Artificial Intelligence and
Internet of Medical Things for Healthcare

Chenyang Lu
Fullgraf Professor
Department of Computer Science & Engineering
https://www.cse.wustl.edu/~lu/
AI for Healthcare

Data-driven tools for healthcare
- Predict clinical outcomes
- Predict treatment effect
- Discover risk factors associated with the outcomes
- Support clinical decisions to improve outcome

Extract knowledge from diverse data

Electronic Health Record (EHR)
- Collected in hospitals
- Complex and high-dimensional data

Internet of Things (IoT)
- Continuous monitoring outside hospitals
- Noisy and lossy data
Bridging the Gap between AI and Medicine

- **AI in medicine**
  - Apply off-the-shelf AI to healthcare
  - Limitations in face of hard clinical problems

- **AI in computer science**
  - New techniques are developed and published at an explosive rate
  - Largely ignored and rarely applied to medicine

- **AI + Medicine**
  - Tackle **hard** and **significant** medical problems with **advanced** AI
Wearables

- Commonly available: step, heart rate, sleep stages
- More sensing modalities
  - Oxygen saturation (SpO2)
  - Skin temperature
  - Breathing rate
  - Heart rate variability
  - ECG (AFib)
  - Stress
- 500+ million wearables sold in 2021

Unprecedented monitoring capability outside hospitals!
Internet of Medical Things

- **Wearables**: wristband, smartwatch, ring...
  - Long-term, non-obtrusive monitoring

- **Connectivity**: Bluetooth, WiFi, cellular
  - Real-time monitoring and intervention

- **Cloud**: computing and storage
  - Scalable to large population

- **Analytics**: machine learning
  - Predict outcomes and support intervention

Clinical decisions and intervention
Wearable Data → Clinical Outcomes

- **Reliable** prediction of clinical outcomes
  - Wearable data are fine-grained, noisy, lossy

- **Personalized** prediction of treatment response
  - Support clinical decisions on intervention

- **Scaling** predictive models
  - Small cohorts: robust models with limited data
  - Large population: exploit deep learning models
Predicting Pancreatic Surgery Outcome

- Pancreatic cancer has a 5-year survival rate less than 5%.

- Surgery is the only cure but is commonly followed by complications.

- Predict postoperative complications before surgery
  - Decision support: suitability for surgery
  - Intervention: pre-habilitation

Joint work with Chet Hammill (Surgery), Jingwen Zhang, Dingwen Li, Ruixuan Dai (CSE)
Goal: predict **postoperative** complications with **preoperative** data
- Wearable data collected by Fitbit (time series)
- Patient clinical characteristics from EHR (static data)

Prospective study: 61 patients undergoing pancreatic surgery
- 25 (40.98%) experienced complications

Machine learning approach
- Small cohort $\rightarrow$ avoid complex (deep) models
- Extract features from wearable time series $\rightarrow$ shallow machine learning models
- Integrate wearable data and patient clinical characteristics
Fitbit Data

- Wristband provides time series data: step count, heart rate, and sleep stage

**Fine-grained, lossy, noisy time-series**
Robust Feature Engineering Pipeline

Imputation for short missing segments

Heart rate Data

Missing values

Step Data

Sleep stages

Imputed values

Daily feature extraction

D: # of daily features
N: # of days

High-level feature extraction

Input:

predictive model

Clinical features

Raw Time Series → Daily Features

- **Statistical features**
  - max, sum, skewness, kurtosis, energy, entropy, inertia…

- **Semantic features**
  - Activity: daily steps, sedentary/active time, active/sedentary bout counts…
  - Sleep: awake counts, sleep efficiency, time of sleep, time before falling asleep…

- **Standardization**: account for missing data

![Diagram]

Daily sedentary time

Duration of wearing the device

\[
\text{Daily sedentary time} \quad \text{\# Data samples}
\]
Daily Features $\rightarrow$ High-level Features

- Singular Spectrum Analysis (SSA)
  - denoise time series of daily features.

