Analysis of Healthcare Data - Explainability using Language Models and Rough-fuzzy possibilities

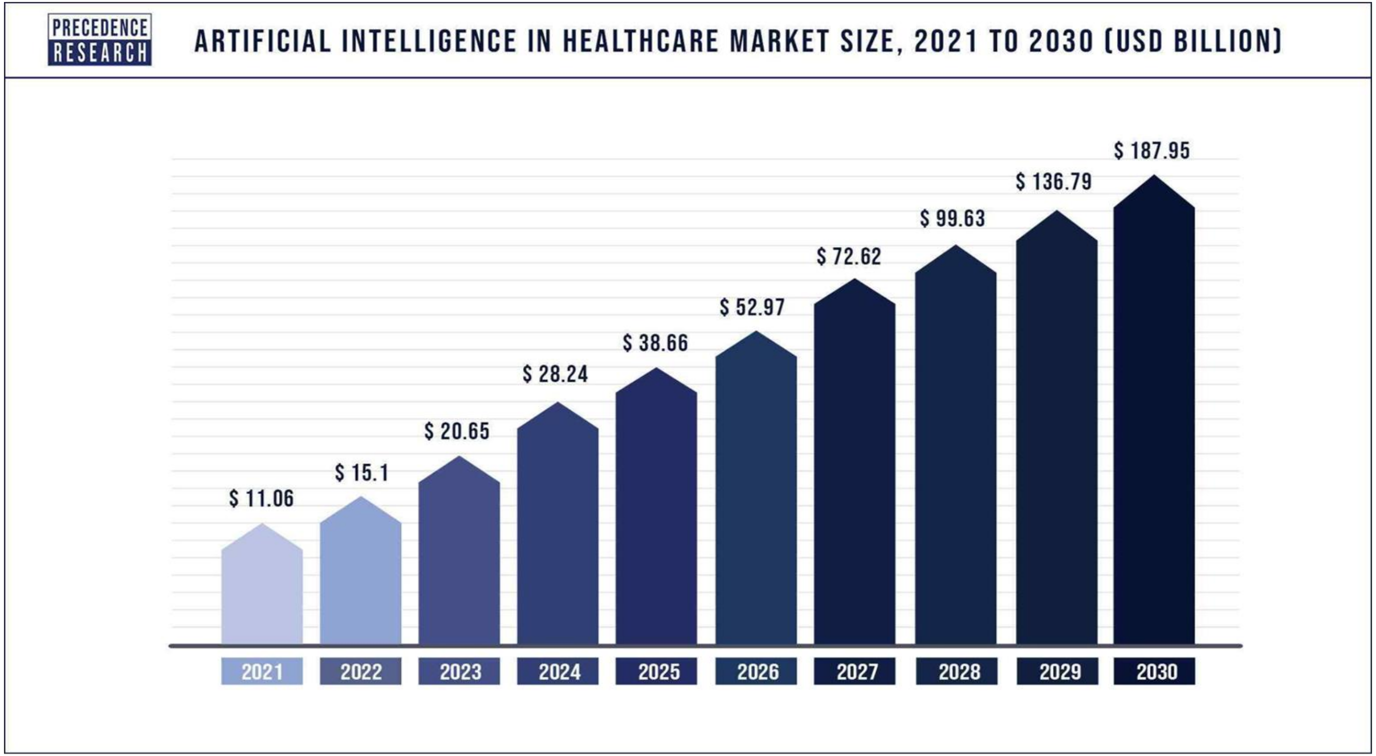
Lipika Dey

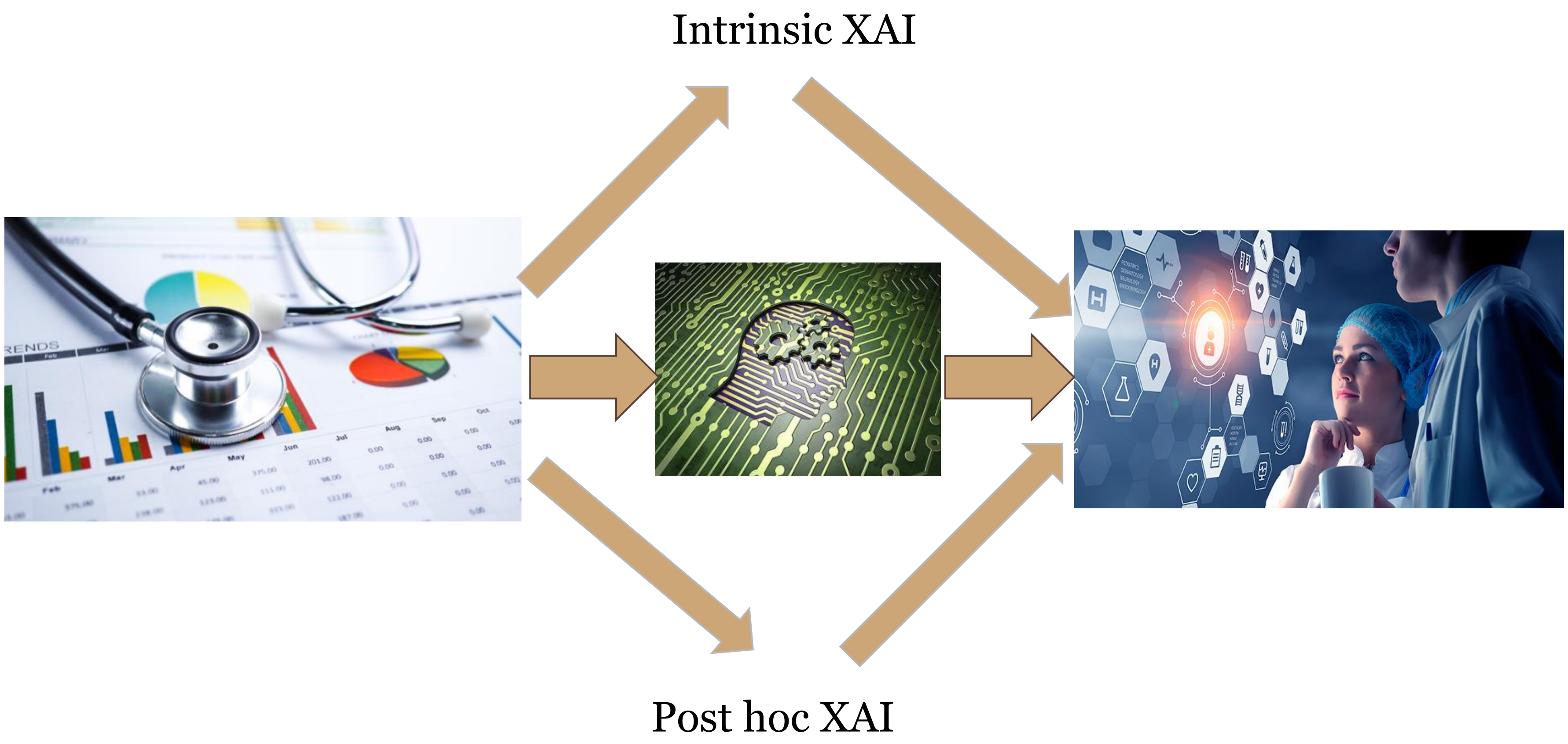
IJCRS 2024

ashoka UNIVERSITY

"AI in healthcare is said to be the hope, the hype, and the promise"

