Real-Time Machine Learning Pipeline: A Clinical Early Warning Score (EWS) Use-Case

Session 1033, February 12, 2019
Prem Timsina, Senior Data Engineer, Mount Sinai Health Systems
Arash Kia, Senior Data Scientist, Mount Sinai Health Systems

DISCLAIMER: The views and opinions expressed in this presentation are those of the author and do not necessarily represent official policy or position of HIMSS.
Conflict of Interest

Prem Timsina, Sc.D.

Arash Kia, M.D., M.Sc.

There is no real or apparent conflicts of interest to report
Clinical Data Science Team

- **Arash Kia**: M.D., M.Sc. - *Senior Data Scientist*
- **Prem Timsina**: D.Sc. - *Senior Data Engineer*
- **Fu-Yuan Cheng**: M.Sc. - *Junior Data Scientist*
- **Prathamesh Parchure**: M.Sc. - *Data Science Analyst*
- **William Otto**: B.Sc. - *System Admin, DevOps Engineer*
- **Robert Freeman**: M.Sc., RN, NE, B.Sc. - *Senior Director of Operations*
- **Matthew A. Levin**: M.D. - *Clinical Lead, Associate Professor, Department of Anesthesiology, Division of Cardiothoracic Anesthesia*
Agenda

- Problem Definition
- Proposed Solution
- Use Cases
- Challenges
- Future Plan and Vision
Learning Objectives

• Explain the rationale behind the need for a real-time machine learning (ML) pipeline in personalized medicine

• Explain the architecture of the healthcare ML pipeline

• Identify the challenges and the lessons learned from developing and deploying process

• Discuss how the ML pipeline can be generalized into various health care settings and cases of use
Problem Definitions
## Conventional AI Applications in Healthcare

<table>
<thead>
<tr>
<th>APPLICATION</th>
<th>POTENTIAL ANNUAL VALUE BY 2026</th>
<th>KEY DRIVERS FOR ADOPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robot-assisted surgery</td>
<td>$40B</td>
<td>Technological advances in robotic solutions for more types of surgery</td>
</tr>
<tr>
<td>Virtual nursing assistants</td>
<td>20</td>
<td>Increasing pressure caused by medical labor shortage</td>
</tr>
<tr>
<td>Administrative workflow</td>
<td>18</td>
<td>Easier integration with existing technology infrastructure</td>
</tr>
<tr>
<td>Fraud detection</td>
<td>17</td>
<td>Need to address increasingly complex service and payment fraud attempts</td>
</tr>
<tr>
<td>Dosage error reduction</td>
<td>16</td>
<td>Prevalence of medical errors, which leads to tangible penalties</td>
</tr>
<tr>
<td>Connected machines</td>
<td>14</td>
<td>Proliferation of connected machines/devices</td>
</tr>
<tr>
<td>Clinical trial participation</td>
<td>13</td>
<td>Patent cliff; plethora of data; outcomes-driven approach</td>
</tr>
<tr>
<td>Preliminary diagnosis</td>
<td>5</td>
<td>Interoperability/data architecture to enhance accuracy</td>
</tr>
<tr>
<td>Automated image diagnosis</td>
<td>3</td>
<td>Storage capacity; greater trust in AI technology</td>
</tr>
<tr>
<td>Cybersecurity</td>
<td>2</td>
<td>Increase in breaches; pressure to protect health data</td>
</tr>
</tbody>
</table>

**Source:** Accenture

Optimization Opportunity

- **Clinical Decision**
  - A process to organized clinical knowledge linked to patient information to improve patient's health and healthcare delivery

- **Strategies for effective use of knowledge**
  - Best knowledge available in terms of adoptability and effectiveness
  - Continuous improvement of knowledge
  - Data-driven CDS
Clinical Decision Support Timeline

1970s

Computer-based Information Systems
  - Transaction processing
  - Operational Research
    - Optimization and Simulation models
    - Behavioral Decision Theory

1980s

Personal Decision Support System
  - Data base Theory
  - Data Warehousing

1990s

Rule Engines
  - Artificial Intelligence
    - AI Based Early Warning System

Present
Current Challenges Cont....