Trend Extraction with Complete Data (30 data samples)
Singular Spectrum Analysis (SSA)  
- **denoise** time series of daily features.
Singular Spectrum Analysis (SSA)
- **denoise** time series of daily features.
- **robust** to missing daily features.

Trend Extraction with Complete Data (30 data samples)  Trend Extraction with Missing Components (20 data samples)
Daily Features \rightarrow High-level Features

- Singular Spectrum Analysis (SSA)
  - **denoise** time series of daily features.
  - **robust** to missing daily features.

Trend Extraction with Complete Data (30 data samples)

Trend Extraction with Missing Components (20 data samples)

- mean
- standard deviation
- slope
Robust Prediction of Surgical Complications

- Machine learning models outperform standard surgical risk scores.
  - x2 AUPRC
  - x3 sensitivity at the same specificity

<table>
<thead>
<tr>
<th>Data Source</th>
<th>AUROC</th>
<th>AUPRC</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random weighted classifier</td>
<td>0.5097 (0.0585)</td>
<td>0.4322 (0.0469)</td>
<td>0.1520 (0.0854)</td>
<td>0.8583 (0.0504)</td>
</tr>
<tr>
<td><strong>NSQIP with Clinical Characteristics</strong></td>
<td><strong>0.6114 (0.0000)</strong></td>
<td><strong>0.4075 (0.0000)</strong></td>
<td><strong>0.2800 (0.0000)</strong></td>
<td><strong>0.8571 (0.0000)</strong></td>
</tr>
<tr>
<td>ML with Clinical Characteristics Wearable Data</td>
<td>0.7632 (0.0085)</td>
<td>0.7374 (0.0206)</td>
<td>0.5800 (0.0699)</td>
<td>0.8583 (0.0083)</td>
</tr>
<tr>
<td><strong>Clinical Characteristics + Wearable Data</strong></td>
<td><strong>0.8802 (0.0050)</strong></td>
<td><strong>0.8871 (0.0087)</strong></td>
<td><strong>0.8320 (0.0160)</strong></td>
<td><strong>0.8583 (0.0083)</strong></td>
</tr>
</tbody>
</table>

- NSQIP: American College of Surgeons National Surgical Quality Improvement Program
- AUROC: Area Under the Receiver Operating Characteristic Curve
- AUPRC: Area Under the Precision-Recall Curve
Robust Prediction of Surgical Complications

- Machine learning models outperform traditional surgical risk scores.
- **Wearable data + clinical characteristics → best predictive performance**

<table>
<thead>
<tr>
<th>Data Source</th>
<th>AUROC</th>
<th>AUPRC</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random weighted classifier</td>
<td>0.5097 (0.0585)</td>
<td>0.4322 (0.0469)</td>
<td>0.1520 (0.0854)</td>
<td>0.8583 (0.0504)</td>
</tr>
<tr>
<td>NSQIP with Clinical Characteristics</td>
<td>0.6114 (0.0000)</td>
<td>0.4075 (0.0000)</td>
<td>0.2800 (0.0000)</td>
<td>0.8571 (0.0000)</td>
</tr>
<tr>
<td>ML with Clinical Characteristics</td>
<td>0.7632 (0.0085)</td>
<td>0.7374 (0.0206)</td>
<td>0.5800 (0.0699)</td>
<td>0.8583 (0.0083)</td>
</tr>
<tr>
<td>Wearable Data</td>
<td>0.7326 (0.0074)</td>
<td>0.7192 (0.0154)</td>
<td>0.5480 (0.0440)</td>
<td>0.8583 (0.0083)</td>
</tr>
<tr>
<td>Clinical Characteristics + Wearable Data</td>
<td><strong>0.8802 (0.0050)</strong></td>
<td><strong>0.8871 (0.0087)</strong></td>
<td><strong>0.8320 (0.0160)</strong></td>
<td><strong>0.8583 (0.0083)</strong></td>
</tr>
</tbody>
</table>

- NSQIP: American College of Surgeons National Surgical Quality Improvement Program
- AUROC: Area Under the Receiver Operating Characteristic Curve
- AUPRC: Area Under the Precision-Recall Curve
Personalized Prediction of Treatment Outcome

