Signal Processing and Machine Learning for Intelligent Patient Monitoring

Xiaopeng Zhao
Department of MABE
National Institute for Mathematical and Biological Synthesis
University of Tennessee, Knoxville
xzhao9@utk.edu
Challenges in Patient Monitoring

• Heterogeneities in data
differences between patients, devices, hospitals, …
• Noise and errors
• Gaps in data
• Desired signals not unavailable
• Nonlinear and nonstationary physiological conditions

Need signal processing and machine learning techniques to predict, estimate, and diagnose, …
Problem Statement: Mind the Gap in Medical Signals

Constant monitoring of multiple physiological signals is essential for clinical diagnosis, treatment, and research. However, disruption or loss of signals can frequently happen, either due to errors in sensors or due to external disturbances.
Method and Results

Results: 2nd place in Physionet Challenge 2010
Blood Pressure

ECG

Photoplethysmography

Respiration
Problem Statement: Improving the Quality of ECGs Collected Using Cell Phones

Artifacts in ECG
• Degrade diagnostic information
• Increase number of false alerts
• Delay proper treatment or cause a wrongful, detrimental treatment
• Increase workload of health providers
Errors in ECG

Normal ECGs may look “Abnormal”  Abnormal ECGs may look “Normal”

Electrodes Correctly Placed

Electrodes Misplaced

Method and Results

Matrix of Regularity

Outcome

Accuracy = 95.1%
Sensitivity = 88.4%
Specificity = 97.0%

Receiver Operating Characteristics
AUC = 0.97

Performance is comparable to human experts

Results: 1st, 1st, 3rd in Physionet Challenge 2011
Problem Statement: ICU Mortality Prediction

• Patient specific mortality prediction using physiological measurements
• Application: examine the efficacy of medications, care guidelines, surgery, and other interventions
Importance of severity evaluation

- ICU improves outcome for seriously ill patients significantly
  - However, it is expensive: in 2005, mean ICU cost is $31,574 for patients requiring mechanical ventilation and $12,931 for those not requiring mechanical ventilation

- Severity evaluation
  - allows to restrict ICUs to patients most at need
  - provides doctors a way to judge the treatment method
Data

- Physionet data: 4000 ICU records
- 5 static variables at admission:
  - Age, Sex, Height, Weight,
  - ICU Type (Coronary Care, Cardiac Surgery Recovery, Medical, or Surgical)
- 37 dynamic variables recorded multiple times in 48 hours
## Dynamic Variables

<table>
<thead>
<tr>
<th>Albumin</th>
<th>HCT</th>
<th>pH</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALP</td>
<td>HR</td>
<td>Platelets</td>
</tr>
<tr>
<td>ALT</td>
<td>K</td>
<td>RespRate</td>
</tr>
<tr>
<td>AST</td>
<td>Lactate</td>
<td>SaO2</td>
</tr>
<tr>
<td>Bilirubin</td>
<td>Mg</td>
<td>SysABP</td>
</tr>
<tr>
<td>BUN</td>
<td>MAP</td>
<td>Temp</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>MechVent</td>
<td>TropI</td>
</tr>
<tr>
<td>Creatinine</td>
<td>Na</td>
<td>TropT</td>
</tr>
<tr>
<td>DiasABP</td>
<td>NIDiasABP</td>
<td>Urine</td>
</tr>
<tr>
<td>FiO2</td>
<td>NIMAP</td>
<td>WBC</td>
</tr>
<tr>
<td>GCS</td>
<td>NISysABP</td>
<td>Weight</td>
</tr>
<tr>
<td>Glucose</td>
<td>PaCO2</td>
<td></td>
</tr>
<tr>
<td>HCO3</td>
<td>PaO2</td>
<td></td>
</tr>
</tbody>
</table>
Occurrence frequency of each variable
Example record