- **Data**
  - Quality and variability of clinical data
  - Interoperability of clinical data
  - Data sharing
  - Standardization of data representation both terminology and ontology level

- **Knowledge**
  - Lack of standardized terminology (lab, medications,...)
  - Computer interpretable representations of guidelines/BPs
  - Environment to trial them at the point of care
  - Translate into routine services
Current Challenges

- **Architecture**
  - Integrated into EHR
  - Context sharing and management

- **Implementation and Integration**
  - Extraordinary variability in clinical practice patterns and workflow
  - Inopportune time in the workflow
  - Too late
  - Difficulty in developing, maintaining and integrating the clinical logic
Current Scenario

Develop Machine Learning Model on ad-hoc Basis

- Development and operational team are in silos
- Operationalization of developed ML model is difficult

Utilize the pre-built Data-driven CDS from EHR System

- CDS Tools are not provider customized
- CDS Tools do not capture local clinical process
- Custom data captured in Health System may not be utilized
# Optimum Scenario

- Easy To Develop and Deploy Clinical Data Science (CDS) App
- Scalable
- *Heterogeneous Data Ingestion, Storage and Consumption*
- Secure
- Automation
- Modular
Proposed Solutions
What is Data Science Pipeline?

Mount Sinai Data Science (DS) Pipeline is end-to-end infrastructure allowing:

• Consumption of a variety of historical and streaming data,
• Feature engineering and transformation,
• Building an CDS App, and
• finally operationalizing the CDS App
Computational Infrastructure

**In-house**
- Mirth Connect
- MongoDB Cluster
- Apache Kafka
- Apache Spark Cluster
- VMs
- MySQL DB

**Cloud**
- Microsoft Azure
## Data Science Toolkit

<table>
<thead>
<tr>
<th>Module</th>
<th>Packages</th>
<th>Descriptions</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>ETL Module</td>
<td>IO Utility</td>
<td>Functions related to read/write from/to any source/destination</td>
<td>writeToHive()</td>
</tr>
<tr>
<td></td>
<td>Normalization</td>
<td>Data normalizations functions</td>
<td>mapGenericLabNameToLONIC()</td>
</tr>
<tr>
<td></td>
<td>Data Service</td>
<td>Receive various type of data</td>
<td>getVitals(), getPatientInfo()</td>
</tr>
<tr>
<td></td>
<td>Transformation</td>
<td>Performs various data transformation</td>
<td>calculateAge(), getLabTimeSeries()</td>
</tr>
<tr>
<td>ML Module</td>
<td>Supervised Learning</td>
<td>Wrapper over the different machine learning algorithms</td>
<td>getCrossValidatedSVMMModel(), getCrossValidatedLSTMMModel()</td>
</tr>
<tr>
<td></td>
<td>Encoding</td>
<td>Functions to transform categorical data into dummy variables</td>
<td>encodeCategoricalColumn()</td>
</tr>
<tr>
<td></td>
<td>Null Value Imputation</td>
<td>Methods to impute numerical null value</td>
<td>imputeNullValueByMean(), imputeNullValueByMode()</td>
</tr>
<tr>
<td></td>
<td>Feature Engineering</td>
<td>Methods to create new features</td>
<td>calculateBMI(), calculateLastVisitLengthOFStay()</td>
</tr>
<tr>
<td></td>
<td>Feature Selection</td>
<td>Methods for Multiple Features Selection Techniques</td>
<td>getTopFeaturesByRecursiveFeatureSelection()</td>
</tr>
<tr>
<td></td>
<td>Sampling</td>
<td>Different Sampling Techniques</td>
<td>conductSMOTESampling(), conductOverSampling()</td>
</tr>
</tbody>
</table>
System Architecture

[Diagram of system architecture with various components and connections, including MSH Centralized Interface, Data Production Environment, Mobile App, Custom UI, MSH, EPIC, MEDS System, Lab System, Home Health Data Stream, ADT HL7, EKG HL7, Lab HL7, Orders HL7, Meds HL7, Flowsheet HL7, and Machine Learning Engine.]
ML Pipeline
Operationalizing the Solution

Build DS App that is Productionalizable

• Use variable that are available at the prediction time
• Model Complexity Vs Operational Cost
• Proper exceptional handling

Deployment Infrastructure

• Continuous Deployment Pipeline
• Dockerizing of a DS App
• Automated Monitoring, Debugging and Alerts of DS App
• Continuous Monitoring of Data Quality
• Model Performance Dashboard
• 24/7 Support and Debugging
Continuous Learning Engine

