Machine Learning at BIDMC: augmenting clinical workflows

Steven Horng, MD MMSc FACEP
Clinical Lead for Machine Learning
June 25, 2019
Increasing ED visits

![Graph showing increasing ED visits over time. The y-axis represents hospitals reporting ED visits, and the x-axis represents years from 1993 to 2003. The graph includes bars and a line chart to illustrate the trend.](image-url)
Life Expectancy

Fertility Rate
Shifting Age Demographic

USA

Asia
Supply - demand mismatch
Need to fundamentally change our care processes
Recent Advances in Machine Learning
Speech Recognition

Loud and clear
Speech-recognition word-error rate, selected benchmarks, %

The Switchboard corpus is a collection of recorded telephone conversations widely used to train and test speech-recognition systems.

Sources: Microsoft; research papers
Image Recognition

Ever cleverer
Error rates on ImageNet Visual Recognition Challenge, %

Sources: ImageNet; Stanford Vision Lab
Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs

Varun Gulshan, PhD; Lily Peng, MD, PhD; Marc Coram, PhD; Martin C. Stumpe, PhD; Derek Wu, BS; Arunachalam Narayanaswamy, PhD; Subhashini Venugopalan, MS; Kasumi Widner, MS; Tom Madams, MEng; Jorge Cuadros, OD, PhD; Ramasamy Kim, OD, DNB; Rajiv Raman, MS, DNB; Philip C. Nelson, BS; Jessica L. Mega, MD, MPH; Dale R. Webster, PhD

Normal

Diabetic Retinopathy
Letter

Dermatologist-level classification of skin cancer with deep neural networks

Andre Esteva, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau & Sebastian Thrun

*Nature* 542, 115–118 (02 February 2017)
doi:10.1038/nature21056

Received: 28 June 2016
Accepted: 14 December 2016
Published online: 25 January 2017

Download Citation

Diagnosis | Machine learning | Skin cancer
Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks

Pranav Rajpurkar*
Awni Y. Hannun*
Masoumeh Haghpanahi
Codie Bourn
Andrew Y. Ng

PRANAVSR@CS.STANFORD.EDU
AWNI@CS.STANFORD.EDU
MHAGHPANAHI@IRHYTHMTECH.COM
CBOURN@IRHYTHMTECH.COM
ANG@CS.STANFORD.EDU

34-layer Convolutional Neural Network

SINUS  SINUS  SINUS  SINUS  AFIB  AFIB  AFIB  AFIB
Artificial intelligence

The latest AI can work things out without being taught

Learning to play Go is only the start
Why the sudden jump?

• Better at math?
• Large amounts of Data
• Increased Computational Capacity
Do we really need doctors?
Do doctors really need computers?
How do we get computers to help us and not hurt us?
What are **humans** really good at?
What are **computers** really good at?

\[
P(Z, W; \alpha, \beta) = \int_{\theta} \int_{\varphi} P(W, Z, \theta, \varphi; \alpha, \beta) \, d\varphi \, d\theta
\]

\[
= \int_{\varphi} \prod_{i=1}^{K} P(\varphi_i; \beta) \prod_{j=1}^{M} \prod_{t=1}^{N} P(W_{j,t|\varphi Z_{j,t}}) \, d\varphi \int_{\theta} \prod_{j=1}^{M} P(\theta_j; \alpha) \prod_{t=1}^{N} P(Z_{j,t|\theta_j}) \, d\theta.
\]
<table>
<thead>
<tr>
<th>Old Paradigm</th>
<th>New Paradigm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capture Once, Use Once</td>
<td>Capture Once, Use Many Times</td>
</tr>
<tr>
<td>Know exactly how data will be used</td>
<td>Can not anticipate future secondary use of data</td>
</tr>
<tr>
<td>Schema on Write (highly structured data)</td>
<td>Schema on Read (less structured data)</td>
</tr>
</tbody>
</table>
Can’t just be about storage
Needs to provide value
Capture Data → Store Data → Retrieve Data → Use Data
Capture Data → Store Data → Retrieve Data → Use Data
1. **Followup (Progress note)**
   - **06/10/2015** Ayad Hamdan (Hematology/Oncology)
   - Heart: Regular rate and rhythm. Normal S1, S2. No murmurs, rubs, or gallops. Lungs: Clear to auscultation bilaterally without rhonchi, rales.

