Deciphering Clinical Narratives – Augmented Intelligence for Decision Making in Health Care Sector

Dr. Lipika Dey
TCS Research, India

19th September, 2023

Keynote Talk
Outline of the talk

- Clinical texts and Decision-Making problems in Health Care Sector
- Predicting Length of stay in ICU using first day’s nursing notes
  - Results and Learnings
- Forming Patient Cohorts
  - Deeper dive into risk factors for each group
- Way forward – some plans
Clinical narratives - Main form of communication within health care

- **Clinical Data**
  - Electronic Medical Records (EMR)/Electronic Health Records (EHR)
  - Physician and Care-giver notes - patient history, assessments and treatments
  - Clinical trials management – trial description, monitoring trial progress

- **Social Media (tweets, Facebook comments, message boards, etc.)**
  - personal accounts of patients – signals for mental health – adverse effects of drugs
  - Health care system feedback

- **Medical Literature**
  - News feeds, Medical journals

- **Insurance Providers (claims from private and government payers)**
  - Underwriter notes

---

Was estimated to be 25,000 petabytes by the end of 2020 – COVID 19 enhanced it by many orders
Most rapidly rising data repository

2018-2025 Data – Compound Annual Growth Rate (CAGR)

- Healthcare: 36%
- Manufacturing: 30%
- Financial Service: 26%
- Media and Entertainment: 25%
- Global DataspHERE: 27%

Challenges of working with Clinical Data

- Chances of Privacy violation
- Not available in large quantities for research
- Makes training and evaluation of Machine Learning techniques difficult
- Non-standard terminologies
- Textual narratives are more so


Decision Support Systems in Health Care Sector

- Targeted Clinical Knowledge
- Hospital Management Data

Improved Healthcare Delivery

- Hospital Logistics Management
- Expectation management for patients and their family
- Personalized Patient Care
Using Clinical Notes

ICU LENGTH OF STAY PREDICTION
About MIMIC-III v1.4 Dataset

- MIMIC-III v1.4 Database contains details of 58976 admission records of 46,520 patients who stayed in critical care units of the Beth Israel Deaconess Medical Center (BIDMC) between 2001 and 2012 - developed by the Laboratory for Computational Physiology, MIT.

- Has pre-existing Institutional Review Board (IRB) approval

- Adheres to stringent anonymization protocols - meticulously safeguards patient privacy

- Ensures heightened privacy protection by obfuscating precise dates and times of events

- Researchers can access the data after successfully completing the training course “Data or Specimens Only Research” provided by the Collaborative Institutional Training Initiative (CITI)
ICU Length of Stay (LOS) Prediction – *why is this important?*

- According to the World Health Organization (WHO) patient Length of Stay (LOS) in hospitals is an important performance measurement and monitoring indicator.

- Intensive Care Units (ICU) are premium facilities in a hospital – comprising hospital resources, manpower, and equipment.
  - Accurate prediction of patient LOS aids healthcare specialists to take medical decisions and allocate medical team and resources appropriately.

- Better logistics planning – ensures **better resource usage** for critically ill patients.

- Patient and insurance companies may use this prediction to manage their budget.