- Nazar et al., 2021

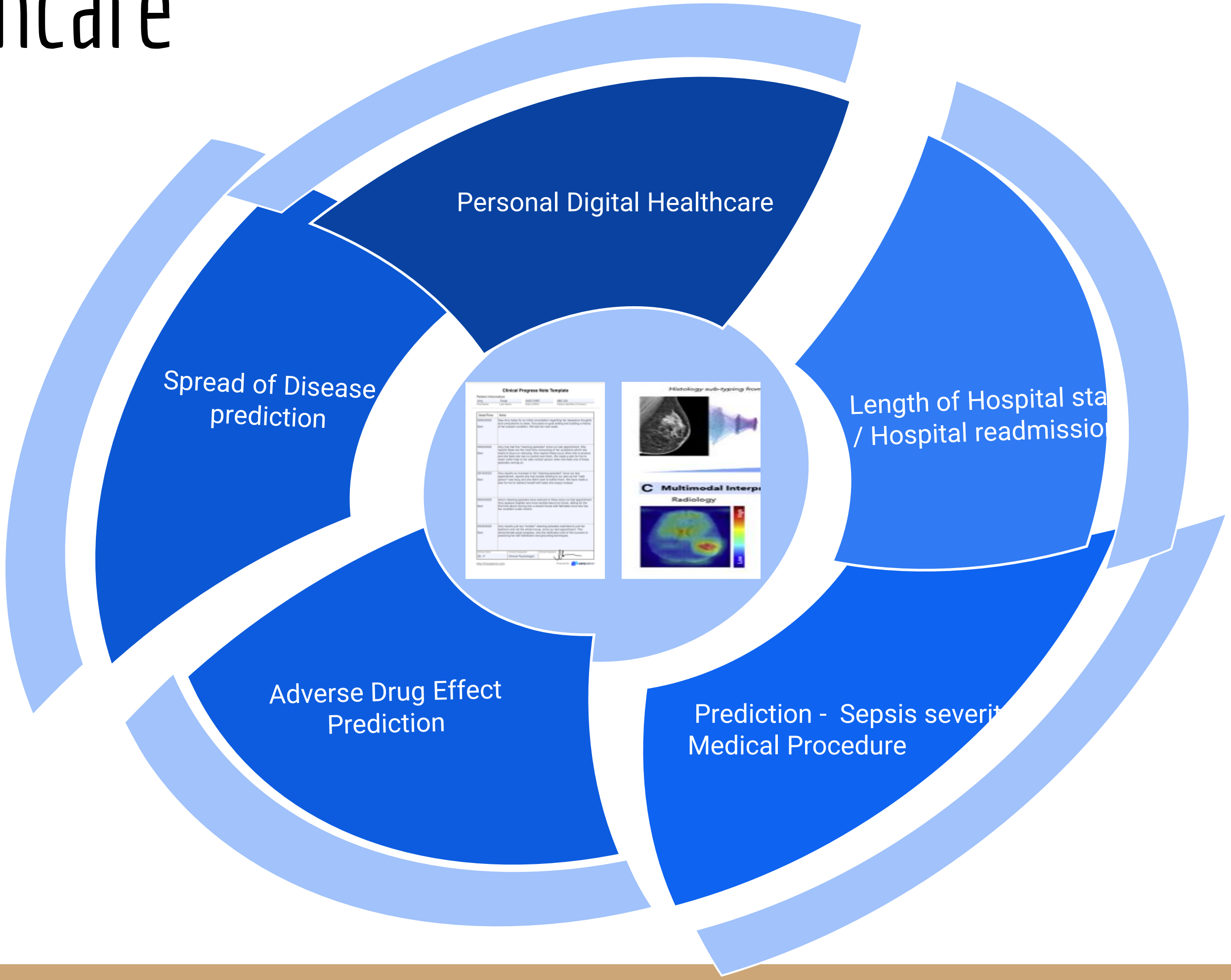




Accuracy ***Explainability*** ***Interpretability***

Reliability

AI in healthcare



Explainable AI - XAI

Explainability of model - knowledge about the model

- Important for a machine learning engineer to design optimal models

Interpretability - ability of the model to provide causal association between input and output

- Meaningful insight about the logic behind the decision arrived at using the data
- Important for Healthcare end-users

Healthcare Data

Patient Vitals

- Heart rate
- Respiratory rate
- Temperature
- Oxygen saturation
- Intubated - Yes /No
- Disease names
- Drugs

Nursing Notes

- Detailed notes from - doctor, nurse, pathologist, medical technologist

Advantage of using text for prediction

1. Details about patient's health status

1. Interpretability is better

Length of Stay at ICU / Hospital - a prediction problem

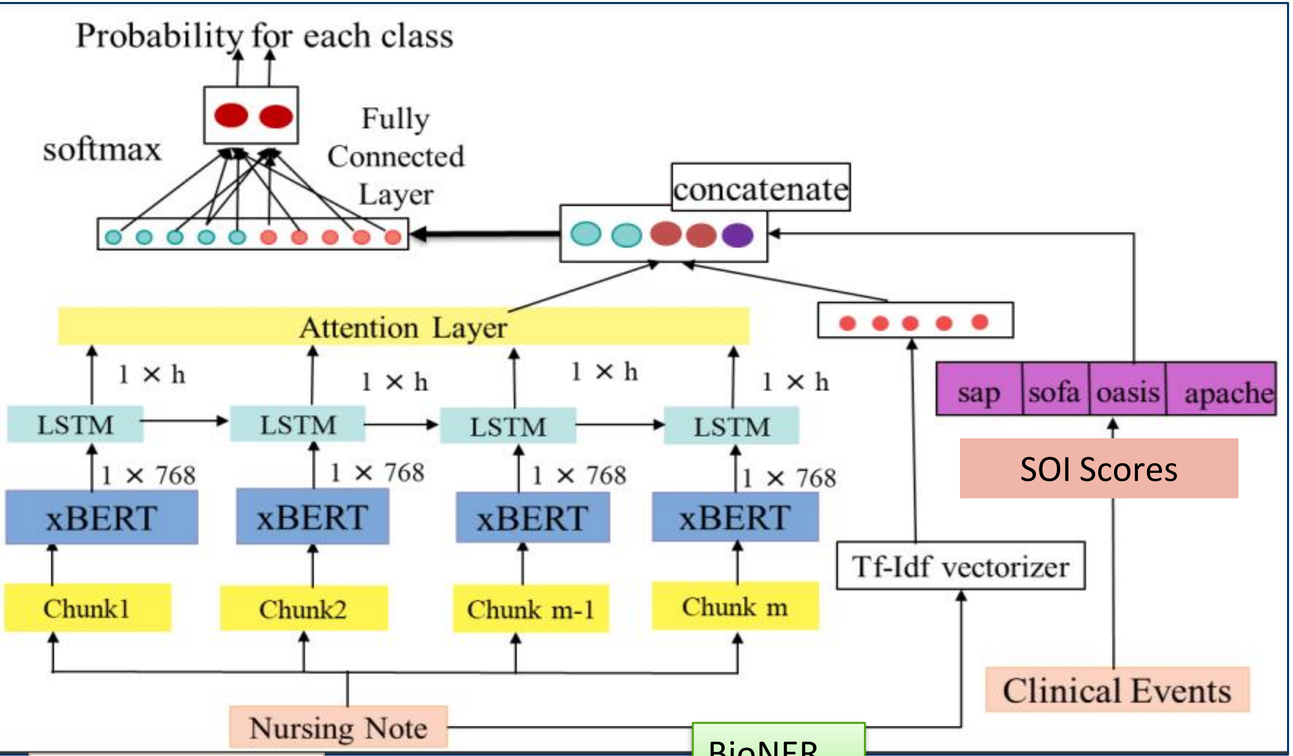
The pt is a complicated 62yo man who was transferred from ajh last evening with mrsa bacteremia and pnx. Pt arrived via EMS,intubated, sedated on Propofol 15mcg/kg-min, on Dopa gtt at 5mcg/kg-min. Transferred to Big Boy bed with 6-person assist, MICU-A monitor, and MICU-A IV pumps.**Dopamine** titrated up to max of 6.5mcg/kg-min with SBP 90's-80's, Propfol 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 repsiratory 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

Can admission statistics predict Length of Stay in ICU / Medical procedure requirement

Can these decisions be interpreted by healthcare professionals?

Deep Neural Networks for Predicting ICU LOS



SOI Scores

BioNER

Transformer + LSTM - handles long notes

TF-IDF weights of Top 2000 significant Linguistic features fed additionally

Accuracy of Prediction improves - but can clinicians interpret why?

notes

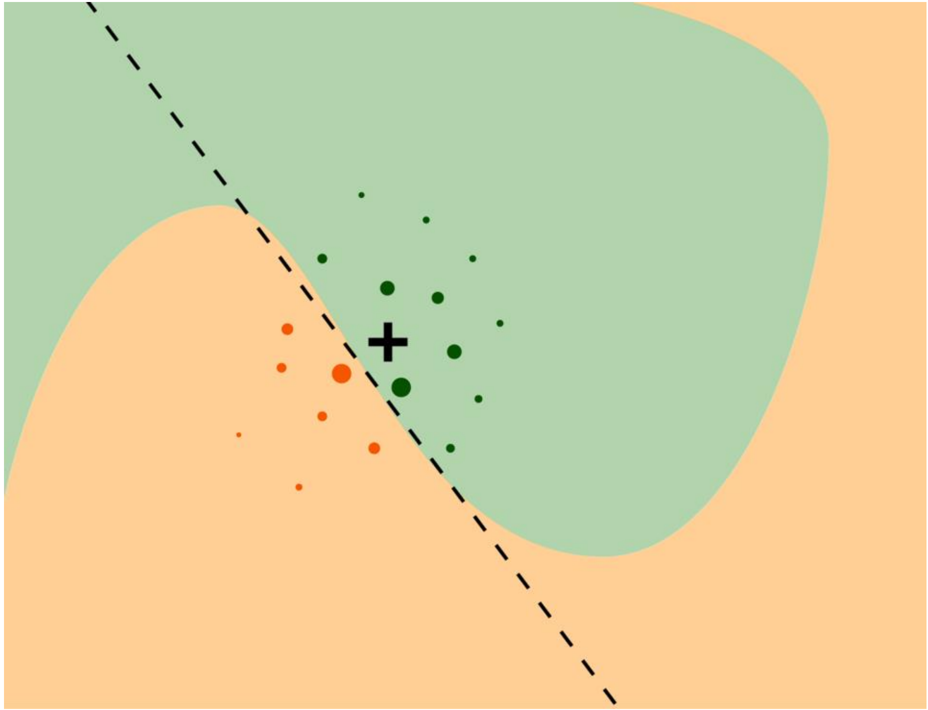
Can pre-trained LLMs be used to summarize the notes?