2. **CHRONIC OBSTRUCTIVE PULMONARY DISEASE (Initial note)**
   - **06/02/2015** Douglas Beach (Pulmonary)

3. **Discharge Summary (Disch Sum)**
   - **04/21/2015** Meghan J Campo (unknown)
   - To assess due to body habitus CARDIAC: tachycardic, S1/S2, no murmurs, gallops, or rubs LUNG: Diffusely wheezing ABDOMEN: nondistended, +BS, nottender...cough compred to prior CV: Regular rate and rhythm, normal S1 + S2, no murmurs, rubs, gallops Abdomen: soft, non-tender, non-distended, no pain on deep palpation.

4. **Follow-up (Progress note)**
   - **09/10/2014** Ayad Hamdan (Hematology/Oncology)
   - Heart: Regular rate and rhythm. Normal S1, S2. No murmurs, rubs, or gallops. Lungs: Clear to auscultation bilaterally without rhonchi, rales.
Contextual Information Retrieval

**DM:** metformin
**HTN:** metoprolol tartrate
**Tobacco:** None
**FH:** Father w/ hx of etoh abuse, passed away from MI @ 70s, first one in 50s. Paternal grandmother with CVA and paternal grandfather with MI in the 70s. mother with OCD. [12/19/2014 17:13]

**Echo**
09/30/2013
Normal left ventricular cavity size with mild global hypokinesis c/w diffuse process (toxin, metabolic, etc., CAD cannot be fully excluded, but less likely). Mild mitral regurgitation with normal valve morphology. Compared with the report of the prior study (images unavailable for review) of 8/19/2004, global left ventricular ejection fraction is more reduced. **CLINICAL IMPLICATIONS:** Based on 2007 AHA endocarditis prophylaxis recommendations, the echo findings indicate prophylaxis is NOT recommended. Clinical decisions regarding the need for prophylaxis should be based on clinical and echocardiographic data. Electronically signed by Warren J. Manning, MD on 10/1/2013 17:43.

**Echo**
10/20/2015
02/09/2006
**NUCLEAR**
Average exercise tolerance. No anginal symptoms or ischemic ST segment changes. Nuclear report sent separately.

**NUCLEAR**
07/28/2005
Non-anginal chest discomfort and no ischemic ST changes at limited/fair exercise capacity. Nuclear imaging under separate report.

**NUCLEAR**
10/11/2004
No anginal type symptoms or ischemic EKG changes. Nuclear report sent separately.

**Echo**
05/25/2004
08/19/2004
1. Coronary angiography of this right dominant circulation revealed single vessel CAD. The LMCA had no significant stenoses. The LAD had minor luminal irregularities. The LCX had a 90% hazy lesion of a large OM1 with TIMI 2 flow. The RCA had no obstructive disease. 2. Resting hemodynamics showed elevated LV filling pressures. 3. Left ventriculography showed no wall motion abnormalities and an EF of 56%. 4. Successful stenting of the OM1 with 2.5 x 20 mm Taxus DES at 14 atm and successful PTCA of the mid OM1 with 2.0 x 15 mm open sail balloon at 6 atm with no residual stenosis in the stented segment, 25% residual stenosis in the mid OM1, no dissection and timi 3 flow. See PTCA comments.