*The Earlier the better*
## Earlier Work – using Patient Health Parameters

<table>
<thead>
<tr>
<th>Comparing with SOTA</th>
<th>Dataset</th>
<th>Feature used</th>
<th>Method</th>
<th>Best Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alghatani et al., 2021</td>
<td>44,000 ICU stays from MIMIC</td>
<td>patient’s vital signs like, heart rate, BP, temp., resp. etc</td>
<td>Random Forest</td>
<td>65% accuracy</td>
</tr>
<tr>
<td>Su et al., 2021</td>
<td>2224 Sepsis patients PICMISD</td>
<td>Age, P(v-a)CO /C(a-v)O, SO, wbc etc</td>
<td>XG-Boost model</td>
<td>F1: 0.69, AUC-ROC:0.76</td>
</tr>
<tr>
<td>Rocheteau, Liò, et al., 2020</td>
<td>eICU critical care dataset</td>
<td>medical features, Gender, Age, Ethnicity, etc.</td>
<td>Temporal convolution</td>
<td>Kappa score = 0.58</td>
</tr>
<tr>
<td>Harutyunyan et al., 2019</td>
<td>42276 ICU stays of 33798 unique patients from mimic database</td>
<td>17 clinical variables like, Capillary refill rate, Diastolic blood pressure etc. from first 24 hours of admission.</td>
<td>LSTM</td>
<td>AUC-ROC : 0.84</td>
</tr>
<tr>
<td>van Aken et al., 2021</td>
<td>38013 admission notes from MIMIC III</td>
<td>Created admission notes from discharge summaries</td>
<td>Pretrained CORE + BioBERT</td>
<td>AUC-ROC : 0.72%</td>
</tr>
</tbody>
</table>
The pt is a complicated 62yo man who was transferred from ajh last evening with mrsa bacteremia and pnx. Pt arrived via EMS,intubated, sedated on Propofol 15mcg/kg-min, on Dopa gtt at 5mcg/kg-min. Transferred to Big Boy bed with 6-person assist, MICU-A monitor, and MICU-A IV pumps. **Dopamine** titrated up to max of 6.5mcg/kg-min with SBP 90's-80's, Propofol weaned from 15mcg/kg-min -> 12. Levo-phed started at 0.05mcg/kg-min and dopa weaned to 4.0mcg at time of shift report. Propofol d/c'd, and fentanyl and midazolam started at 25mcg/kg-min and 0.5mg/hr, respectively......Pt turned upon admission; SBP by a-line dropped to 70's with + wave-form, and SpO2 dropped to 80's......**Skin breakdown noted** over back of neck, sloughed skin with serosanguinous drng OTA on arrival......NaHCO3 3 amps given after 2nd ABG when acidosis was worsening with respiratory intervention....Daughter in to see pt. Gravity of pt status discussed with daughter.....Patient remains on mechanical ventilation;switched to PCV due to high Paw.PIP improved as well as Pa02, but patient still has **significant metabolic acidosis. Renal insufficiency**,patient may need to be dialysised.BS diminished,suctioned for small amount of clear thick tenacious type of secretion. the micu team is concerned that the pt may have mrsa endocarditis and the plan is to obtain a tte later today. he continued to have a difficult noc with persistent fevers, hypoxia, acidosis, copd/emphysema, asbesteosis, endometriosis with hematuria as well as requiring an increase in pressor support.....

Nursing notes are very detailed, time-stamped account of a patient’s stay in a hospital

- Observations made by Doctors about symptoms and diseases
- Prior medical history including drug allergies etc.
- Critical examinations suggested and / or reports
- Treatment plans
  - Medicines prescribed
  - Procedures suggested
- General conditions – appetite, mobility, pain
Predicting ICU stay on the day of admission to ICU

- Most of the earlier work used vital parameters to predict LOS

- Van Aken et al., 2021 used Discharge Summaries - contains entire history of patient during admission

- **Our work**
  - *Use the first day's Nursing Note prepared after patient’s admission to ICU along with vital parameters to predict the LOS*

- Though called LOS –
  - usually prediction is about LONG / SHORT stay
    - SHORT <= MEDIAN, LONG > MEDIAN
  - Exact stay is decided based on multiple other factors like age, availability of facility, patient willingness etc.
scispaCy is a Python package containing spaCy models for processing biomedical, scientific or clinical text.

Patient is a 83 yo female, recently admitted for treatment of severe multilobar pnx.......CXR revealed worsening multifocal pnx and free air under diaphragm.....Currently on 100%NRB with 4l NC. O2 sat 94-97%, rr 14-18, Bp 80-110/40-50, HR 80-105 Hypoactive BS, abdomen soft, distended and tender to touch. No stool thus far....Alert and oriented, pleasant and cooperative...No sob, No abdominal pain...no bs heard....triple antibiotic coverage..also to start amphotericin tonoc....had some oozing from neck line..redressed with gelfoam..no further bleeding noted..no other sites...Family is abroad...back on Wednesday per patient......Plan: Monitor resp status, monitor hemodynamic status, monitor temp, wbc's, follow cultures, continue amphotericin.
Severity of Illness Score

- **Acute Physiology and Chronic Health Evaluation (APACHE-II) score** – Uses 12 physiological variables that include mean arterial pressure, temperature, heart rate, respiratory rate, oxygenation, GCS, pH, sodium, potassium, creatinine, hematocrit, and WBC level in the blood.

- **Simplified Acute Physiology Score (SAPS-II)** - Uses logistic regression techniques to predict the SOI using 12 physiological variables, age, type of admission such as surgical or medical, and three variables related to acquired immuno-deficiency syndrome, metastatic cancer, and hematologic malignancy.

- **Sepsis-related Organ Failure Assessment (SOFA) score** - Used to measure a person’s organ function or rate of failure during the stay in an ICU. This score is based on six different values coming from the assessment of the respiratory, cardiovascular, hepatic, coagulation, renal, and neurological systems.

