“Big Data” In Analysis Of Streaming Physiologic Data: Implications For Health Care

Rishi Kamaleswaran, PhD
Center for Biomedical Informatics,
Division of Critical Care Medicine,
Department of Pediatrics,
The University of Tennessee Health Science Centre, Memphis TN
@rkamaleswaran
ADEPT4 Workshop 2017
Big Data to Knowledge Initiative

PIC-SURE Precision Medicine June 2015 Patient Driven meeting (BD2K Center of Excellence, Harvard University)

“How do we securely collect and analyze distinct data streams in real-time to guide medical decisions?”
Big Data in the NICU

182 million readings a day per patient.
1 - 5 readings an hour
Characteristics of Physiologic Data

• Unlike other sensor-based big data sources
  • Has variances in signal quality, strength and frequently non-stationary
  • Changes over time as the infant grows and matures
  • New normal are established day-by-day
  • Contains numerous artefacts
  • Wasn’t meant to be streamed!
Bringing analytics to the bed-side

• Access the physiologic data to analyze every bit of information that can lead to *personalized* and *earlier interventions*

• The provision of real-time physiologic analysis requires a multidimensional approach:
  • *Multiple conditions*
  • *Multiple streams* of data
  • For which *multiple behaviours* can exist

• In addition, integrate of
  • Real-time synchronous medical device data
  • Asynchronous clinical data
Sensor Analytics - Architectural Requirements and Development

Interactive Exploration of Event Streams
Event Streams—A hierarchy

- Primitive Event
- Complex Event
- Multidimensional Event
- Physical Event
- Sensor_1
- Sensor_2
- Environment
Event Stream Processing (InfoSphere Streams)
Event Stream Processing (Spark Streaming)
Spark Streaming – Structured Streaming

Structured Streaming Model
users express query on streaming data using a batch API; Spark incrementalizes them to run on streams
System Architecture for Sensor Analytics at Le Bonheur Children’s Hospital
Analytics at the Point of Care - Past work

“Right information at the right time”
Kamaleswaran et al., 2010: A method for clinical and physiological event stream processing

Mean blood pressure vs Gestational age [1]

Kamaleswaran et al., 2012: Integrating complex business processes for knowledge-driven clinical decision support systems
Clinical Domain Challenges—Apnea of Prematurity

• Apnoea of prematurity: *cessation of breathing* for more than 20 seconds [2]

• Challenge to *visually assess* the etiology and type of apnoea

• Requires extensive monitoring to determine diagnosis

• In lieu of that monitoring, bedside staff broadly classify any cardiorespiratory event as a “Neonatal Spell” [3]


Clinical Domain Challenges—Nosocomial Infection

• **Nosocomial infection** (sepsis) is a common hospital-borne infection for babies receiving care in the NICU [4]

• A cocktail of antibiotics are prepared and administered even before diagnosis.

• **Heart-Rate Variability** seen as a potential indicator of the onset of neonatal sepsis [5]

Neonatal Sepsis and Spells

Question: Can we identify neonatal spells prior to onset of nosocomial infection?

- Thommandram et al., built a neonatal **spells algorithm** [3] for detecting spells activity in real-time

**Classifications:**
- central apnoea
- vagal apnoea
- obstructive
- obstructive central
- central obstructive
- isolated blood oxygen desaturation
- isolated bradycardia.

[Diagram showing baseline, resp pause, spell, and ECG, modified from Sale, 2010]
## Events as Sequences