**Echo**
05/25/2004
03/30/2004
**NUCLEAR**
Capture Data → Store Data → Retrieve Data → Use Data
Structured Chief Complaints

- Almost always entered as free text
- Structured chief complaints have great utility
  - Clinical Decision support
  - Research
  - QA
  - Epidemiology
<table>
<thead>
<tr>
<th></th>
<th>Small vocabulary</th>
<th>Large vocabulary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy to learn and use</td>
<td></td>
<td>Difficult to learn and use</td>
</tr>
<tr>
<td>Reproducible</td>
<td></td>
<td>Quality dependent on user training and motivation</td>
</tr>
<tr>
<td></td>
<td>Too general, often times not accurately representing the chief complaint</td>
<td>Very specific, well representing the chief complaint if used properly</td>
</tr>
</tbody>
</table>
Probabilistic Inference for Chief Complaint UI

• **Essentially a Diagnosis Problem**

• Probabilistic Inference based on all available data (Age, Sex, Vital signs, Emergency Severity Index, Medical Problems, labs, etc.)

• Implemented as Contextual Auto-complete
Chief Complaint UI
Chief Complaint UI
Performance Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Multiclass SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best-5</td>
<td>0.757</td>
</tr>
<tr>
<td>Best-10</td>
<td>0.825</td>
</tr>
<tr>
<td>Discounted Cumulative Gain</td>
<td>0.613</td>
</tr>
</tbody>
</table>
Compliance

- Structured chief complaint went from ~35% to ~95% (p<0.001)

- Nurses don’t realize it’s there
Can be applied to usage for any standard ontology

- Snomed CT
- ICD10
- CPT
Automated Triggering

• Clinical pathways
• Decision Support
• Order Sets
• Research eligibility screening
Benefits of Automated Triggers

• Doesn’t rely on user knowing that a function exists (important given our large number of transient users)

• Independent of user’s knowledge and motivation
<table>
<thead>
<tr>
<th>Clinical State</th>
<th>Description</th>
<th>Value</th>
<th>Anchors</th>
<th>State</th>
<th>Reasoning</th>
</tr>
</thead>
<tbody>
<tr>
<td>chestpain_acute</td>
<td>cardiac chest pain</td>
<td>.81</td>
<td></td>
<td></td>
<td># chest pain</td>
</tr>
<tr>
<td>cardiac_acute</td>
<td></td>
<td>.07</td>
<td></td>
<td></td>
<td>also in ddx is angina or mi</td>
</tr>
<tr>
<td>cancer_history</td>
<td></td>
<td>.04</td>
<td></td>
<td></td>
<td>collins is a simnum year old female with a history of htn , mdd , and asthma who presents to the ed for chest pain and shortness of breath</td>
</tr>
<tr>
<td>abdominalpain_acute</td>
<td>acute abdominal pain</td>
<td>.03</td>
<td></td>
<td>chest pain</td>
<td></td>
</tr>
<tr>
<td>infection_acute</td>
<td></td>
<td>.03</td>
<td></td>
<td>cxr</td>
<td></td>
</tr>
<tr>
<td>asthma-copd_acute</td>
<td>active asthma/copd</td>
<td>.02</td>
<td></td>
<td></td>
<td>acute onset chest pressure with dyspnea possibly related to history of asthma and could be an exacerbation however she is not wheezy on exam</td>
</tr>
<tr>
<td>diabetes_history</td>
<td>diabetes</td>
<td>.02</td>
<td></td>
<td>age_63</td>
<td></td>
</tr>
<tr>
<td>syncope_acute</td>
<td></td>
<td>.02</td>
<td></td>
<td></td>
<td>she was also so uncomfortable that she had to get out of bed and lay on the floor</td>
</tr>
<tr>
<td>uti_acute</td>
<td></td>
<td>.01</td>
<td></td>
<td>ms</td>
<td></td>
</tr>
<tr>
<td>backpain_acute</td>
<td>acute back pain</td>
<td>.01</td>
<td></td>
<td># chest pain</td>
<td></td>
</tr>
<tr>
<td>allergicreaction_acute</td>
<td>acute allergic reaction</td>
<td>.01</td>
<td></td>
<td></td>
<td>acute onset chest pressure with dyspnea possibly related to history of asthma and could be an exacerbation however she is not wheezy on exam</td>
</tr>
<tr>
<td>chf_acute</td>
<td></td>
<td>.01</td>
<td></td>
<td></td>
<td># dyspnea :</td>
</tr>
<tr>
<td>psych_acute</td>
<td>acute psychiatric condition</td>
<td>0</td>
<td></td>
<td></td>
<td>she was also so uncomfortable that she had to get out of bed and lay on the floor</td>
</tr>
<tr>
<td>headache_acute</td>
<td>active headache</td>
<td>0</td>
<td></td>
<td>a / p :</td>
<td></td>
</tr>
</tbody>
</table>
Learning a Health Knowledge Graph from Electronic Medical Records