- **Oxford Acute Severity of Illness Score (OASIS)** - computed from 10 variables: elective surgery, age, pre-ICU length of stay, and seven physiological measurements.

For our model, all four scores are calculated using data from first 24 h of ICU stay only – using libraries.
Multimodal DNN for Predicting ICU LOS

Clinical BioBERT - developed by further training the BioBERT (base version; 110 million parameters with 12 layers, 768 hidden units, and 12 attention heads) using clinical text from the MIMIC-III corpus.

BlueBERT - language model trained on Biomedical and Clinical texts – PUBMED abstracts and MIMIC III clinical notes
LIME Framework to add explainability

Local Interpretable Model-agnostic Explanations

Attempts to understand the model by perturbing the input of data samples and observe how the predictions change

- Weighted average from multi-head attention models

Output

Long ICU Stays - “HR dropping”, “requiring mask ventilation for resp. failure”, “couldn’t breathe”

Short ICU Stays - “good effect from Ativan”, “comfortable breathing”, “hemodynamically stable”
Results obtained from 22789 admissions which had Nursing Notes for first day

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy</th>
<th>F1 score Class ‘Short’</th>
<th>F1 score Class ‘Long’</th>
<th>AUC-ROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>BlueBERT+LSTM+Attn+TF-IDF+SOI</td>
<td>0.797</td>
<td>0.810</td>
<td>0.790</td>
<td>0.872</td>
</tr>
<tr>
<td>BlueBERT+LSTM+Attn+TF-IDF</td>
<td>0.792</td>
<td>0.800</td>
<td>0.790</td>
<td>0.873</td>
</tr>
<tr>
<td>BlueBERT+CNN+TF-IDF</td>
<td>0.789</td>
<td>0.800</td>
<td>0.780</td>
<td>0.872</td>
</tr>
<tr>
<td>BlueBERT</td>
<td>0.776</td>
<td>0.790</td>
<td>0.760</td>
<td>0.833</td>
</tr>
<tr>
<td>Clinical BioBERT+LSTM+TF-IDF+SOI</td>
<td>0.780</td>
<td>0.790</td>
<td>0.770</td>
<td>0.871</td>
</tr>
<tr>
<td>Clinical BioBERT+LSTM+TF-IDF</td>
<td>0.778</td>
<td>0.780</td>
<td>0.770</td>
<td>0.857</td>
</tr>
<tr>
<td>Clinical BioBERT+CNN+TF-IDF</td>
<td>0.776</td>
<td>0.780</td>
<td>0.770</td>
<td>0.864</td>
</tr>
<tr>
<td>Clinical BioBERT</td>
<td>0.741</td>
<td>0.750</td>
<td>0.730</td>
<td>0.818</td>
</tr>
</tbody>
</table>
Studying the Discharge Notes to gain insights about wrong predictions

- Patients developed complications mid-way

- Complications could be about other organs and not related to the disease on admission

- Initial Nursing Notes not always indicative of type of procedures patients underwent later
  - *Admitted for Pneumonia – underwent Bypass Surgeries*
  - *Delay in scheduling unplanned procedure*

- Four most common procedures
  - Bypass Surgery
  - Stenting
  - Tracheotomcy
  - Cholecystectomy

Reformulated Problem – Predict LOS and Major Procedure requirements

- Helps in better planning of resources
- Better expectation management for patient’s family
Multitask, Multimodal Architecture for predicting LOS and Major Procedures
### Results

#### ICU LOS Prediction

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy</th>
<th>F1 score ‘Short’</th>
<th>F1 score ‘Long’</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multitask BILSTM-blueBERT with tf_idf and SOI</td>
<td>0.84</td>
<td>0.86</td>
<td>0.82</td>
<td>0.89</td>
</tr>
<tr>
<td>Multitask BILSTM-blueBERT with tf_idf</td>
<td>0.83</td>
<td>0.84</td>
<td>0.80</td>
<td>0.87</td>
</tr>
<tr>
<td>Multitask BILSTM-blueBERT</td>
<td>0.79</td>
<td>0.79</td>
<td>0.78</td>
<td>0.85</td>
</tr>
<tr>
<td>BILSTM-blueBERT with tf_idf and SOI</td>
<td>0.79</td>
<td>0.81</td>
<td>0.79</td>
<td>0.86</td>
</tr>
<tr>
<td>BILSTM-blueBERT with tf_idf</td>
<td>0.79</td>
<td>0.80</td>
<td>0.78</td>
<td>0.84</td>
</tr>
<tr>
<td>BILSTM-blueBERT</td>
<td>0.77</td>
<td>0.77</td>
<td>0.76</td>
<td>0.83</td>
</tr>
</tbody>
</table>

