</td>
<td>0.779</td>
<td>0.776</td>
<td></td>
</tr>
<tr>
<td>Psychiatric length of stay</td>
<td>957</td>
<td>48-400-4</td>
<td>0.2</td>
<td>74</td>
<td>76</td>
<td></td>
</tr>
<tr>
<td>Intensive care outcome</td>
<td>284</td>
<td>27-18-1</td>
<td>0.5</td>
<td>0.82</td>
<td></td>
<td>0.82</td>
</tr>
<tr>
<td>Skin tumor</td>
<td>150</td>
<td>18</td>
<td></td>
<td>80</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>Evoked potentials</td>
<td>100</td>
<td>14-4-3</td>
<td>3.8</td>
<td>77</td>
<td></td>
<td>77</td>
</tr>
<tr>
<td>Head injury</td>
<td>500</td>
<td>6-3-3</td>
<td>20</td>
<td>66</td>
<td></td>
<td>77</td>
</tr>
<tr>
<td>Psychiatric outcome</td>
<td>289</td>
<td>41-10-1</td>
<td>0.7</td>
<td>79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tumor classification</td>
<td>53</td>
<td>8-9-3</td>
<td>1.4</td>
<td>99</td>
<td></td>
<td>88</td>
</tr>
<tr>
<td>Dementia</td>
<td>75</td>
<td>80-10-7-7</td>
<td>0.6</td>
<td>61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pulmonary embolism</td>
<td>607</td>
<td>50-4-1</td>
<td>2.9</td>
<td>0.82</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td>Heart disease</td>
<td>460</td>
<td>35-16-8-2</td>
<td>3</td>
<td>83</td>
<td></td>
<td>93</td>
</tr>
<tr>
<td>Thyroid function</td>
<td>3,600</td>
<td>21-16-8-3</td>
<td>22</td>
<td>98</td>
<td></td>
<td>93</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>350</td>
<td>9-4-4-2</td>
<td>10</td>
<td>97</td>
<td></td>
<td>96</td>
</tr>
<tr>
<td>Diabetes</td>
<td>384</td>
<td>8-4-4-2</td>
<td>12</td>
<td>77</td>
<td></td>
<td>75</td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>2,856</td>
<td>291-1</td>
<td>9.8</td>
<td>85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hepatitis</td>
<td>39</td>
<td>4-4-3</td>
<td>3.3</td>
<td>74</td>
<td></td>
<td>79</td>
</tr>
<tr>
<td>Psychiatric admission</td>
<td>319</td>
<td>53-1-1</td>
<td>6.0</td>
<td>91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiac length of stay</td>
<td>713</td>
<td>15-12-1</td>
<td>3.5</td>
<td>0.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anti-cancer agents</td>
<td>127</td>
<td>60-7-6</td>
<td>1.5</td>
<td>91</td>
<td>86</td>
<td></td>
</tr>
<tr>
<td>Ovarian cancer</td>
<td>75</td>
<td>6-6-2</td>
<td>2.6</td>
<td>84</td>
<td></td>
<td>81</td>
</tr>
<tr>
<td><strong>MEDIAN VALUE</strong></td>
<td><strong>350</strong></td>
<td><strong>20</strong></td>
<td></td>
<td><strong>2.8</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*For reference citations, see the reference list
†P = prior probability of most prevalent category.
‡D = ratio of training examples to weights per output.
§A single integer in the accuracy column denotes percentage overall classification rate and a single real number between 0 and 1 indicates the AUROCC value. Neural = accuracy of neural net, Other = accuracy of best other method.
Outline for morning

1. Brief history of AI and ML in healthcare
2. Why *now*?
3. Examples of how ML will transform healthcare
4. What is *unique* about ML in healthcare?
Why now?

- Cost of health care expenditures in the US are over $3 trillion, and rising
- Over half of adults have one or more chronic health conditions:
  - Often diagnosed late
  - Often inappropriately managed
  - Leads to otherwise preventable complications
Adoption of Electronic Health Records (EHR) has increased 9x since 2008.

[Henry et al., ONC Data Brief, May 2016]
Large datasets

De-identified health data from ~40K critical care patients

Demographics, vital signs, laboratory tests, medications, notes, ...