Predicted Length of Stay:

Given the complex medical condition, including acute respiratory distress, pneumonia, hypertension, and ESRD, the patient may require an extended hospital stay for stabilization and management of his multiple health issues. The length of stay could range from several days to potentially a week or more, depending on the response to treatment, resolution of acute symptoms, and achievement of stable medical condition. Close monitoring of respiratory status, blood pressure control, antibiotic therapy, and renal replacement therapy transition would be necessary during the hospitalization. Discharge from the ICU may be considered if the patient becomes medically stable, but overall, a prolonged length of stay is anticipated for comprehensive management and optimization of the patient's health.

infusion. The patient is aware of and agreeable to the plan of care.

Local Interpretable Model-Agnostic Explanations



Explains each individual prediction

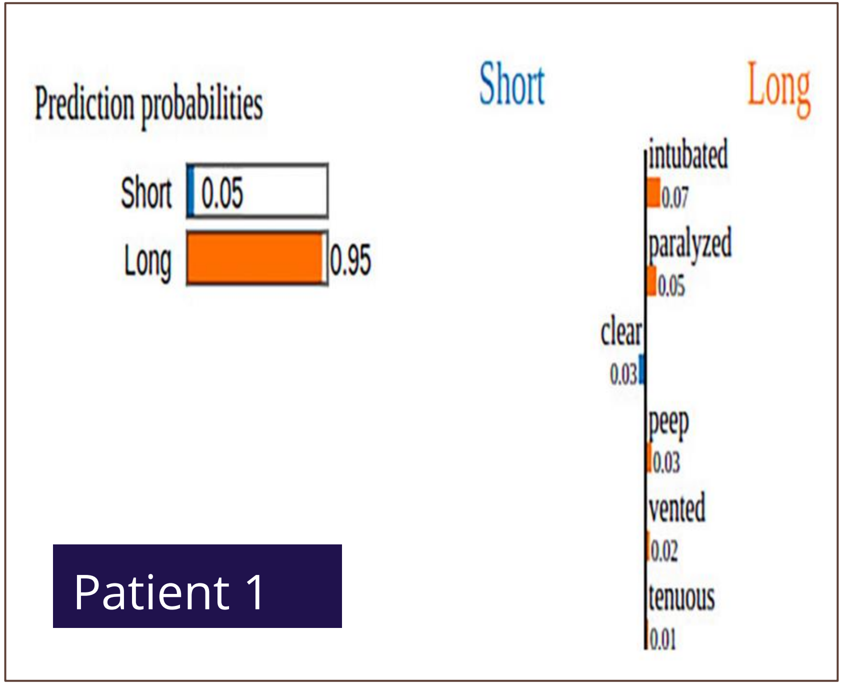
For a given black-box classifier

For an input - generates “similar” samples through perturbation

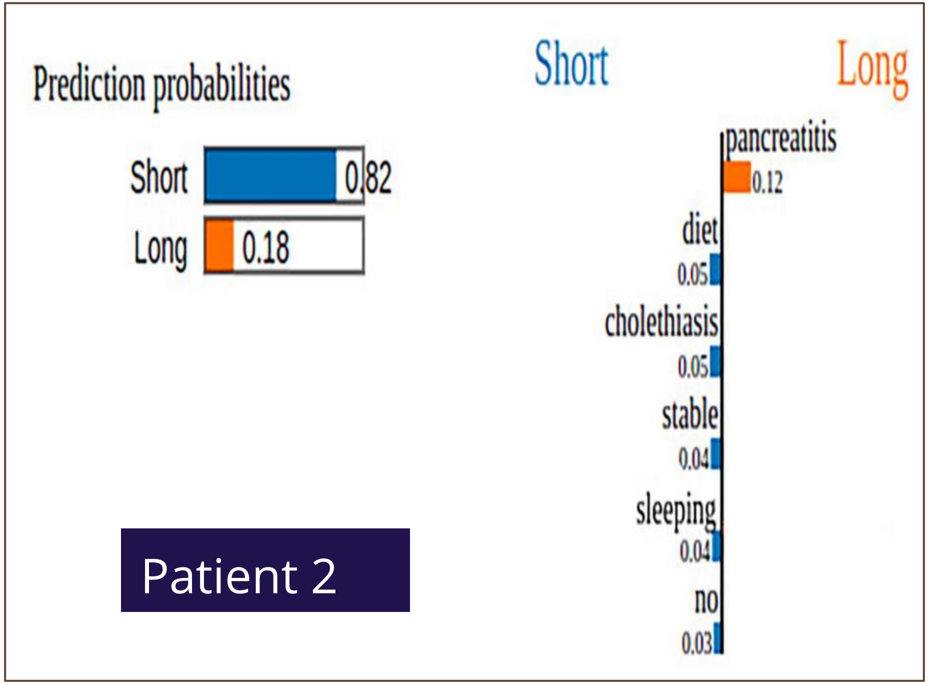
$$argmin_g \, \mathcal{L}(f, g, \pi_x) + \Omega(g)$$

minimizes the locality-aware loss $L(f,g,\Pi_x)$ - measuring how unfaithful g approximates the model to be explained f in its vicinity Π_x with least complexity

Understanding individual predictions



Patient 1



Patient 2

Contribution of each term in Clinical note towards the final prediction

Local interpretability not robust

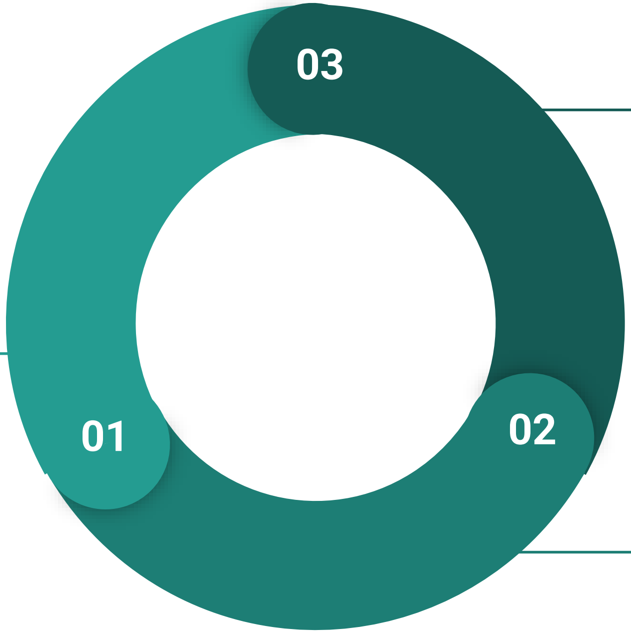
- LIME focuses on individual instances - ***overlooks global patterns***
- Clinicians require both pictures to make informed decision -

  understand each patient with respect to a cohort
- Good for structured - can’t handle text well
  - Discrete Words don’t provide the full picture
    - “No fever” vs “fever” can change the sense
    - Handling synonyms - “diabetes” vs “high blood sugar”

We want them all!

Complex Data

Clinical notes plus patient health parameters



High Interpretability

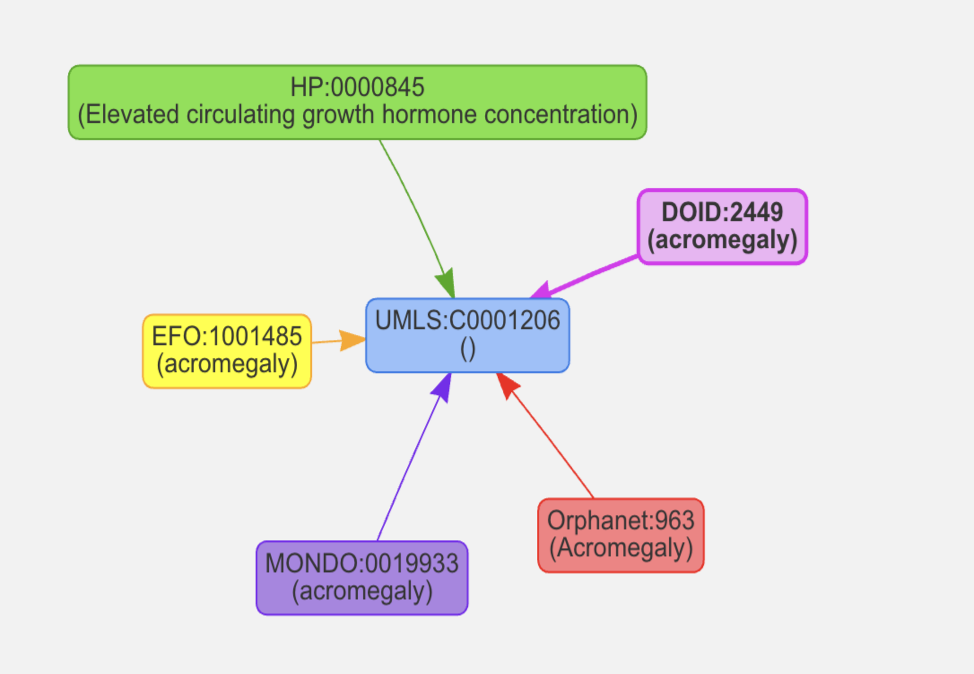
Local and Global context - individual patient along with cohort