#### Procedure Prediction

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy</th>
<th>F1 score “Dypass”</th>
<th>F1 score “Stent”</th>
<th>F1 score “Tracheotomy”</th>
<th>F1 score “Cholecystectomy”</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multitask BILSTM-blueBERT with tf_idf and SOI</td>
<td>0.82</td>
<td>0.89</td>
<td>0.83</td>
<td>0.55</td>
<td>0.54</td>
<td>0.86</td>
</tr>
<tr>
<td>Multitask BILSTM-blueBERT with tf_idf</td>
<td>0.81</td>
<td>0.86</td>
<td>0.81</td>
<td>0.53</td>
<td>0.51</td>
<td>0.85</td>
</tr>
<tr>
<td>Multitask BILSTM-blueBERT</td>
<td>0.60</td>
<td>0.85</td>
<td>0.78</td>
<td>0.51</td>
<td>0.48</td>
<td>0.83</td>
</tr>
<tr>
<td>BILSTM-blueBERT with tf_idf and SOI</td>
<td>0.60</td>
<td>0.65</td>
<td>0.79</td>
<td>0.49</td>
<td>0.49</td>
<td>0.83</td>
</tr>
<tr>
<td>BILSTM-blueBERT with tf_idf</td>
<td>0.78</td>
<td>0.61</td>
<td>0.79</td>
<td>0.48</td>
<td>0.46</td>
<td>0.83</td>
</tr>
<tr>
<td>BILSTM-blueBERT</td>
<td>0.77</td>
<td>0.81</td>
<td>0.79</td>
<td>0.47</td>
<td>0.46</td>
<td>0.81</td>
</tr>
</tbody>
</table>

AUC: area under the receiver operating characteristic; SOI: severity of illness.

We report accuracies, AUC scores of the model, and F1 scores of all four classes. Bold values indicate the best performance of our experiments.
# Utility of Prediction Model

<table>
<thead>
<tr>
<th>Procedure Name</th>
<th>Recall</th>
<th>Precision</th>
<th>Gain = (TP-Mentioned_in_firstnote)/total_number)%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bypass Surgery</td>
<td>0.98</td>
<td>0.82</td>
<td>47.97</td>
</tr>
<tr>
<td>Stenting</td>
<td>0.75</td>
<td>0.92</td>
<td>31.92</td>
</tr>
<tr>
<td>Tracheotomy</td>
<td>0.71</td>
<td>0.44</td>
<td>68.15</td>
</tr>
<tr>
<td>Cholecystectomy</td>
<td>0.41</td>
<td>0.78</td>
<td>30.93</td>
</tr>
</tbody>
</table>

Model was able to predict many procedures on first day itself – *doctors did it much later*

*Symptom correlations to procedures were identified by model*
Digging Deeper

- We observed that - Though “long” / “short” predictions have improved and are good enough for hospital management – *actual length prediction of stay is not so good*

- Key observations
  - Two patients with very similar symptoms for major diseases have very different outcomes in terms of Length of Stay
  - Prescriptions / Treatments varied for patients with similar symptoms
  - Actual outcomes may be dependent on comorbidities present

- Individual Outcomes for patients - Risk Assessment

- Patient cohorts - can provide better understanding of diseases when studied in smaller groups
Risk Stratification and Patient Cohorts

- Process to identify individuals who are at different levels of risk for a particular disease
  - in terms of disease progression, complications, or adverse events
- Risk scores are calculated by assigning weights to patient's clinical parameters
- Cohort studies are a way to understand the factors

**What is a Patient cohort**

It is a term used in medical research to define *groupings of individuals with common traits, such as social and health factors*

Patient cohorts are integral to researching and developing effective medical interventions

Observational medical studies begin with selecting a patient cohort
Identifying Patient Cohorts using Machine Learning

- Patient Cohorts are selected by clinical researchers
- Studies are rarely reproducible
- How do we obtain cohorts and understand risk factors automatically from open research data?