If you use MIMIC data or code in your work, please cite the following publication:

Large datasets

“Data on nearly 230 million unique patients since 1995”
Large datasets

President Obama’s initiative to create a 1 million person research cohort

Core data set:
- Baseline health exam
- Clinical data derived from electronic health records (EHRs)
- Healthcare claims
- Laboratory data

[Precision Medicine Initiative (PMI) working Group Report, Sept. 17 2015]
Diversity of digital health data

- Lab tests
- Proteomics
- Imaging
- Phone
- Social media
- Vital signs
- Devices
- Genomics
Standardization

• Diagnosis codes: ICD-9 and ICD-10 (International Classification of Diseases)

  ICD-9 codes 290–319: mental disorders
  ICD-9 codes 320–359: diseases of the nervous system
  ICD-9 codes 360–389: diseases of the sense organs
  ICD-9 codes 390–459: diseases of the circulatory system
  ICD-9 codes 460–519: diseases of the respiratory system
  ICD-9 codes 520–579: diseases of the digestive system
  ICD-9 codes 580–629: diseases of the genitourinary system
  ICD-9 codes 630–679: complications of pregnancy, childbirth, ...

  [https://en.wikipedia.org/wiki/List_of_ICD-9_codes]

[https://blog.curemd.com/the-most-bizarre-icd-10-codes-infographic/]
Standardization

- Diagnosis codes: ICD-9 and ICD-10 (International Classification of Diseases)
- Laboratory tests: LOINC codes
- Pharmacy: National Drug Codes (NDCs)
- Unified Medical Language System (UMLS): millions of medical concepts

[http://loinc.com/newsletter/index_May08.htm]
# Standardization

## Level 1
Basic framework on which the specification is built

| Foundation | Base Documentation, XML, JSON, REST API + Search, Data Types, Extensions |

## Level 2
Supporting Implementation, and binding to external specifications

- **Implementer Support**
  - Downloads, Common Use Cases, Testing
- **Security & Privacy**
  - Security, Consent, Provenance, AuditEvent
- **Conformance**
  - StructureDefinition, CapabilityStatement, ImplementationGuide, Profiling
- **Terminology**
  - CodeSystem, ValueSet, ConceptMap, Terminology Svc
- **Linked Data**
  - RDF

## Level 3
Linking to real world concepts in the healthcare system

| Administration | Patient, Practitioner, Device, Organization, Location, Healthcare Service |

## Level 4
Record-keeping and Data Exchange for the healthcare process

- **Clinical**
  - Allergy, Problem, CarePlan, DetectedIssue, RiskAssessment, etc.
- **Diagnostics**
  - Observation, Report, Specimen, ImagingStudy, Genomics, etc.
- **Medications**
  - Order, Dispense, Administration, Statement, Immunization, etc.
- **Workflow**
  - Task, Appointment, Schedule, Referral, PlanDefinition, etc.

## Level 5
Providing the ability to reason about the healthcare process

| Clinical Reasoning | Library, ServiceDefinition & GuidanceResponse, Measure/MeasureReport, etc. |
Standardization

OMOP Common Data Model v5.0
Breakthroughs in machine learning

Ever cleverer
Error rates on ImageNet Visual Recognition Challenge, %

Why now?
• Big data
• Algorithmic advances
• Open-source software

Sources: ImageNet; Stanford Vision Lab

Economist.com
Breakthroughs in machine learning

• Major advances in ML & AI
  – Learning with high-dimensional features (e.g., l1-regularization)
  – Semi-supervised and unsupervised learning
  – Modern deep learning techniques (e.g. convnets, variants of SGD)

• Democratization of machine learning
  – High quality open-source software, such as Python’s scikit-learn, TensorFlow, Torch, Theano
Industry interest in ML & healthcare
DIGITAL HEALTH FUNDING
2011-H1 2018

TOTAL VENTURE FUNDING

# OF DEALS

Source: Rock Health Funding Database
1: Shadowed portion shows projections for entire year of 2018, assuming current funding pace continues.
Note: Only includes U.S. deals >$2M; data through June 30, 2018
1.1 Why Now?