Create DS App that is always adapting to change in data and practice
Practice Flow

- Ingesting stream of data from EHR, LAB, MED, ADT platforms -> outdated data
- Real-time computation → too late
- Real-time integration into point of care → too late
- Using machine learning approaches to maintain clinical logic consistent and patient specific
- Applying machine learning to optimize the golden time and minimize the failure to rescue → too late
- Streaming Data Service to extract and normalize data → real-time data collection/ data quality check/ data standardization for computation
- Product development cycle starts from frontline clinicians → change in the users’ mental mode / minimize irrelevancy
Incorporate ML Process into Current Care Practice
Monitoring and Debugging
ML Pipeline Monitoring: Apps Monitoring — Apache Spark UI

Streaming Statistics

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0-0.1 sec</td>
<td>18 records</td>
<td>0 ms</td>
<td>19 s</td>
<td>19 s</td>
<td>15/15</td>
</tr>
<tr>
<td>0.1-0.2 sec</td>
<td>22 records</td>
<td>1 ms</td>
<td>19 s</td>
<td>19 s</td>
<td>15/15</td>
</tr>
<tr>
<td>0.2-0.3 sec</td>
<td>34 records</td>
<td>1 ms</td>
<td>22 s</td>
<td>22 s</td>
<td>15/15</td>
</tr>
<tr>
<td>0.3-0.4 sec</td>
<td>7 records</td>
<td>0 ms</td>
<td>18 s</td>
<td>18 s</td>
<td>15/15</td>
</tr>
<tr>
<td>0.4-0.5 sec</td>
<td>16 records</td>
<td>0 ms</td>
<td>21 s</td>
<td>21 s</td>
<td>15/15</td>
</tr>
<tr>
<td>0.5-0.6 sec</td>
<td>24 records</td>
<td>0 ms</td>
<td>20 s</td>
<td>20 s</td>
<td>15/15</td>
</tr>
</tbody>
</table>

Completed Batches (last 1000 out of 3653)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2018/12/06 17:16:30</td>
<td>18 records</td>
<td>0 ms</td>
<td>19 s</td>
<td>19 s</td>
<td>15/15</td>
<td>15/15</td>
</tr>
<tr>
<td>2018/12/06 17:19:30</td>
<td>22 records</td>
<td>1 ms</td>
<td>19 s</td>
<td>19 s</td>
<td>15/15</td>
<td>15/15</td>
</tr>
<tr>
<td>2018/12/06 17:15:30</td>
<td>34 records</td>
<td>1 ms</td>
<td>22 s</td>
<td>22 s</td>
<td>15/15</td>
<td>15/15</td>
</tr>
<tr>
<td>2018/12/06 17:13:30</td>
<td>7 records</td>
<td>0 ms</td>
<td>18 s</td>
<td>18 s</td>
<td>15/15</td>
<td>15/15</td>
</tr>
<tr>
<td>2018/12/06 17:14:30</td>
<td>16 records</td>
<td>0 ms</td>
<td>21 s</td>
<td>21 s</td>
<td>15/15</td>
<td>15/15</td>
</tr>
<tr>
<td>2018/12/06 17:14:00</td>
<td>24 records</td>
<td>0 ms</td>
<td>20 s</td>
<td>20 s</td>
<td>15/15</td>
<td>15/15</td>
</tr>
</tbody>
</table>
ML Pipeline Monitoring: Apps Monitoring — Custom UI

- ADT Channel
- ECG Channel
- Lab Channel
- Flowsheet Channel
- Output Prediction Channel
DS Infrastructure Monitoring Dashboard

- Nagios
- Mongo Ops Manager
ML Model Performance Evaluation Dashboard—Custom UI

Mainnutrition Model Utilization and Performance Report

**Background**
- Historical rate of documented malnutrition one year ago was 3.8%
- After hiring additional RDs, rate jumped to 7.6%
- UHC top 10% average is 8.2%

**Purpose**
- Evaluate the Registered Dietitians' Workload for all the Units
- Evaluate the utilization of the Prediction Model
- Evaluate the model performance

Dashboard generated at 12/10/2018 05:00
USE CASE I: Early Warning System for Patient Deterioration

Analysis of Main Hospital
Classical Approach

Limitations:

- BPA Alert--No Golden Time for Intervention
- Low Performance
  - Low Accuracy (~ 70%)
    - Median of the MEWS (~ 1). It suggests that MEWS fails in appraising the impact of the presence of multiple co-morbidities
  - Low Sensitivity (~ 60%)

## Machine Learning Model To Predict Patient’s Deterioration After 16 hours

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>FPR</th>
<th>FNR</th>
<th>Accuracy</th>
<th>F₁ Score</th>
<th>AUC ROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Forest</td>
<td>81.6</td>
<td>75.5</td>
<td>24.5</td>
<td>18.4</td>
<td>75.7</td>
<td>0.189</td>
<td>0.85</td>
</tr>
<tr>
<td>SVM</td>
<td>77.3</td>
<td>73.7</td>
<td>26.3</td>
<td>22.7</td>
<td>73.9</td>
<td>0.17</td>
<td>0.83</td>
</tr>
<tr>
<td>Logistic Regression</td>
<td>50.2</td>
<td>83.2</td>
<td>16.8</td>
<td>49.8</td>
<td>82.0</td>
<td>0.162</td>
<td>0.75</td>
</tr>
<tr>
<td>Classical MEWS</td>
<td>64.5</td>
<td>66.6</td>
<td>33.4</td>
<td>33.5</td>
<td>66.5</td>
<td>0.11</td>
<td>0.67</td>
</tr>
</tbody>
</table>

![ROC Curve Graph](image)
Comparing Classical MEWS with MEWS-Plus

Comparing Sensitivity and FPR

Comparing Specificity and FNR
Use Case II: Early Detection of the Risk of Severe Malnutrition in the Inpatient Setting

Analysis of Main Hospital
Prediction Performance: Selected Model

<table>
<thead>
<tr>
<th>Model</th>
<th>Test Size</th>
<th>Malnutrition Rate (%)</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
<th>False Neg. Rate (%)</th>
<th>False Pos. Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF Classifier</td>
<td>2017</td>
<td>12.6</td>
<td>77.2</td>
<td>73.2</td>
<td>73.7</td>
<td>22.8</td>
<td>26.8</td>
</tr>
</tbody>
</table>

Model Parameters

- # of Trees: 500
- Max # of Bins: 32
- Max Depth of Trees: 10
- # of Folds in CV: 10
- Feature Subset Strategy: 1/3 features (394) = 131

ROC Curve: ROC Score = 0.82
Prediction Performance: Selected Model vs. MUST

Comparing Sensitivity and False Positive Rate(%)

- Classical MUST: Sensitivity 22.9, False Positive Rate 3.3
- mustPlus: Sensitivity 77.2, False Positive Rate 26.8

Comparing Specificity and False Negative Rate(%)

- Classical MUST: Specificity 96.7, False Negative Rate 77.1
- mustPlus: Specificity 73.2, False Negative Rate 22.8
Challenges and Future Vision
Policy and Governance

- Financial incentives go frequently against data sharing for example in fee for service business model
- Lack of Data governance initiatives in Health systems
- Paradigm shift in guidelines and best practices from empirical intervention to proactive preventive intervention in acute care facilities
- Reinforce clinical documentation templates and flow sheets
Technical Challenges

- **Multiple data sources and the lack of a unified data dictionary across the EHR systems:** The data science engine ingests data from multiple data producers each having their own data representation as well as custom extract, transform, and load logic.

- **Data Quality:** The majority of data were highly sparse and qualitative in nature, which required considerable effort in terms of data preparation.

- **Lack of Standardized Blue Print:** Engineering/infrastructure teams are yet to determine the optimal process
Future Plan and Vision

• Incorporate Outpatient Setting Data
• Incorporate Imaging, and Genomics Data
• Create Application Programming Interface (API) over Data Science Toolkit, and publish API for usage within the Health Systems
• Move pipeline to cloud
• Provide DS App as a service for other providers nationally and internationally
Acknowledgement

• All of this is possible thanks to a strong dedication to IT investments from Mount Sinai Health System President & CEO Dr. Ken Davis & CIO Kumar Chatani

• Project sponsored by the Mount Sinai Hospital President Dr. David Reich
Questions

Contacts

Prem Timsina
Email: Prem.timsina@mssm.edu
LinkedIn: https://www.linkedin.com/in/timsinaprem/

Arash Kia
Email: arash.kia@mssm.edu
LinkedIn: https://www.linkedin.com/in/KiaArash/

Please complete online session evaluation