Powerful Model

Model accuracy cannot be compromised

Unstructured ⟶ Structured Representation

Unified Medical Language System (UMLS) - brings together many health and biomedical vocabularies and standards to enable interoperability



Generating Discrete Representations from text

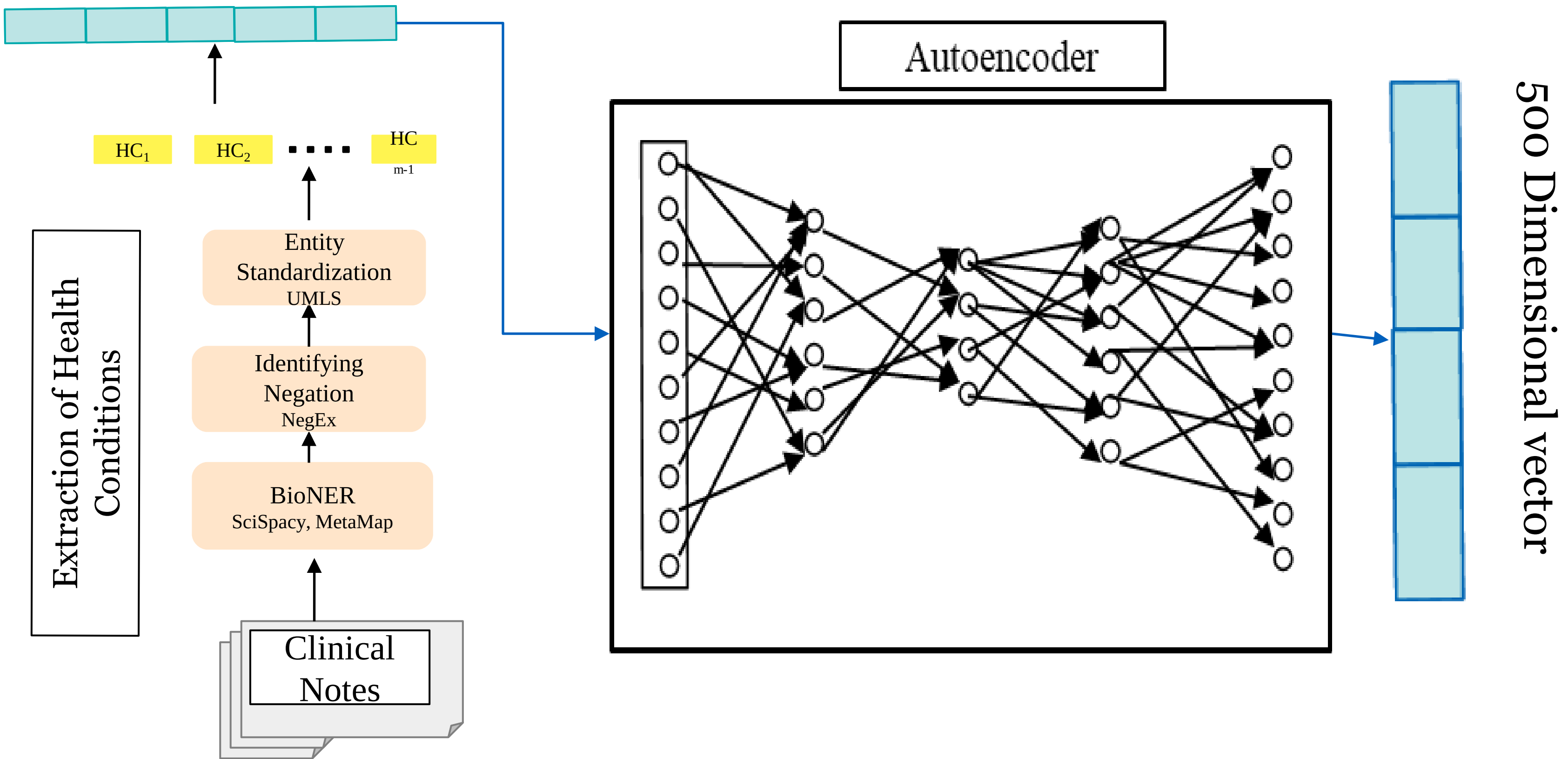
Number of features - ***very large***

O(4000 - 5000) features for patients admitted with one major disease

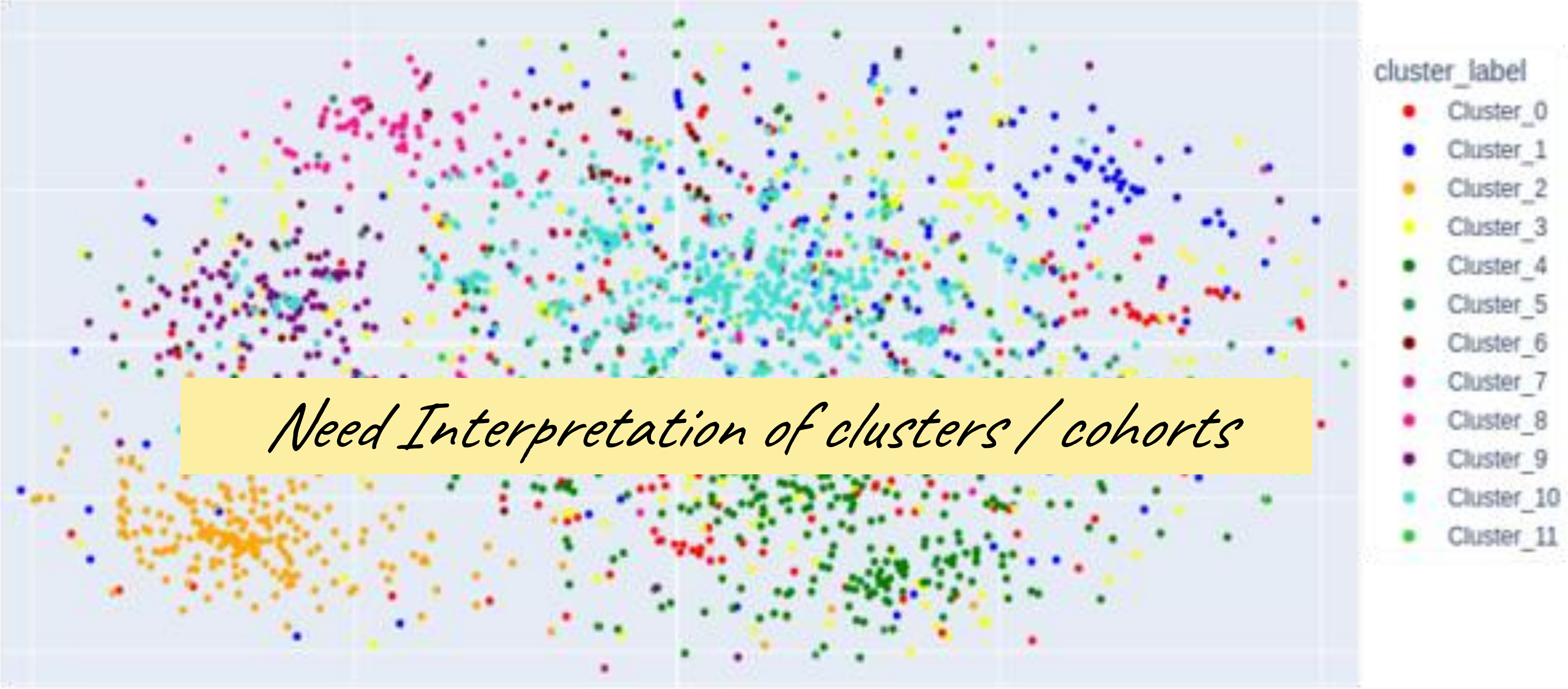
Features	Patient 1	Patient 2
accidental falls	1	1
anxiety	1	0
asleep	1	0
bowel ischemic	1	0
diabetes	1	0
hematochezia	1	0
jaundice	1	0
oliguria	1	0
pain	1	0
premature ventricular contraction	1	1
activity tolerance	0	1
anterior fascicular block	0	1
basilar rales	0	1
bundle-branch block	0	1
cough	0	1
discomfort	0	1
dyspnea	0	1
knee discomfort	0	1
pneumonia	0	1
premature ventricular contraction	1	1
sepsis	0	1