- Statistical analysis → **population-level** effectiveness of treatment

- **Personalized** prediction → **precision** medicine

- **Machine learning from RCT data**
  - Clinical data (baseline): age, anxiety score, PTSD…
  - Fitbit data (2 months): heart rate, sleep
  - Depression outcome (at 6 month)

---

**Randomized Controlled Trial of Depression Therapy**

**Group randomization**

- 106 patients

- Random split (71 : 35)

**6-month trial period**

- **Intervention**
  - Behavior therapy

- **Control**
  - No treatment

**Baseline clinical measurements**

- Continuous wearable data

**Joint work with**

Thomas Kannampallil (Informatics), Jun Ma (Medicine, UIC), Ruixuan Dai, Jingwen Zhang, (CSE)
Predicting Personalized Treatment Response

Group-specific models

- Intervention group
- Treatment-specific model
- Outcome with therapy

- Control group
- Risk model
- Outcome without therapy

Multi-task learning (MTL) with a unified model

- MTL model
- Outcome with therapy
- Outcome without therapy

- Small groups → overfitting
- Exploit similarity
- Capture differences

Detect Mental Disorders in the Community

- Mental disorders are prevalent
  - Over 50% of patients are not recognized or treated.

- Detect mental disorder (depression & anxiety) using
  - wearable data: daily features
  - static data: age, race, ethnicity, gender, education, smoke, alcohol

- All of Us: 8,996 participants with wearables (1,247 with mental disorders)

- WearNet: end-to-end deep model learning directly from raw daily features
  - No need for feature engineering

Joint work with Thomas Kannampallil (Informatics), Laura Jean Bierut (Psychiatry), Ruixuan Dai (CSE)
IoMT for Precision Medicine

- **Reliable** prediction of clinical outcomes
  - Wearable data → clinical outcomes

- **Personalized** prediction of treatment response
  - RCT data → personalized intervention

- **Scaling** predictive models
  - Small patient cohorts → robust feature engineering
  - Large population → deep learning models

Unleash the power of wearables for healthcare!
High Dimensionality of EHR Data

- High dimensionality
  - BJC perioperative data: 500+ variables
  - National COVID Cohort Collaborative: >10k measurement features
  - Complex correlations among variables

- Missingness: only a subset of variables are collected for each patient

- Lead to “brittle” models suffering performance deterioration and instability in the real world.
Clinical Variational Autoencoder (cVAE)

- Learn implicit, nonlinear relationship between input features
- High-dimensional input → low-dimensional representation
- Prediction guided → improve predictive performance
- Disentangled → retain interpretability

Electronic Health Records (EHR)

Encoder: compress into latent representation

Decoder: Reconstructs the original representation

Electronic Health Records (EHR)

Joint work with York Jiao, Thomas Kannampallil, Bradley Fritz, Christopher King, Joanna Abraham, Michael Avidan (Anesthesiology), Bing Xue (CSE)
Predicting Postoperative Delirium

- **BJC perioperative data:** 12,904 patients
- **Lower dimensionality:** $562 \rightarrow 10$
- **Better** predictive performance
- **No need for a separate predictor**