Maya Rotmensch ¹, Yoni Halpern ², Abdulhakim Tlimat ³, Steven Horng ³,⁴ & David Sontag ⁵,⁶

Demand for clinical decision support systems in medicine and self-diagnostic symptom checkers has substantially increased in recent years. Existing platforms rely on knowledge bases manually compiled through a labor-intensive process or automatically derived using simple pairwise statistics. This study explored an automated process to learn high-quality knowledge bases linking diseases and symptoms directly from electronic medical records. Medical concepts were extracted from 273,174 de-identified patient records and maximum likelihood estimation of three probabilistic models was used to automatically construct knowledge graphs: logistic regression, naïve Bayes classifier and a Bayesian network using noisy OR gates. A graph of disease-symptom relationships was elicited from the learned parameters and the constructed knowledge graphs were evaluated and validated, with permission, against Google’s manually-constructed knowledge graph and against expert physician opinions. Our study shows that direct and automated construction of high-quality health knowledge graphs from medical records using rudimentary concept extraction is feasible. The noisy OR model produces a high-quality knowledge graph reaching precision of 0.85 for a recall of 0.6 in the clinical evaluation. Noisy OR significantly outperforms all tested models across evaluation frameworks (p < 0.01).
Figure 2. Workflow of modeling the relationship between diseases and symptoms and knowledge graph construction, for each of our 3 models (naive Bayes, logistic regression and noisy OR).
Cardiac CP Order Set

To be drawn immediately

Initial
- IVs: Place IV (saline lock); flush per protocol
- Noninvasive Patient Monitoring: Continuous Cardiac monitoring
- Noninvasive Patient Monitoring: Continuous Pulse oximetry

EKG (pick 1)
- EKG (to be performed): Indication: Chest Pain
- EKG (to be performed): Indication: Dyspnea

Laboratory
- CBC + Diff
- Chem-7
- Troponin

Aspirin (pick 1)
- Aspirin 243 mg PO *Allergy
- Aspirin 324 mg PO *Allergy
- Aspirin taken before arrival

Imaging
- XR Chest PA & Lateral
- XR Chest AP Portable

Stress
- Exercise Stress Test (ETT)
- Stress Echo
- Myocardial Perf with Exercise
- Myocardial Perf with Pharm Stress

+ Other

Order More          Order+Sign
Contextual Information Retrieval

---

**Cardiac Info (10)**

**DM:** metformin  
**HTN:** metoprolol tartrate  
**Tobacco:** None  
**FH:** Father w/ hx of etoh abuse, passed away from MI @ 70s, first one in 50s. Paternal grandmother with CVA and paternal grandfather with MI in the 70s. mother with OCD. [12/19/2014 17:13]

**10/20/2015**  
**Echo**  
Normal left ventricular cavity size with mild global hypokinesis c/w diffuse process (toxin, metabolic, etc., CAD cannot be fully excluded, but less likely). Mild mitral regurgitation with normal valve morphology. Compared with the report of the prior study (images unavailable for review) of 8/19/2004, global left ventricular ejection fraction is more reduced. CLINICAL IMPLICATIONS: Based on 2007 AHA endocarditis prophylaxis recommendations, the echo findings indicate prophylaxis is NOT recommended. Clinical decisions regarding the need for prophylaxis should be based on clinical and echocardiographic data. Electronically signed by Warren J. Manning, MD on 10/1/2013 17:43.