Autoencoders for representation of Nursing Notes

Binary vector of 4541 health conditions – Large and Sparse



Clustering Patients –obtaining patient cohorts



Need Interpretation of clusters / cohorts

2000+ Pneumonia patients

SHAP explainers

- Uses game theory concepts to compute feature contributions for each prediction
- Provides both local and global explanations
  - Offers insights into individual predictions (local explanations)
  - Quantifies feature importance across the entire dataset (global explanations)
- Decomposes the model's prediction into the contributions of individual features based on their importance

SHAP value computation

The SHAP value ϕ_i for feature i is calculated using the following equation:

$$\phi_i = \sum_{\mathcal{T} \subseteq \mathcal{S} \setminus \{i\}} \frac{|\mathcal{T}| \cdot |\mathcal{S}|! \cdot (|\mathcal{S}| - |\mathcal{T}| - 1)!}{|\mathcal{S}|!} [f(x_{\mathcal{T} \cup \{i\}}) - f(x_{\mathcal{T}})]$$

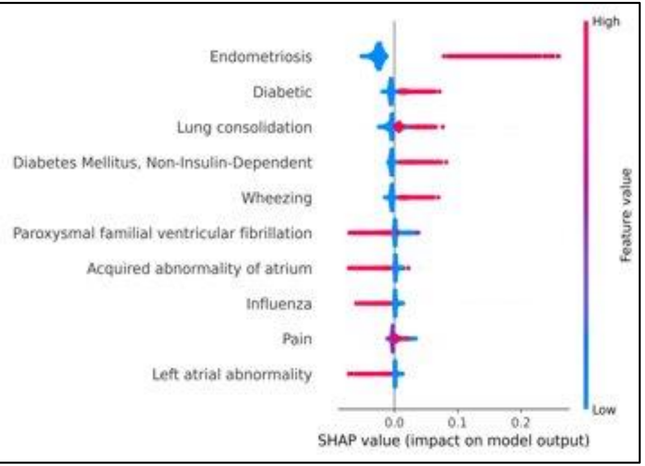
where:

- $\mathcal{T}$ represents a subset of $\mathcal{S}$ excluding feature i.
- $x_{\mathcal{T}}$ and $x_{\mathcal{T} \cup \{i\}}$ are instances with only the features in $\mathcal{T}$ and with features $\mathcal{T}$ plus feature i, respectively.
- $|\mathcal{T}|$ is the number of features in subset $\mathcal{T}$.

- *Computes for members of Power-set of S*
- *Finds features that maximize intracluster and minimize inter-cluster similarity*

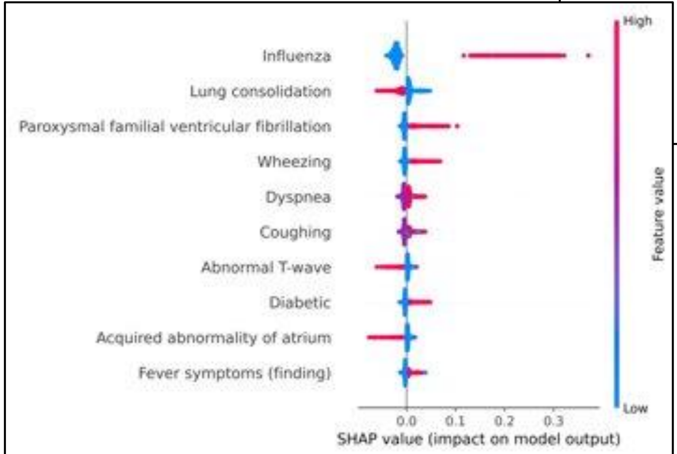
SHAP explanations for clusters

Each cluster could be described uniquely using top three disease / symptom as attributes (present / absent)

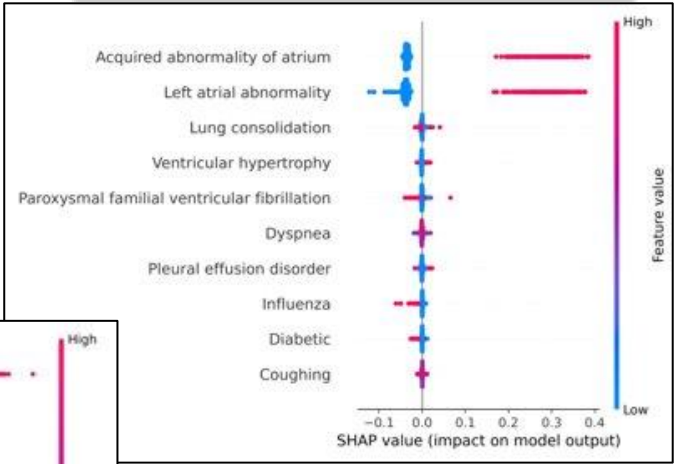


Cluster 0 – Diabetic Patients with Endometriosis

Cluster 1 – Influenza, Lung Consolidation absent



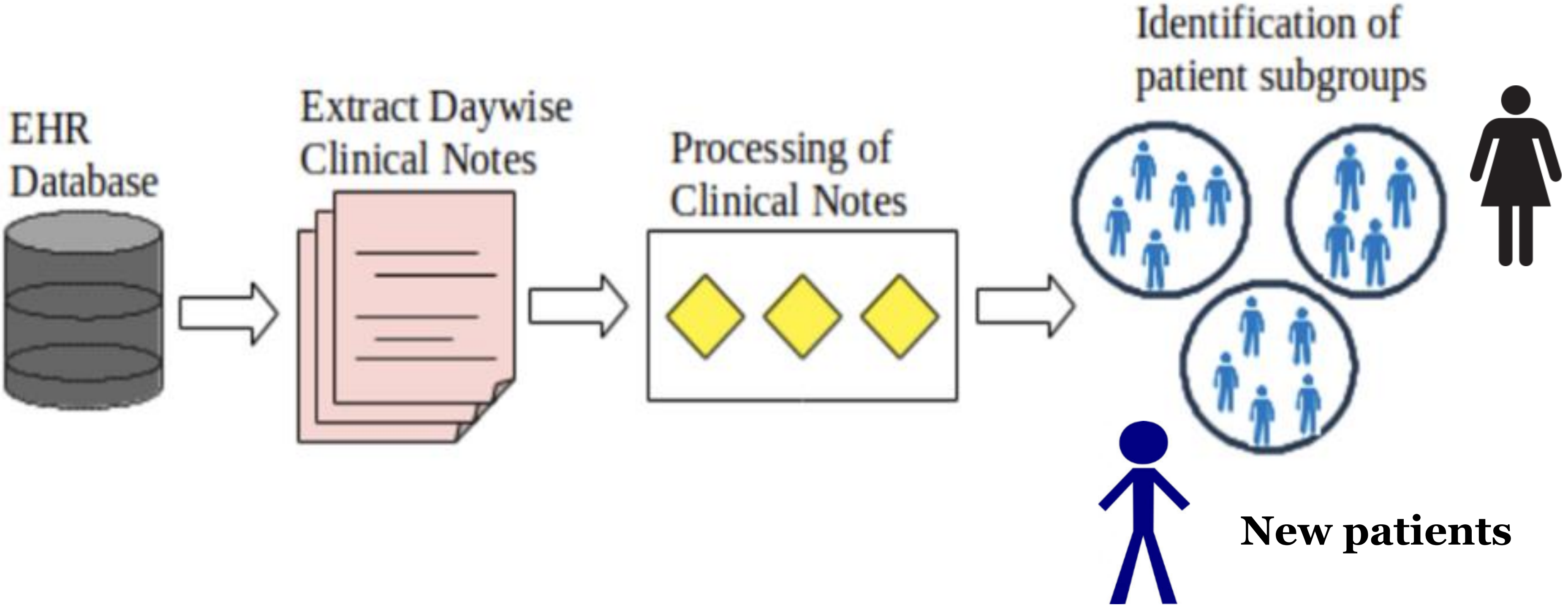
Cluster 2 – Acquired Abnormality of Atrium, Left Atrial Abnormality