- Time, Parameter, Value
- 00:00,RecordID,12539
- 00:00,Age,54
- 00:00,Gender,0
- 00:00,Height,-1
- 00:00,ICUType,4
- 00:00,Weight,-1
- 00:07,GCS,15
- 00:07,HR,73
- 00:07,NISysABP,147
- 00:07,RespRate,19
- 00:07,Temp,35.1
- 00:07,Urine,900
- 00:37,HR,77
- 00:37,NIDiasABP,58
- 00:37,NIMAP,91
- 00:37,NISysABP,157
- 00:37,RespRate,19
- 00:37,Temp,35.6
- 01:37,NIDiasABP,62
- 01:37,NIMAP,87
- 01:37,NISysABP,137
- 01:37,RespRate,18
- 01:37,Urine,30
- 02:37,HR,62
- 02:37,NIDiasABP,52
- 02:37,NISysABP,123
- 02:37,RespRate,19
- 02:37,Urine,170
- 03:08,HCT,33.7
- 03:37,GCS,15
- 03:37,HR,80
- 03:37,NIDiasABP,114
- 03:37,NISysABP,114
- 03:37,RespRate,20
- 03:37,Temp,37.8
- 03:37,Urine,60
- 04:37,HR,74
- 04:37,RespRate,20
- 05:37,NISysABP,110
- 05:37,RespRate,17
- 05:37,Urine,170
- 07:37,GCS,15
- 07:37,HR,64
- 07:37,NIDiasABP,49
- 07:37,NIMAP,68.33
- 07:37,NISysABP,107
- 07:37,RespRate,15
- 07:37,Temp,38.1
- 07:37,Urine,120
- 08:37,HR,64
- 08:37,NIDiasABP,56
- 08:37,NIMAP,71.33
- 08:37,NISysABP,102
- 08:37,RespRate,14
- 08:37,Urine,80
- ............ For 48 hours
Data Preprocessing

- Data like a time-series; but, not really....

Need to extract features from raw data:

- Mean
- Median
- First
- Last
- Min
- Max
- Number of values
- Trend
- Standard Deviation
- Sum (urine)
- ...

- Have to deal with missing data values
- Have to transform data
A Deep Learning Approach

• Use unsupervised learning to learn a model of the data that uses multiple layers of features (features of features)

• Use these multiple feature layers to initialize a deep neural network and then discriminatively fine-tune with labeled data
The Restricted Boltzmann Machine (Binary Visible and Hidden Units)

Model parameters: \( \theta = \{W, b, c\} \)

The probability of joint configuration \((v, h)\): 

\[
p(v, h; \theta) = \frac{e^{-E(v,h;\theta)}}{Z(\theta)}
\]

The energy function:

\[
E(v, h; \theta) = -h^T W v - b^T v - c^T h
\]

The partition function:

\[
Z(\theta) = \sum_{v^*} \sum_{h^*} e^{-E(v^*,h^*;\theta)}
\]
Preliminary Results: skewed data

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Died in-hospital (Positive)</th>
<th>Survivor (Negative)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Died in-hospital</td>
<td>True Positive (TP)</td>
<td>False Positive (FP)</td>
</tr>
<tr>
<td>(Positive)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survivor (Negative)</td>
<td>False Negative (FN)</td>
<td>True Negative (TN)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

accuracy = \frac{TP + TN}{TP + FP + TN + FN}, \quad sensitivity = \frac{TP}{TP + FN}, \quad specificity = \frac{TN}{TN + FP},

PPV = \frac{TP}{TP + FP}, \quad NPV = \frac{TN}{TN + FN}.

<table>
<thead>
<tr>
<th>Accuracy (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
<th>Score (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>86.23±0.14</td>
<td>50.29±0.22</td>
<td>92.01±0.21</td>
<td>50.29±0.50</td>
<td>92.01±0.00</td>
<td>50.29±0.22</td>
</tr>
</tbody>
</table>

Score = \text{Min}(PPV, Sensitivity)

Random guessing: score = 13.86\%
Other ongoing work

- **ECG Analysis**
  - Fetal ECG, T-wave alternans, Sport ECG, Cloud software
- **Scalp EEG analysis for cognitive deficits**
- **Intraoperative monitoring**
- **Critical care data analysis**
- **Spatiotemporal dynamics in the heart**
Summary

- Physiological monitoring faces various challenges
- Integration between signal processing and machine learning may improve performance on automated decision making