**10/10/2008**  
**NUCLEAR**  
Average exercise tolerance. No anginal symptoms or ischemic ST segment changes. Nuclear report sent separately.

**02/09/2008**  
**NUCLEAR**  
No anginal symptoms or ischemic EKG changes noted. Nuclear report sent separately.

**07/28/2005**  
**NUCLEAR**  
Non-anginal chest discomfort and no ischemic ST changes at limited/fair exercise capacity. Nuclear imaging under separate report.

**10/11/2004**  
**NUCLEAR**  
No anginal type symptoms or ischemic EKG changes. Nuclear report sent separately.

**08/19/2004**  
**Echo**  
05/25/2004  
**Cath**  
1. Coronary angiography of this right dominant circulation revealed singe vessel CAD. The LMCA had no significant stenoses. The LAD had minor luminal irregularities. The LCX had a 90% hazy lesion of a large OM1 with TIMI 2 flow. The RCA had no obstructive disease.  
2. Resting hemodynamics showed elevated LV filling pressures.  
3. Left ventriculography showed no wall motion abnormalities and an EF of 56%.  
4. Successful stenting of the OM1 with 2.5 x 20 mm Taxus DES at 14 atm and successful PTCA of the mid OM1 with 2.0 x 15 mm open sail balloon at 6 atm with no residual stenosis in the stented segment, 25% residual stenosis in the mid OM1, no dissection and timi 3 flow. See PTCA comments.

**05/25/2004**  
**Echo**

**03/30/2004**  
**NUCLEAR**
Checklists

• Unstructured Data

• Scanned Images (consent forms, DNR/DNI, etc.)
Risk Stratification
Triage waiting room sorting

• Triage is most necessary when demand > supply

• Same time when very difficult to allocate additional resources to re-prioritize / monitor patients

• Automated methods using already collected data would be helpful
Probabilistic Inference for Triage Rack Sorter

• **Essentially a Risk Stratification Problem**

• Probabilistic Inference based on all available data (Age, Sex, Vital signs, Emergency Severity Index, Medical Problems, labs, etc.)

• Implemented as graphical risk stratification meter
## Risk Stratification

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Algorithm AUC</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>derivation</td>
<td>Validation</td>
<td></td>
</tr>
<tr>
<td>All ED Patients</td>
<td></td>
<td>(n=75,992)</td>
<td>(n=18,981)</td>
</tr>
<tr>
<td>Infection</td>
<td>0.88</td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td>28 day mortality</td>
<td>0.94</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>ICU admission</td>
<td>0.91</td>
<td>0.90</td>
<td></td>
</tr>
</tbody>
</table>

> Predicting 28 day in-hospital mortality

*True Positive Sensitivity vs 1-Specificity False Positive*  

AUC: 0.92
Triage
Quantification of pulmonary edema as a surrogate for fluid status
# Table 1

Radiologic scoring of pulmonary edema on bedside chest radiographs

<table>
<thead>
<tr>
<th>Variables</th>
<th>Score</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mild</td>
<td>Moderate</td>
<td>Severe</td>
<td></td>
</tr>
<tr>
<td>Hilar vessels</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Enlarged</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Increased in density</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Blurred</td>
<td>3</td>
<td>6</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Kerley B lines</td>
<td>4</td>
<td>-</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Micronoduli</td>
<td>4</td>
<td>-</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Widening of interlobular fissure</td>
<td>4</td>
<td>8</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Peribronchial and perivascular cuffs</td>
<td>4</td>
<td>8</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Extensive perihilar haze</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Diffuse increase in density</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td></td>
</tr>
</tbody>
</table>

kappa 0.68
AUC=0.75
No Edema vs. Mild/Moderate/Severe edema

AUC=0.80
No/Mild Edema vs Moderate/Severe Edema
Summary

- Big data / Machine Learning will help us leverage the data already being generated
- Right information at the right time
- Design for the user (save time, facilitate workflow)
- Build decision support directly into the user interfaces