- Clustering!
OBJECTIVE:

BETTER RISK ASSESSMENT FOR EACH COHORT

BETTER PREDICTION OF RESOURCES

Identifying cohorts within Patient Data
Choosing the disease to study

- Working with entire dataset for cohort detection did not yield meaningful insights
  - **Cohort identification is usually done on subsets of patients admitted for a particular disease**

- Nursing Notes of NewBorn is noisy – contains information about mother and neo-natal – **pneumonia**

**Top 5 Disease Categories (%) for admissions to hospital**

- Newborn: 13%
- Pneumonia: 3%
- Sepsis: 2%
- Congestive Heart Failure: 2%
- Coronary Artery Disease: 1%

**Chose Pneumonia as a subset for experimentation**
Clustering Nursing Notes directly did not yield good results. *Use of Bio-medical Resources can help*

**Using SciSpacy and Metamap**

- SciSpacy detects Disease names
  - Metamap detects Disease or Syndrome
  - Acquired Abnormality
  - Congenital Abnormality
  - Mental or Behavioral Dysfunction
  - Injury or Poisoning
  - Mental process
  - Anatomical Abnormality
  - Sign or Symptom

**Detect Negation and its argument using Negex**

- Example: patient denied chest pain → neg (chest pain)

**Standardization using Unified Medical Language System (UMLS) - Entity Resolution**

- Example: Hypertension, High Blood pressure, Hypertensive Disorder, Arterial Hypertension

**Union of all entities used as Basis for representation**

- 4000+ unique entities

**UMLS - A set of files and software that brings together many health and biomedical vocabularies and standards to enable interoperability between computer systems.**

The Metathesaurus is the biggest component of the UMLS. It is a large biomedical thesaurus that is organized by concept, or meaning, and it links similar names for the same concept from nearly 200 different vocabularies.
Autoencoders for representation of Nursing Notes

Binary vector of 4541 health conditions – Large and Sparse
LOS Prediction for Pneumonia Patients using Autoencoders and CNN

**Observation**

Though accuracy improved - actual values predicted were still way off.

Shorter stays were more difficult to predict.

Varied combination of comorbidities led to different outcomes.

<table>
<thead>
<tr>
<th>Hospital Stay</th>
<th>Recall</th>
<th>Precision</th>
<th>F1-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short</td>
<td>0.64</td>
<td>0.94</td>
<td>0.76</td>
</tr>
<tr>
<td>Long</td>
<td>0.97</td>
<td>0.80</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Accuracy Score : 0.83
**K-means Clustering of Auto encoded Nursing Notes**

- Input Auto-encoder vectors
- Distance Metric - Euclidean distance
- Set $k = 2$
- Iterate to find right value of $k$ by optimizing the Silhouette Score
  - Measure Silhouette score of each point $i$ belonging to $C_i$ as
    \[ s'(i) = \frac{b'(i) - a'(i)}{\max\{a'(i), b'(i)\}} \]
    \[ a'(i) = d(i, \mu_{C_i}) \quad b'(i) = \min_{C_j \neq C_i} d(i, \mu_{C_j}) \]
  - where
  - Compute Silhouette Coefficient as
    \[ SC' = \max_k \frac{1}{N} \sum_i s'(i) \]
- Find $k$ which minimizes Silhouette Coefficient
Clustering 2106 Pneumonia Patients – 2D visualization using tSNE

How to interpret the clusters?
SHAP Values for interpreting clusters

- SHAP values are based on game theory
  - Determines the importance of each feature towards label assignment
- Features with positive SHAP values positively impact the prediction, while those with negative values have a negative impact
- Magnitude measures the strength of impact

- Using the Cluster IDs as label
- The original 4541 entity vector was used as input to a Random Forest Classifier
- Models were fed to SHAP tree explainer
- Obtained the SHAP values for each feature towards determination of the labels
SHAP Explainability for a few clusters