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central</td>
<td>↓ Resp.</td>
<td>↓ Heart Rate</td>
<td>↓ O₂</td>
<td>↑ Resp.</td>
<td>↑ Heart Rate</td>
</tr>
<tr>
<td>Vagal</td>
<td>↓ Resp.</td>
<td>↓ O₂</td>
<td>↑ Resp.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>↓ Heart Rate</td>
<td></td>
<td>↑ Heart Rate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obstructive</td>
<td>↓ Heart Rate</td>
<td>↓ O₂</td>
<td></td>
<td>↑ Heart Rate</td>
<td></td>
</tr>
<tr>
<td>Obstructive (Incremental)</td>
<td>↓ Heart Rate</td>
<td>↓ O₂</td>
<td>↑ Resp.</td>
<td>↑ Heart Rate</td>
<td></td>
</tr>
<tr>
<td>Obstructive Central</td>
<td>↓ Heart Rate</td>
<td>↓ O₂</td>
<td>↓ Resp.</td>
<td>↑ Resp.</td>
<td>↑ Heart Rate</td>
</tr>
<tr>
<td>Central Obstructive</td>
<td>↓ Resp.</td>
<td>↓ Heart Rate</td>
<td>↓ O₂</td>
<td>↑ Resp.</td>
<td>↓ Heart Rate</td>
</tr>
<tr>
<td>Desaturation</td>
<td>↓ O₂</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bradycardia</td>
<td>↓ Heart Rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
PhysioEx: Visual analysis of physiological event streams [6]

Central (56 Secs.) Time: July 25, 2010 04 PM

1. Heartrate (HR)
2. Oxygen (SpO₂)
3. Respiratory (IRW)
Data-driven Approaches: Current Work

Knowledge Discovery for Point of Care Applications
Continuous feature extraction: Probabilistic Symbolic Pattern Recognition (PSPR) [7]

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>[-∞      0.10]</td>
</tr>
<tr>
<td>b</td>
<td>[0.10    0.20]</td>
</tr>
<tr>
<td>c</td>
<td>[0.20   0.40]</td>
</tr>
<tr>
<td>d</td>
<td>[0.40  1.0]</td>
</tr>
<tr>
<td>e</td>
<td>[1.0    ∞]</td>
</tr>
</tbody>
</table>

Paroxysmal Atrial Fibrillation (PAF) Prediction

Sampling frequency indicates influence on accuracy of predictive model for PAF prediction

<table>
<thead>
<tr>
<th>Sampling frequency, $f_s$ (Hz)</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training data</td>
</tr>
<tr>
<td>128</td>
<td>76.67</td>
</tr>
<tr>
<td>64</td>
<td>81.67</td>
</tr>
<tr>
<td>32</td>
<td>85</td>
</tr>
<tr>
<td>16</td>
<td>85.33</td>
</tr>
<tr>
<td>8</td>
<td>84.33</td>
</tr>
</tbody>
</table>
Deep Learning – Convolutional Neural Networks for Sensor Timeseries Analytics

• ConvNets provide a method of automatic feature extraction using a combination of convolutional filters and dimensional reduction
• Accuracies greater than humans have been observed using ConvNets on image data
• Features are extracted using multiple filters applied to the same image
Deep Learning – Convolutional Neural Networks

• ECGs are continuously sampled waveform data most remarkably distinguished by the prominent ‘R’ peak.
• An ECG signal can be sampled anywhere from a couple of minutes to hours
• We abstract a continuous waveform and segment it by it’s individual complex

A sample ECG waveform at 300 Hz

A sample ECG complex at 300 Hz
Deep Learning – Convolutional Neural Networks

• Each complex is segmented into a square matrix that is then fed into the ConvNet for feature extraction
Influences of Noise on Sensor Data
Deep Convolutional Neural Network Topology

Total params: 2,042,820
Deep Convolutional Neural Network Results

Train Score: acc: 89.14% Test Score: acc: 84.69%
Take home message

• Every breath, every beat – Critical Care
  • Event stream processing a potent tool to uncover new knowledge
  • Collection of high quality data is still a pressing challenge
• Not enough to simply “analyze” the data
  • Knowledge is a holistic experience
  • Feedback loop reinforces algorithm learning
• Deep learning methods offer novel means to classify difficult and very noisy ECG sensor data
Thank You!

Rishi Kamaleswaran

rkamesales@uthsc.edu
@rkamesaleswaran