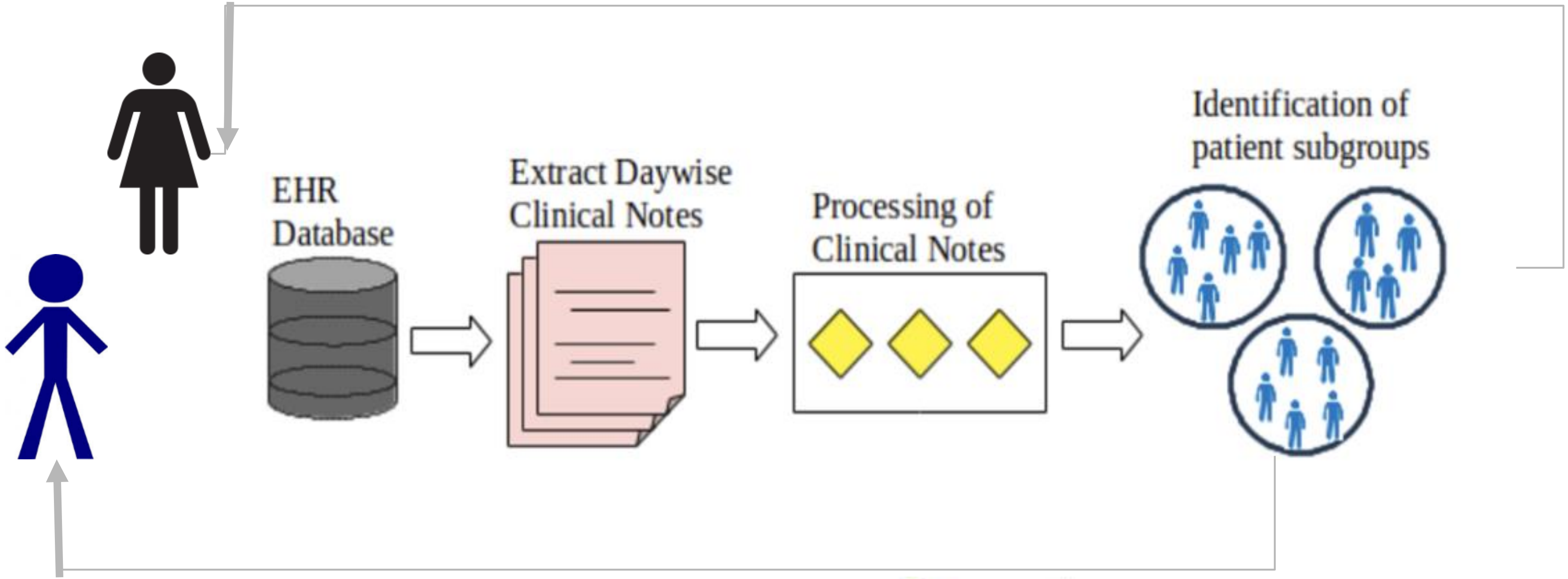
Clusters for Pneumonia Patients

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

Using Cohorts to understand new patients

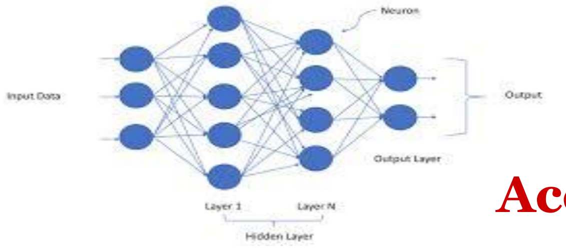


Prediction for new patient - using cohort information



Cluster ID	F1	F2	…	F500

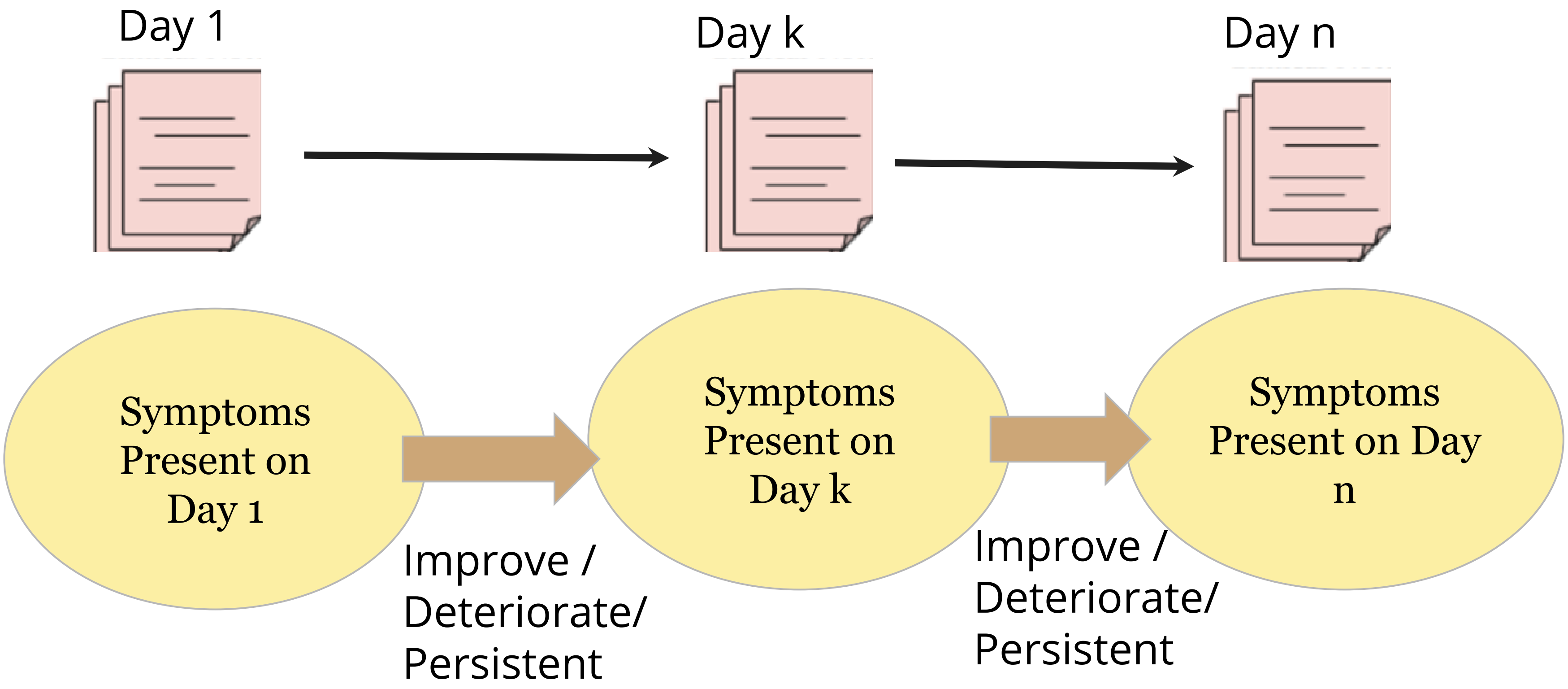
Cohort closeness information



Length of Stay

Accuracy Improves

Risk Prediction and Recovery Pathways



Transition Probabilities can be computed for each cluster

Predicting Risk for Sepsis Patients

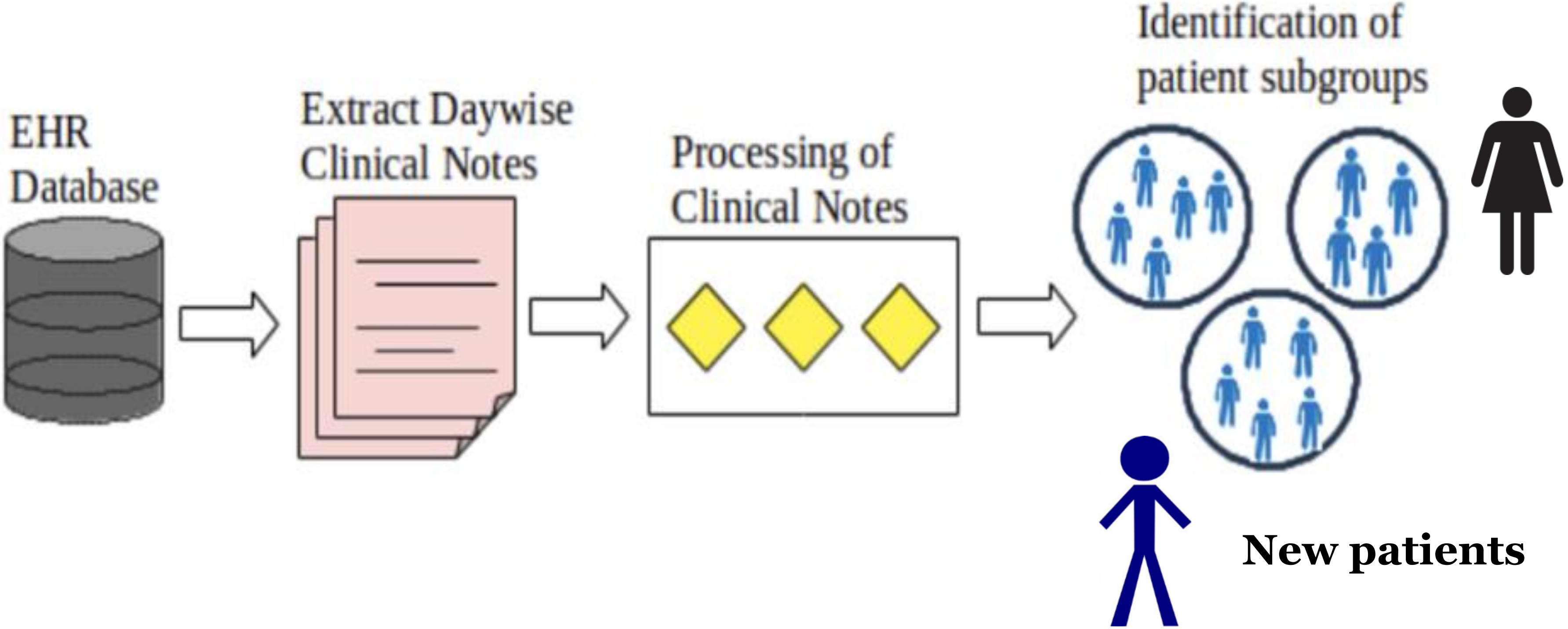
SEPSIS is a generic term - Severity score is determined based on multiple conditions