Cluster 0 – Diabetic Patients with Endometriosis

Cluster 1 – Influenza, Lung Consolidation absent

Cluster 2 – Acquired Abnormality of Atrium, Left Atrial Abnormality

Each cluster could be described uniquely using top three disease / symptom as attributes (present / absent)
<table>
<thead>
<tr>
<th>Clusters</th>
<th>Number of patient</th>
<th>key diseases present</th>
<th>key disease absence</th>
<th>Key drug category administrated</th>
<th>Key drug category not administrated</th>
<th>Mode LOS</th>
<th>Most patients in the age group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 0</td>
<td>175</td>
<td>Endometriosis (86%), Diabetes (74%)</td>
<td>Paroxysmal familial ventricular fibrillation, Acquired abnormality of atrium, Influenza</td>
<td>antidiabetic hormone, insulin, hypnotic/anesthetic</td>
<td>hypoglycemic agent</td>
<td>7 days</td>
<td>60-80 years</td>
</tr>
<tr>
<td>Cluster 1</td>
<td>174</td>
<td>Influenza(70%)</td>
<td>Lung Consolidation, Acquired abnormality of atrium</td>
<td>antiviral medication for influenza</td>
<td>hypoglycemic agent</td>
<td>7 days</td>
<td>60-80 years</td>
</tr>
<tr>
<td>Cluster 2</td>
<td>205</td>
<td>Acquired abnormality of atrium(99%), Left atrial abnormality(98%)</td>
<td>Paroxysmal familial ventricular fibrillation</td>
<td>analgesic/antiplatelet, proton pump inhibitor</td>
<td>vasopressor</td>
<td>4 days</td>
<td>60-80 years</td>
</tr>
<tr>
<td>Cluster 3</td>
<td>228</td>
<td>Paroxysmal familial ventricular fibrillation(91%)</td>
<td>Influenza, Acquired abnormality of atrium, Left atrial abnormality</td>
<td>antidiabetic hormone, vasopressor</td>
<td>hypoglycemic agent, corticosteroid</td>
<td>9 days</td>
<td>60-80 years</td>
</tr>
<tr>
<td>Cluster 4</td>
<td>405</td>
<td>Lung Consolidation(93%)</td>
<td>Paroxysmal familial ventricular fibrillation, Acquired abnormality of atrium, Left atrial abnormality, Endometriosis</td>
<td>carbohydrate supplement, hypoglycemic agent, antiplatelet</td>
<td>antidiabetic hormone, insulin, analgesic/antiplatelet</td>
<td>6 days</td>
<td>60-80 years</td>
</tr>
<tr>
<td>Cluster 5</td>
<td>65</td>
<td>Pleural effusion disorder(92%), Bilateral pleural effusion(78%)</td>
<td>Lung Consolidation, Paroxysmal familial ventricular fibrillation</td>
<td>hypoglycemic agent, diuretic, analgesic/antiplatelet</td>
<td>antidiabetic hormone</td>
<td>5 days</td>
<td>Above 80 years</td>
</tr>
<tr>
<td>Cluster 6</td>
<td>51</td>
<td>Atrial Premature Complexes (84%)</td>
<td>Lung Consolidation, Paroxysmal familial ventricular fibrillation</td>
<td>Vasopressor, antidiabetic hormone, proton pump inhibitor, beta-blocker, anticonvulsant/neuropathic pain agent</td>
<td>Insulin</td>
<td>8 days</td>
<td>Above 80 years</td>
</tr>
<tr>
<td>Cluster 7</td>
<td>31</td>
<td>Pneumonia (71%), Edema (50%)</td>
<td>Lung Consolidation, Paroxysmal familial ventricular fibrillation</td>
<td>Corticosteroid, Vasopressor</td>
<td>antidiabetic hormone, proton pump inhibitor, hypoglycemic agent, hypnotic/anesthetic, diuretic</td>
<td>9 days</td>
<td>60-80 years</td>
</tr>
<tr>
<td>Cluster 8</td>
<td>101</td>
<td>Left anterior fascicular block(81%), Left axis deviation(60%)</td>
<td>Lung Consolidation, Pleural effusion disorder</td>
<td>hypoglycemic agent, beta-blocker</td>
<td>antidiabetic hormone, proton pump inhibitor</td>
<td>7 days</td>
<td>Above 80 years</td>
</tr>
<tr>
<td>Cluster 9</td>
<td>120</td>
<td>Abnormal T-wave (95%)</td>
<td>Lung Consolidation, Acquired abnormality of atrium, Left atrial abnormality</td>
<td>carbohydrate supplement, analgesic/antiplatelet, beta-blocker</td>
<td>proton pump inhibitor, antidiabetic hormone, benzodiazepine</td>
<td>5 days</td>
<td>60-80 years</td>
</tr>
<tr>
<td>Cluster 10</td>
<td>516</td>
<td>No diseases predominantly occurred</td>
<td>Lung Consolidation, Paroxysmal familial ventricular fibrillation, Pleural effusion disorder, Acquired abnormality of atrium</td>
<td>hypoglycemic agent</td>
<td>antidiabetic hormone, Vasopressor, proton pump inhibitor</td>
<td>6 days</td>
<td>60-80 years</td>
</tr>
<tr>
<td>Cluster 11</td>
<td>35</td>
<td>Ventricular hypertrophy(89%)</td>
<td>Lung Consolidation, Paroxysmal familial ventricular fibrillation, Abnormal T-wave</td>
<td>hypoglycemic agent, analgesic/antiplatelet</td>
<td>antidiabetic hormone, Vasopressor</td>
<td>6 days</td>
<td>Above 80 years</td>
</tr>
</tbody>
</table>
SHAP values for Drug Associations with each cluster

Cluster 0 – Anti diabetic hormone, Insulin, vasopressor
Cluster 2 – Analgesic, proton pump inhibitor