AI has been around for decades and its promise to revolutionize our lives has been frequently raised, with many of the promises remaining unfulfilled. Fueled by the growth of capabilities in computational hardware and associated algorithm development, as well as some degree of hype, AI research programs have ebbed and flowed. The JASON 2017 report gives this history and also comments on the current AI revolution stating:

"Starting around 2010, the field of AI has been jolted by the broad and unforeseen successes of a specific, decades-old technology: multi-layer neural networks (NNs). This phase-change reenergizing of a particular area of AI is the result of two evolutionary developments that together crossed a qualitative threshold: (i) fast hardware Graphics Processor Units (GPUs) allowing the training of much larger—and especially deeper (i.e., more layers)—networks, and (ii) large labeled data sets (images, web queries, social..."
Industry interest in ML & healthcare

• Major acquisitions to get big data for ML:
  – Merge ($1 billion purchase by IBM, 2015) *medical imaging*
  – Truven Health Analytics ($2.6 billion purchase by IBM, 2016) *health insurance claims*
  – Flatiron Health ($1.9 billion purchase by Roche, 2018) *electronic health records (oncology)*
Outline for morning

1. Brief history of AI and ML in healthcare
2. Why now?
3. Examples of how ML will transform healthcare
4. What is unique about ML in healthcare?
ML will transform every aspect of healthcare

The stakeholders:

Source for figure:
Emergency Department:

- Limited resources
- Time sensitive
- Critical decisions
What will the ER of the future be like?

Behind-the-scenes reasoning about the patient’s conditions (current and future)

- Better triage
- Faster diagnosis
- Early detection of adverse events
- Prevent medical errors

Automatically extracted from electronic health record
What will the ER of the future be like?

Propagating best practices

The ED Dashboard decision support algorithms have determined that this patient may be eligible for the Atrius Cellulitis pathway. Please choose from the following options:

- Enroll in pathway
- Decline

You can include a comment for the reviewers: **Mandatory if Declining**

Below are links to the pathway and/or other supporting documents:

Atrius Cellulitis Pathway
What will the ER of the future be like?

Anticipating the clinicians’ needs
What will the ER of the future be like?

Reducing the need for specialist consults

Input
Chest X-Ray Image

CheXNet
121-layer CNN

Output
Pneumonia Positive (85%)

Arrhythmia?

Figure sources: Rajpurkar et al., arXiv:1711.05225 '17
Rajpurkar et al., arXiv:1707.01836, '17
What will the ER of the future be like?

Automated documentation and billing

<table>
<thead>
<tr>
<th>Triage note</th>
<th>Predicted chief complaints</th>
<th>Contextual auto-complete</th>
</tr>
</thead>
<tbody>
<tr>
<td>69 y/o M Patient with severe intermittent RUQ pain. Began soon after eating. Also is a heavy drinker.</td>
<td>RUQ abdominal pain, Allergic reaction, L Knee pain, Rectal pain, Right sided abdominal pain.</td>
<td>Transfer, MCI</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chief Complaints:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RUQ abdominal pain</td>
</tr>
<tr>
<td></td>
<td>Allergic reaction</td>
</tr>
<tr>
<td></td>
<td>L Knee pain</td>
</tr>
<tr>
<td></td>
<td>Rectal pain</td>
</tr>
<tr>
<td></td>
<td>Right sided abdominal pain</td>
</tr>
<tr>
<td></td>
<td>Transfer</td>
</tr>
<tr>
<td></td>
<td>MCI</td>
</tr>
</tbody>
</table>

References

Triage note

What will the ER of the future be like?

Automated documentation and billing

<table>
<thead>
<tr>
<th>Triage note</th>
<th>Predicted chief complaints</th>
<th>Contextual auto-complete</th>
</tr>
</thead>
<tbody>
<tr>
<td>69 y/o M Patient with severe intermittent RUQ pain. Began soon after eating. Also is a heavy drinker.</td>
<td>RUQ abdominal pain, Allergic reaction, L Knee pain, Rectal pain, Right sided abdominal pain.</td>
<td>Transfer, MCI</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chief Complaints:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RUQ abdominal pain</td>
</tr>
<tr>
<td></td>
<td>Allergic reaction</td>
</tr>
<tr>
<td></td>
<td>L Knee pain</td>
</tr>
<tr>
<td></td>
<td>Rectal pain</td>
</tr>
<tr>
<td></td>
<td>Right sided abdominal pain</td>
</tr>
<tr>
<td></td>
<td>Transfer</td>
</tr>
<tr>
<td></td>
<td>MCI</td>
</tr>
</tbody>
</table>

References
TYPE 2

BODY CANNOT USE INSULIN PROPERLY
- Can develop at any age
- Most cases can be prevented

Currently, at least 1 out of 3 people will develop the disease in their lifetime
More than 5,000 youth diagnosed each year in 2008 and 2009

RISK FACTORS FOR TYPE 2 DIABETES:

- BEING OVERWEIGHT
- HAVING A FAMILY HISTORY
- HAVING DIABETES WHILE PREGNANT (GESTATIONAL DIABETES)

WHAT CAN YOU DO?