Sepsis Severity score = $\boldsymbol{f}$ (Temp., Heart rate, Resp. rate, WBC count, Organ failures)

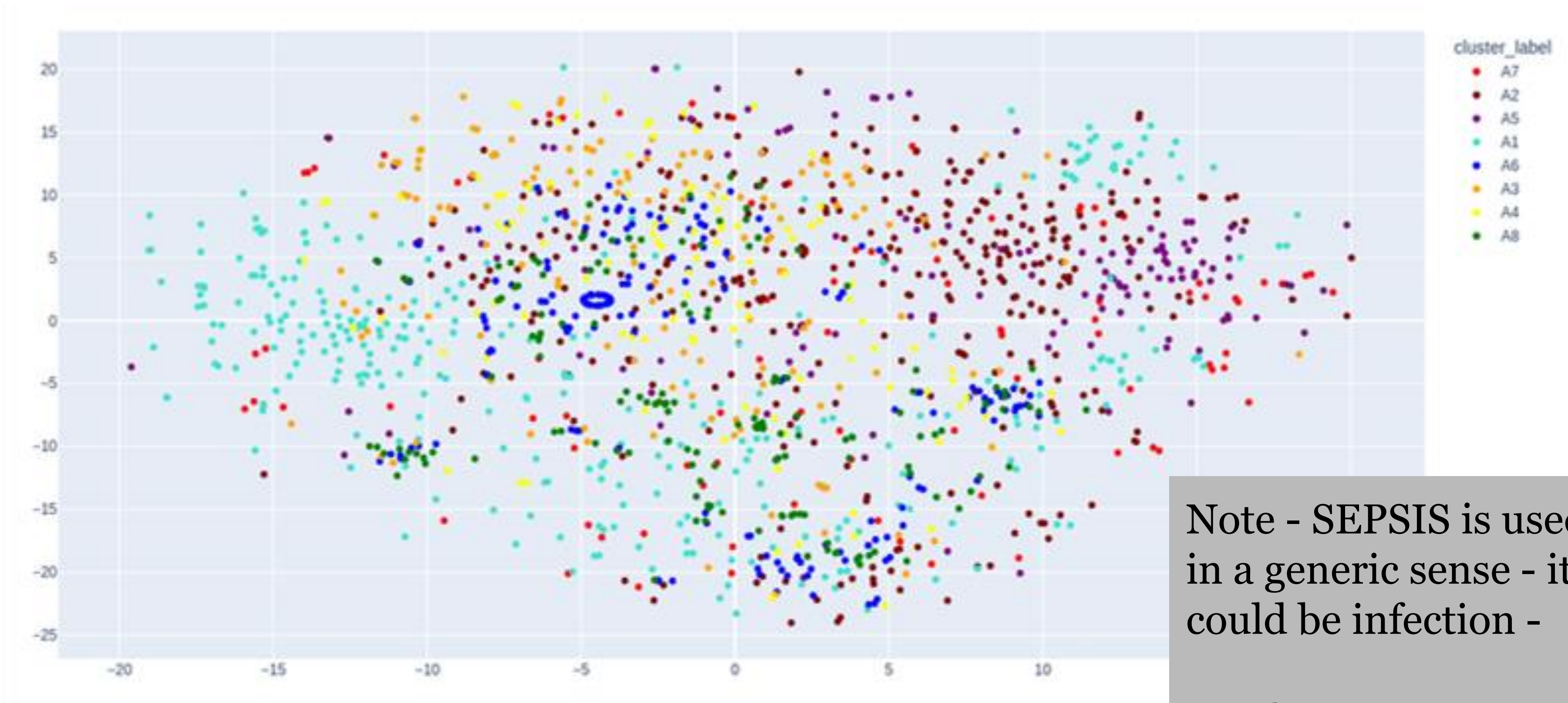
Information gathered from clinical notes

- SIRS - Severity Score 1
- Sepsis - Severity Score 2
- Severe Sepsis - Severity Score 3
- Septic Shock - Severity Score : 4