Drug – Cluster Associations were not significantly unique
More or less same drugs were administered to patients of all clusters
Causality Analysis of Drugs - DoWHY

Disease and Symptoms of each patient were given as attributes – LOS was target variable – drugs as interventions

<table>
<thead>
<tr>
<th>Drug Class</th>
<th>Cluster 0</th>
<th>Cluster 1</th>
<th>Cluster 2</th>
<th>Cluster 3</th>
<th>Cluster 4</th>
<th>Cluster 5</th>
<th>Cluster 6</th>
<th>Cluster 7</th>
<th>Cluster 8</th>
<th>Cluster 9</th>
<th>Cluster 10</th>
<th>Cluster 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adrenocorticotropic hormone</td>
<td>3.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analgesic/ antiplatelet</td>
<td>-0.06</td>
<td>0.93</td>
<td>-0.31</td>
<td>-0.96</td>
<td>-2.02</td>
<td>1.43</td>
<td>-4.69</td>
<td>1.66</td>
<td>-0.36</td>
<td>-1.06</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td>Analgesic/ antipyretic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antibiotic/ antiprotozoal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.88</td>
</tr>
<tr>
<td>Anticonvulsants/ neuropathic pain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-3.03</td>
</tr>
<tr>
<td>Antidepressant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-1.30</td>
</tr>
<tr>
<td>Antidiabetic hormone</td>
<td>-0.46</td>
<td>1.41</td>
<td>1.59</td>
<td>1.28</td>
<td>1.41</td>
<td>-0.88</td>
<td>2.09</td>
<td>6.70</td>
<td>-0.17</td>
<td>1.37</td>
<td>-0.86</td>
<td>2.70</td>
</tr>
<tr>
<td>Antiprotein</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.81</td>
</tr>
<tr>
<td>Antiviral medication for influenza</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.70</td>
</tr>
<tr>
<td>Benodiazepine</td>
<td>4.74</td>
<td></td>
<td>1.66</td>
<td></td>
<td>0.94</td>
<td>2.31</td>
<td>0.77</td>
<td>4.03</td>
<td>2.10</td>
<td>0.30</td>
<td>1.19</td>
<td>-2.46</td>
</tr>
<tr>
<td>Beta-blocker</td>
<td></td>
<td></td>
<td></td>
<td>1.88</td>
<td>-1.36</td>
<td></td>
<td>-2.74</td>
<td>1.02</td>
<td>0.25</td>
<td>-3.98</td>
<td>-0.25</td>
<td>1.76</td>
</tr>
<tr>
<td>Bronchodilator</td>
<td>3.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-2.77</td>
</tr>
<tr>
<td>Carbohydrate supplement</td>
<td>1.62</td>
<td>0.47</td>
<td>-0.14</td>
<td>-0.20</td>
<td>0.24</td>
<td>-2.39</td>
<td>1.51</td>
<td>-6.27</td>
<td>1.77</td>
<td>2.37</td>
<td>-1.71</td>
<td>2.19</td>
</tr>
<tr>
<td>Corticosteroid</td>
<td>0.06</td>
<td>2.53</td>
<td>2.47</td>
<td>-0.86</td>
<td></td>
<td>-3.18</td>
<td>5.32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-2.92</td>
</tr>
<tr>
<td>Diuretic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electrolysate supplement</td>
<td>2.96</td>
<td>1.34</td>
<td>0.01</td>
<td>1.12</td>
<td>0.75</td>
<td>0.86</td>
<td>-5.05</td>
<td>-7.54</td>
<td>1.57</td>
<td>1.53</td>
<td>-0.30</td>
<td>-1.30</td>
</tr>
<tr>
<td>H2 blocker</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.26</td>
</tr>
<tr>
<td>Hypnotic/ anesthetic</td>
<td>6.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.60</td>
</tr>
<tr>
<td>Hypoglycemic agent</td>
<td>0.64</td>
<td>-0.46</td>
<td>-1.12</td>
<td>-1.98</td>
<td>1.10</td>
<td>-2.75</td>
<td>1.26</td>
<td>-0.17</td>
<td>0.68</td>
<td>-2.43</td>
<td>4.11</td>
<td></td>
</tr>
<tr>
<td>Insulin</td>
<td>1.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.21</td>
</tr>
<tr>
<td>Iron supplement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.27</td>
</tr>
<tr>
<td>I.v. solution electolyte</td>
<td>-2.77</td>
<td>-2.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I.v. solution electolyte/ carbohydrate</td>
<td>1.64</td>
<td></td>
<td>-1.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laxative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.08</td>
</tr>
<tr>
<td>Multivitamin supplement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.45</td>
</tr>
<tr>
<td>Opioid analgesic</td>
<td>3.33</td>
<td>1.11</td>
<td>1.45</td>
<td>2.11</td>
<td>1.57</td>
<td>-2.42</td>
<td>2.33</td>
<td>0.64</td>
<td>-0.29</td>
<td>0.19</td>
<td>1.31</td>
<td></td>
</tr>
<tr>
<td>Proton pump inhibitor</td>
<td>2.75</td>
<td>0.40</td>
<td>0.51</td>
<td>0.60</td>
<td>1.23</td>
<td>-0.67</td>
<td>1.00</td>
<td>2.45</td>
<td>-0.08</td>
<td>1.03</td>
<td>0.90</td>
<td>0.19</td>
</tr>
<tr>
<td>Statin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.14</td>
</tr>
<tr>
<td>Stool softener</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.69</td>
</tr>
<tr>
<td>Thyroid hormone</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.47</td>
</tr>
<tr>
<td>Vaccine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.10</td>
</tr>
<tr>
<td>Vasopressor</td>
<td>2.42</td>
<td>3.35</td>
<td>5.52</td>
<td>1.30</td>
<td>1.72</td>
<td>-0.45</td>
<td>-1.02</td>
<td>-2.68</td>
<td>1.47</td>
<td>1.53</td>
<td>0.94</td>
<td>6.77</td>
</tr>
<tr>
<td>Vitamin supplement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-1.17</td>
</tr>
</tbody>
</table>

