A Framework for Artificial Intelligence in the Perioperative Period

Ira S. Hofer, MD
Assistant Professor
Department of Anesthesiology, UCLA

Personal Background

- Assistant Professor of Anesthesiology at UCLA
- Director, Division of Bioinformatics, Department of Anesthesiology
- 10 Years working in EMR Data
- Research Interest: Understanding, quantifying and mitigating perioperative risk
- Financial Disclosures: Copyright on software to extract data from EMR
- Funding: NIH 1R01AG059815-01 Co-I
What is Machine Learning

- Features
  - Feature A
  - Feature B
  - Feature C
  - Feature D
  - ...
  - Feature n

- Model
  - DNN
  - Random Forest
  - Elastic Net

Result

---

Value Proposition of Smart Data

---

Opening the Black Box: Understanding the Science Behind Big Data and Predictive Analytics

Ira S. Hofer, MD,* Eran Halperin, PhD,†† and Maxime Cannesson, MD, PhD*
Triangulation: Using multiple data points to arrive at a better definition

Algorithm: >99% Correct

Manual Review: ~70% Correct
Patient Phenotyping

- Phenotype patients based on RCRI
  - Diabetes, CAD, CHF, Cerebrovascular Disease
- Algorithms look at
  - Past medical history
  - Previous procedures
  - Previous coding
  - Lab results
  - Medication usage

• Prevalence of each of the disease (A: CHF, B: CAD, C: CVD, D: DM) as determined by each of the four methods: diagnosis algorithms, ICD codes, Anesthesiologist preoperative note, and Manual review.

Automated Assessment of Existing Patient’s Revised Cardiac Risk Index Using Algorithmic Software

Ira S, Hoffer, MD, * Drew Cheng, MD, * Tristan Grogan, MS, * Yohoi Fujimoto, MD, PhD, † Takashiye Yamasita, MD, PhD, ‡ Lauren Beck, MD, * Maxime Corning, MD, PhD, * and Aman Minhajian, MD, PhD
AUC plots for predicting in hospital mortality

Single data points are never sufficient to classify a patient
What is Machine Learning

Features
- Feature A
- Feature B
- Feature C
- Feature D
- ...
- Feature n

Model
- DNN
- Random Forest
- Elastic Net

Result

The **right** intervention for the **right** patient at the **right** time – *every* time
Figure 2: Progressive Value of Smart Data

Table 1: Select scoring systems available for assessment of postoperative risk.

<table>
<thead>
<tr>
<th>Scoring system</th>
<th>Year</th>
<th>Number of variables</th>
<th>Intraperative variables</th>
<th>Outcome-predicted</th>
<th>Simplicity</th>
<th>Objectivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Postoperative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>APACHE II</td>
<td>1984</td>
<td>2</td>
<td>No</td>
<td>No</td>
<td>Simple</td>
<td>Subjective</td>
</tr>
<tr>
<td>APACHE III</td>
<td>1991</td>
<td>12</td>
<td>Yes</td>
<td>Mortality &amp; morbidity</td>
<td>Complex</td>
<td>Objective</td>
</tr>
<tr>
<td>NSQIP</td>
<td>1997</td>
<td>5</td>
<td>Yes</td>
<td>Mortality &amp; morbidity</td>
<td>Complex</td>
<td>Objective</td>
</tr>
<tr>
<td>SACS</td>
<td>2001</td>
<td>3</td>
<td>Yes</td>
<td>Mortality &amp; morbidity</td>
<td>Complex</td>
<td>Objective</td>
</tr>
<tr>
<td>Intensive care</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>APACHE II†</td>
<td>1984</td>
<td>&gt; 10</td>
<td>No</td>
<td>Mortality</td>
<td>Complex</td>
<td>Objective</td>
</tr>
<tr>
<td>SAPS II†</td>
<td>1985</td>
<td>&gt; 10</td>
<td>No</td>
<td>Mortality</td>
<td>Complex</td>
<td>Objective</td>
</tr>
<tr>
<td>MPM-II*</td>
<td>1995</td>
<td>&gt; 10</td>
<td>No</td>
<td>Mortality</td>
<td>Complex</td>
<td>Objective</td>
</tr>
<tr>
<td>SOFA</td>
<td>1994</td>
<td>6</td>
<td>No</td>
<td>None</td>
<td>Complex</td>
<td>Objective</td>
</tr>
<tr>
<td>MODS</td>
<td>1995</td>
<td>6</td>
<td>No</td>
<td>None</td>
<td>Complex</td>
<td>Objective</td>
</tr>
</tbody>
</table>

* †Based on the version of the scoring system used.