You can **prevent** or **delay** type 2 diabetes
- LOSE WEIGHT
- EAT HEALTHY
- BE MORE ACTIVE

You can **manage** diabetes
- WORK WITH A HEALTH PROFESSIONAL
- EAT HEALTHY
- STAY ACTIVE

LEARN MORE AT [www.cdc.gov/diabetes/prevention](http://www.cdc.gov/diabetes/prevention) OR SPEAK TO YOUR DOCTOR

LEARN MORE AT [www.cdc.gov/diabetes/ndep](http://www.cdc.gov/diabetes/ndep) OR SPEAK TO YOUR DOCTOR
PROGRESSION OF CHRONIC KIDNEY DISEASE (CKD)

NORMAL  INCREASED RISK  KIDNEY DAMAGE  REDUCED KIDNEY FUNCTION  KIDNEY FAILURE

Time

Figure credit: https://www.cdc.gov/kidneydisease/prevention-risk.html
Progression of Multiple Myeloma

Asymptomatic

Symptomatic

Smoldering myeloma or MGUS

First-line therapy

Second-line therapy

Third-line or later therapy

Duration of remission decreases with each line of therapy

Figure credit: http://www.myelomarevealed.com/multiple-myeloma-definition-and-statistics/
What is the future of how we treat chronic disease?

• Predicting a patient’s future disease progression

  When will a specific individual with smoldering multiple myeloma (a rare blood cancer) transition to full-blown multiple myeloma?

• Precision medicine

  Which second-line diabetes treatment should we give this patient?
Me

???????

Next 20 years
Me

Similar patient 1

Similar patient 2

Next 20 years
Me

Similar patient 1

Similar patient 2

Drug A or Drug B

???

time
What is the future of how we treat chronic disease?

• Early diagnosis, e.g. of diabetes, Alzheimer's, cancer

Figure sources: NIH, https://www.roche.com/research_and_development/what_we_are_working_on/oncology/liquid-biopsy.htm
What is the future of how we treat chronic disease?

- Continuous monitoring and coaching, e.g. for the elderly, diabetes, psychiatric disease

Figure source (left): http://www.emeraldforhome.com/
What is the future of how we treat chronic disease?

- Discovery of new disease subtypes; design of new drugs; better targeted clinical trials

Figure sources: Haldar et al., Am J Respir Crit Care Med, 2008
Outline for today’s class

1. Brief history of AI and ML in healthcare
2. Why now?
3. Examples of how ML will transform healthcare
4. What is unique about ML in healthcare?
What makes healthcare different?

• Life or death decisions
  – Need robust algorithms
  – Checks and balances built into ML deployment
  – (Also arises in other applications of AI such as autonomous driving)
  – Need fair and accountable algorithms

• Many questions are about unsupervised learning
  – Discovering disease subtypes, or answering question such as “characterize the types of people that are highly likely to be readmitted to the hospital”?

• Many of the questions we want to answer are causal
  – Naïve use of supervised machine learning is insufficient
What makes healthcare different?

- Very little labeled data

Recent breakthroughs in AI depended on *lots* of labeled data!
What makes healthcare different?

• Very little labeled data
  – Motivates semi-supervised learning algorithms

• Sometimes small numbers of samples (e.g., a rare disease)
  – Learn as much as possible from other data (e.g. healthy patients)
  – Model the problem carefully

• Lots of missing data, varying time intervals, censored labels
What makes healthcare different?

• Difficulty of de-identifying data
  – Need for data sharing agreements and sensitivity

• Difficulty of deploying ML
  – Commercial electronic health record software is difficult to modify
  – Data is often in silos; everyone recognizes need for interoperability, but slow progress
  – Careful testing and iteration is needed
Goals for these two days

• Intuition for working with healthcare data
• How to set up as machine learning problems
• Understand which learning algorithms are likely to be useful and when
• Appreciate subtleties in safely & robustly applying ML in healthcare