<table>
<thead>
<tr>
<th>Transformation Method ($d=10$)</th>
<th>Direct Prediction</th>
<th>LR</th>
<th>XGBoost</th>
<th>SVM</th>
<th>DNN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ROC AUC</td>
<td>Average Precision</td>
<td>ROC AUC</td>
<td>Average Precision</td>
<td>ROC AUC</td>
</tr>
<tr>
<td>PCA</td>
<td>-</td>
<td>-</td>
<td>.717 (.009)</td>
<td>.739 (.015)</td>
<td>.706 (.013)</td>
</tr>
<tr>
<td>ICA</td>
<td>-</td>
<td>-</td>
<td>.747 (.009)</td>
<td>.769 (.012)</td>
<td>.672 (.007)</td>
</tr>
<tr>
<td>GMM</td>
<td>.720 (.007)</td>
<td>.732 (.010)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AE</td>
<td>-</td>
<td>-</td>
<td>.643 (.007)</td>
<td>.657 (.015)</td>
<td>.641 (.009)</td>
</tr>
<tr>
<td>VAE</td>
<td>-</td>
<td>-</td>
<td>.647 (.007)</td>
<td>.667 (.009)</td>
<td>.650 (.010)</td>
</tr>
<tr>
<td>pi-VAE</td>
<td>-</td>
<td>-</td>
<td>.690 (.014)</td>
<td>.714 (.018)</td>
<td>.667 (.011)</td>
</tr>
<tr>
<td>cVAE-P</td>
<td>-</td>
<td>-</td>
<td>.656 (.010)</td>
<td>.672 (.007)</td>
<td>.655 (.011)</td>
</tr>
<tr>
<td>cVAE-D</td>
<td>-</td>
<td>-</td>
<td>.760 (.010)</td>
<td>.776 (.015)</td>
<td>.760 (.010)</td>
</tr>
<tr>
<td>cVAE</td>
<td><strong>.776 (.009)</strong></td>
<td><strong>.794 (.015)</strong></td>
<td><strong>.773 (.010)</strong></td>
<td><strong>.790 (.017)</strong></td>
<td><strong>.774 (.009)</strong></td>
</tr>
<tr>
<td>Raw Data ($d=562$)</td>
<td>-</td>
<td>-</td>
<td>.758 (.009)</td>
<td>.780 (.015)</td>
<td>.737 (.010)</td>
</tr>
</tbody>
</table>

Heterogeneity of EHR Data

- Early warning system: predict clinical deterioration of cancer patients

- Inpatient data from EHR
  - 128 static variables
  - 41 time-series variables

- Static and time series variables
  - make complementary contributions to prediction of clinical deterioration
  - have cross-modal correlation

Joint work with Patrick Lyons, Marin Kollef, Brian Gage (Medicine), Dingewen Li (CSE)
CrossNet

- Unified deep recurrent model for integrating static and time-series inputs
- Multi-modal fusion: integrating heterogeneous input data
- Cross-modal imputation: exploiting cross-modal correlation

CrossNet detects $10\times$ deterioration events than MEWS while generating the same number of false alarms.

<table>
<thead>
<tr>
<th>Model</th>
<th>Alarm rate control</th>
<th>False alarm control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sensitivity</td>
<td>Specificity</td>
</tr>
<tr>
<td>MEWS</td>
<td>0.3358(0.0115)</td>
<td>0.8257(0.0142)</td>
</tr>
<tr>
<td>C. BRITS</td>
<td>0.3899(0.0134)</td>
<td>0.9394(0.0097)</td>
</tr>
<tr>
<td>S. BRITS</td>
<td>0.3891(0.0122)</td>
<td>0.9396(0.0105)</td>
</tr>
<tr>
<td>CrossNet</td>
<td><strong>0.4218(0.0130)</strong></td>
<td><strong>0.9486(0.0093)</strong></td>
</tr>
</tbody>
</table>

MEWS: Modified Early Warning Scores

AI and IoT for Medicine (AIM)

- Healthcare should benefit from the vast amount of data available.
- Advanced AI allows us to extract knowledge from complex data.
- Collaboration between medicine and AI/IoT researchers is essential.

Electronic Health Record (EHR)
- Collected in hospitals
- Complex and high-dimensional data

Internet of Things (IoT)
- Continuous monitoring outside hospitals
- Noisy and lossy data
Proposal: The AIM Institute

- Solve significant healthcare problems with advanced AI and IoT
  - Hard medical problems call for advanced AI and IoT
  - Advanced AI and IoT bring significant value to medicine

- Hub for interdisciplinary collaboration on AIM
  - Connect AI/IoT faculty and medical researchers

- Incubator for interdisciplinary research of AIM
  - Fund and develop pilot studies combining medicine and advanced AI/IoT

- Establish WashU as a world leader in AIM