ASA: American Society of Anesthesiologists; APACHE: Acute Physiology and Chronic Health Evaluation; SAPS: Simplified Acute Physiology Score; MPM: Mortality Probability Model; SOFA: Sequential Organ Failure Assessment; MODS: Multi-Organ Dysfunction Score; P-POSSUM: Portsmouth Physiological and Operative Severity Score for the Enumeration of Mortality and Morbidity; F-POSSUM: Estimation of Physiological, Ability and Surgical Stress; NSQIP: National Surgical Quality Improvement Program; SAS: Surgical Age Risk Score.
Precision vs Accuracy

- High Accuracy Low Precision
- Low Accuracy High Precision
- Low Accuracy Low Precision
- High Accuracy High Precision
Preoperative predictions of in-hospital mortality using electronic medical record data


Fig. 2. Receiver operating characteristic (ROC) and precision recall curves for the random forest model. Plots were generated using cross-validated predictions on the entire dataset. ROC curves (a) show the false positive rate on the x-axis and the true positive rate on the y-axis. The optimal point is the upper-left corner. Precision-recall curves (b) show the recall on the x-axis and precision on the y-axis. The optimal point is in the upper-right corner.

Fig. 3. Heatmap of Preoperative Risk vs. Postoperative Risk. Preoperative (x-axis) and postoperative (y-axis) risk scores were binned by percentile, and the count per bin visualized as a heatmap in log scale. In (a) all patients are displayed, and in (b) only the in-hospital mortalities are shown. 78% of patients who die and have a pre-operative risk percentile below 95% have an increased postoperative risk percentile.
SMART Screen
Figure 2: Progressive Value of Smart Data
Predicting Blood Pressure Response to Fluid Bolus Therapy Using Attention-Based Neural Networks for Clinical Interpretability

Uma M. Girkar, Rya Uchinode, Liwei H. Leibman, Peter Sadowski, Lee O'Keefe, and Wei-Hung Wong

1Computer Science and Artificial Intelligence Laboratory, MIT, Cambridge, MA, USA
2Emergency Department, Beth Israel Deaconess Medical Center, Boston, MA, USA
3Laboratory for Computational Physiology, MIT, Cambridge;
ruchinoidb@med.harvard.edu

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Timsteps</th>
<th>Accuracy</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Raw features</td>
<td>Raw features</td>
</tr>
<tr>
<td>L1-regularized logistic regression</td>
<td>-</td>
<td>0.706</td>
<td>0.680</td>
</tr>
<tr>
<td>L2-regularized logistic regression</td>
<td>-</td>
<td>0.699</td>
<td>0.695</td>
</tr>
<tr>
<td>Multiple layer perception</td>
<td>-</td>
<td>0.712</td>
<td>0.688</td>
</tr>
<tr>
<td>LSTM</td>
<td>12</td>
<td>0.751</td>
<td>0.705</td>
</tr>
<tr>
<td>GRU</td>
<td>12</td>
<td>0.748</td>
<td>0.721</td>
</tr>
<tr>
<td>LSTM + Attention</td>
<td>12</td>
<td>0.747</td>
<td>0.727</td>
</tr>
<tr>
<td>LSTM</td>
<td>12</td>
<td>0.747</td>
<td>0.716</td>
</tr>
<tr>
<td>LSTM</td>
<td>36</td>
<td>0.814</td>
<td>0.827</td>
</tr>
<tr>
<td>LSTM</td>
<td>36</td>
<td>0.820</td>
<td>0.794</td>
</tr>
<tr>
<td>LSTM + Attention</td>
<td>36</td>
<td>0.819</td>
<td>0.820</td>
</tr>
<tr>
<td>LSTM</td>
<td>36</td>
<td>0.812</td>
<td>0.777</td>
</tr>
<tr>
<td>LSTM</td>
<td>72</td>
<td>0.843</td>
<td>0.834</td>
</tr>
<tr>
<td>LSTM</td>
<td>72</td>
<td>0.848</td>
<td>0.836</td>
</tr>
<tr>
<td>LSTM + Attention</td>
<td>72</td>
<td>0.852</td>
<td>0.831</td>
</tr>
</tbody>
</table>

Table 1: Model performance in accuracy and AUC between different experimental settings. **Boldface** denotes the best performance in each group.

Value Proposition of Smart Data

SMART DATA

Cognitive
Prescriptive
Predictive
Descriptive

BIG DATA
Velocity Variety Volume Veracity

Figure 2: Progressive Value of Smart Data
Cognitive

Thank you

• UCLA PDW Team
  • Eilon Gabel MD
  • Theodora Wingert MD
  • Vali Salari PhD
  • Drew Cheng MD
  • Johnny Quach
  • Olivia Vallejo
  • Evan Moh

• Collaborators
  • Eran Halperin, PhD
  • Sriram Sankararaman, PhD
  • Eleazar Eskin, PhD
  • Christine Lee
  • Brian Hill
  • Nadav Rakocz
  • Robert Brown

• Mentors
  • David Reich MD
  • Adam Levine MD
  • Maxime Cannesson MD, PhD
  • Barbara Van de Wiele, MD
Use Clustering to group patients