**Example Observations**

**Adrenocorticotropic Hormone** → seriously ill patients Suffering from Endometriosis

For patients suffering from **arterial premature complexes in cluster 6** - drugs effective in reducing stay were

- **Corticosteroid (-3.17)**,
- **Anticonvulsants/ neuropathic pain agent (-3.02)**,
- **Opioid analgesic (-2.4)**,
- **Vasopressor (-1.02)**

---

[Image of a table with drug names and their interactions with disease attributes]
Cluster-wise Recovery Status (identified from Discharge Notes)

High Risk Pneumonia Patient cohorts identified as those with Risk of Death > 20%

Cluster 7 – Edema, Hypotensive, Left axis deviation;

Cluster 0 – Endometriosis and Diabetes;

Cluster 11 – Ventricular hypertrophy

Percentage of Deceased relatively high for clusters 7, 0, 11 (> 20%)
Risk calculation – Probability of Discharge State given Initial Symptoms

\[ P(\text{Discharge State}/\text{Symptom}) \] - computed cluster-wise and over whole dataset

*End-state is mostly dependent on Comorbidities rather than majority symptoms – rare events and combinations are more informative*

- Initial Comorbidities that have high probability of State Deceased
  - Chronic multifocal osteomyelitis \( \rightarrow \) Deceased (0.85)
  - Portal Vein Thrombosis \( \rightarrow \) Deceased (0.6)
  - Renal Osteodystrophy \( \rightarrow \) Deceased (0.6)
  - Diverticulosis of sigmoid colon \( \rightarrow \) Confusion, Ambulation with Assistance (0.7)
- *Most of the above symptoms are rare (0.01% each) and distributed across clusters*
Learnings from Patient Cohorts

- Clustering and Classification models mostly learn majority features and feature-class associations
  - Actual risk factors may be the rare symptoms

- Each patient belonging to a cluster may share majority features but differ in certain unique aspects which determine the true value of risk and also actual hospital stay
  - Not aptly captured by models

- A lot depends on correct encoding of observations
  - discovery of “Chronic multifocal osteomyelitis” after 7 days – is it a new disease or existing disease discovered?
  - If this was present in first day - would prediction be better?
  - Would it help if these factors are known a priori? Would hospitals test for these?
To Conclude

- Predictive power of Nursing notes are relatively less explored
  - Can improve prediction of Hospital stay and procedure requirements effectively
  - Accuracy of LOS prediction improves when patients of a single disease are considered
  - Explainable mechanisms can provide insights about predictions

- Better insights are possible to obtain from Patient Cohorts
  - Combination of auto-encoders and SHAP explainability offers rich insights about Risks associated

Work in Progress

- Modeling mid-stage records
  - Tracking progression of symptoms – old and new
  - Reconstruction of Recovery Pathways with probabilities
    - Modeling path to recovery as a set of transitions and onset of new knowledge
    - Explainable Risk assessment framework
  - Using other modes of data from MIMIC Database
Thank You