Using Cohorts to predict outcome for newly admitted patients after 3 days



MIMIC-III, v1.4, 1593 patients identified with SEPSIS on admission

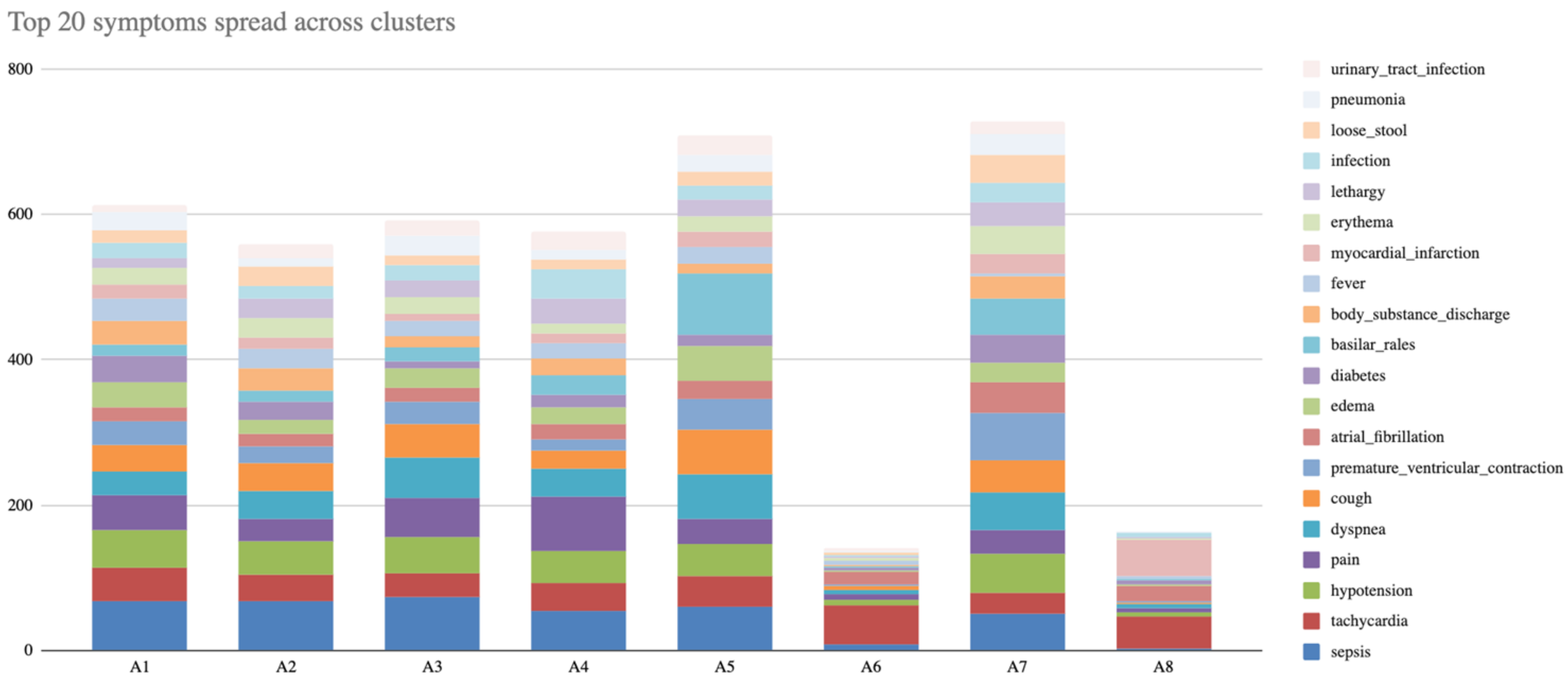


Initially grouped into 8 cohorts

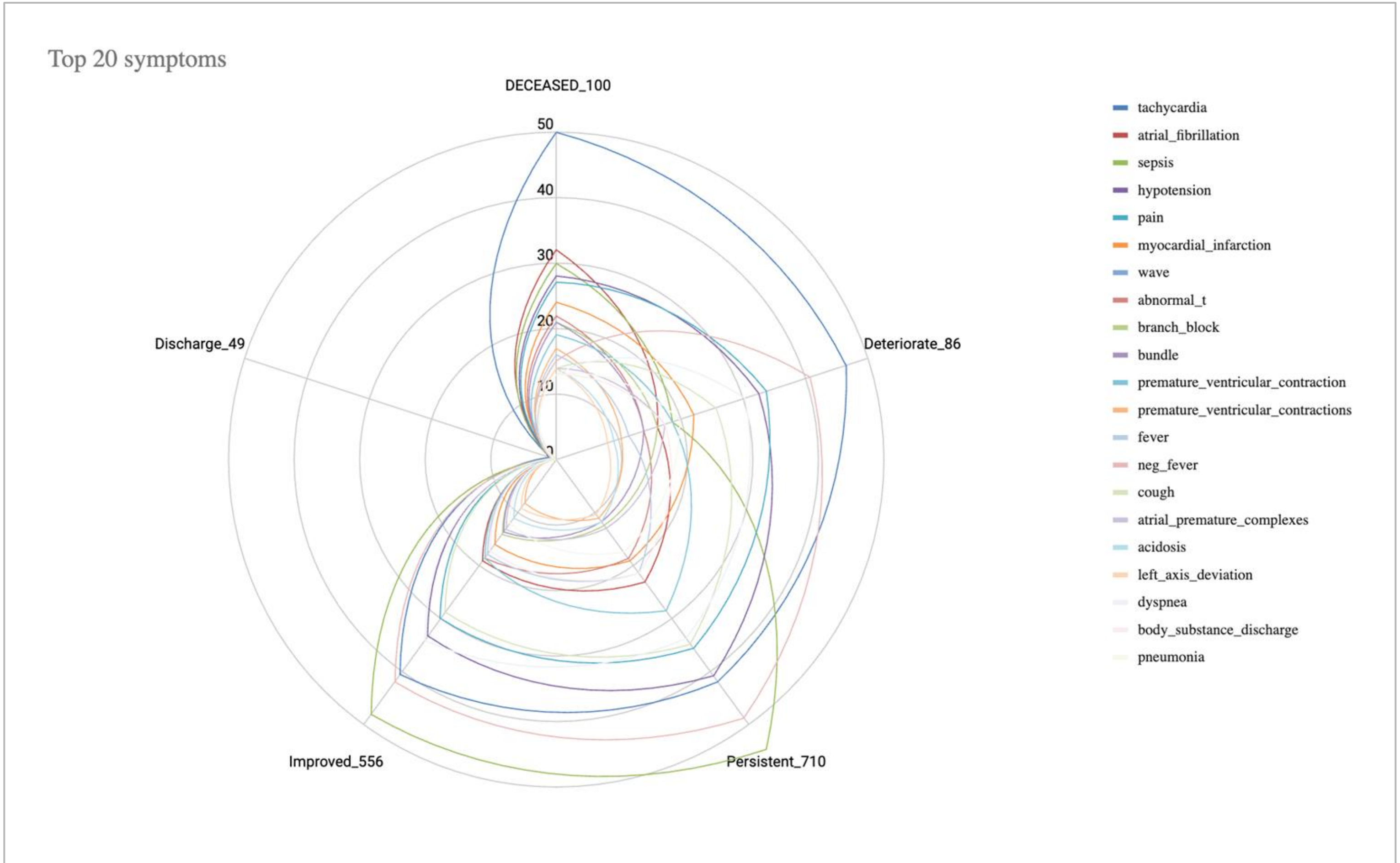
Note - SEPSIS is used in a generic sense - it could be infection -

It's the severity score that matters

Distribution of features across clusters



Feature distribution across outcomes



Discharged patients had very few symptoms

Deceased and Deteriorate, Persistent and Improved have similar distribution of features

No clear picture emerging at individual symptom level

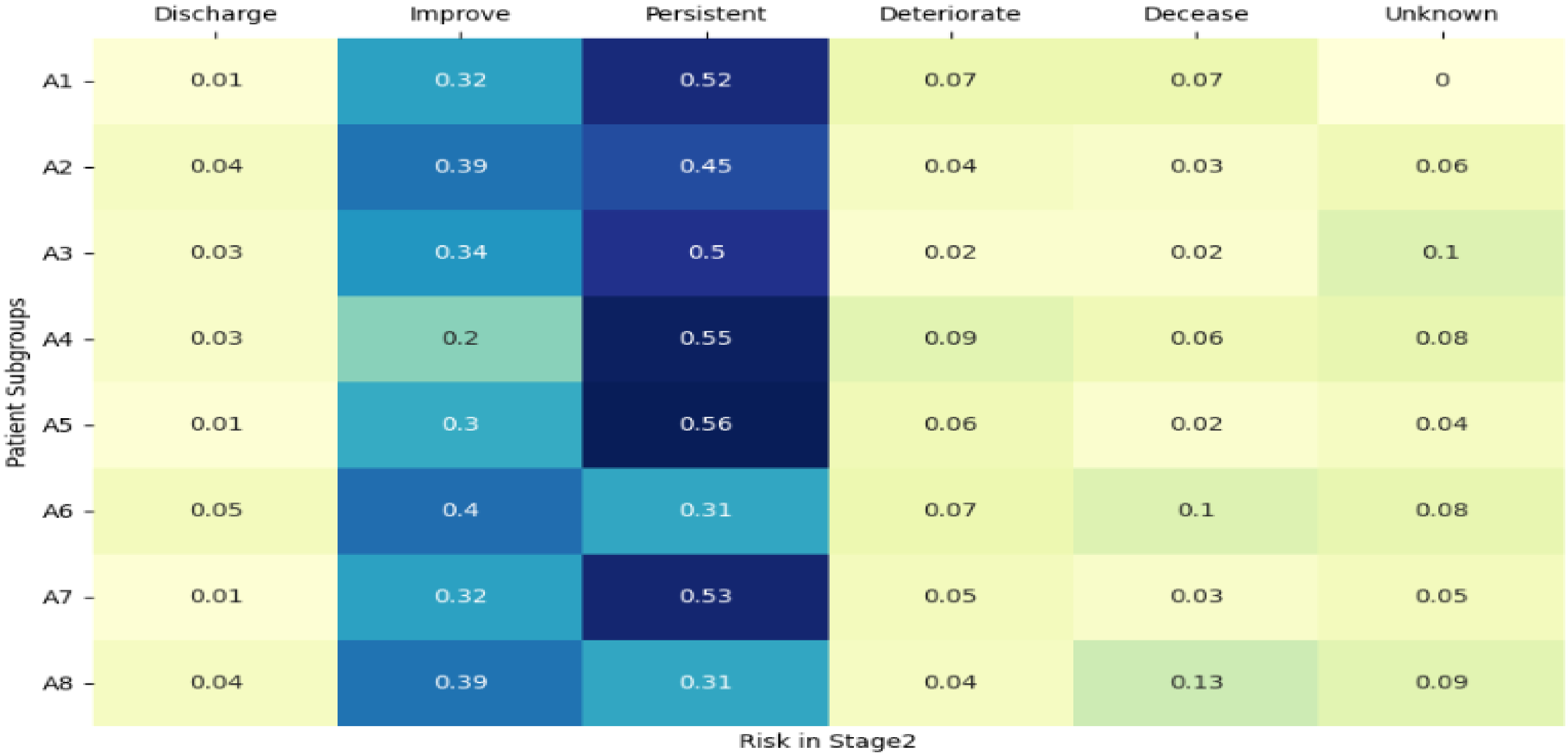
Cluster descriptions obtained using SHAP explainers

Subgroup	#Patients	prominent diseases or symptoms
A1	298	sepsis with hypotension, acidosis, diabetes, respiratory distress, pain, tachycardia
A2	339	sepsis with loose stool, hypotension absence of acidosis, pain, fever
A3	155	sepsis with dyspnea, pain, hypotension, airway disease absence of diabetes, acidosis
A4	105	sepsis with hypotension, skin infection, pain, urinary tract infection, kidney diseases
A5	128	sepsis with basilar rales, dyspnea, hypotension, edema, premature ventricular contraction (PVC), urinary tract infection (UTI), heart disease
A6	284	sepsis with tachycardia, atrial fibrillation, atrial premature complexes
A7	91	sepsis with premature ventricular contraction (PVC), hypotension, thick sputum, loose stool, diabetes, erythema, basilar rales, atrial fibrillation
A8	193	sepsis with myocardial infarction, bundle-branch block, ventricular hypertrophy, anterior fascicular block

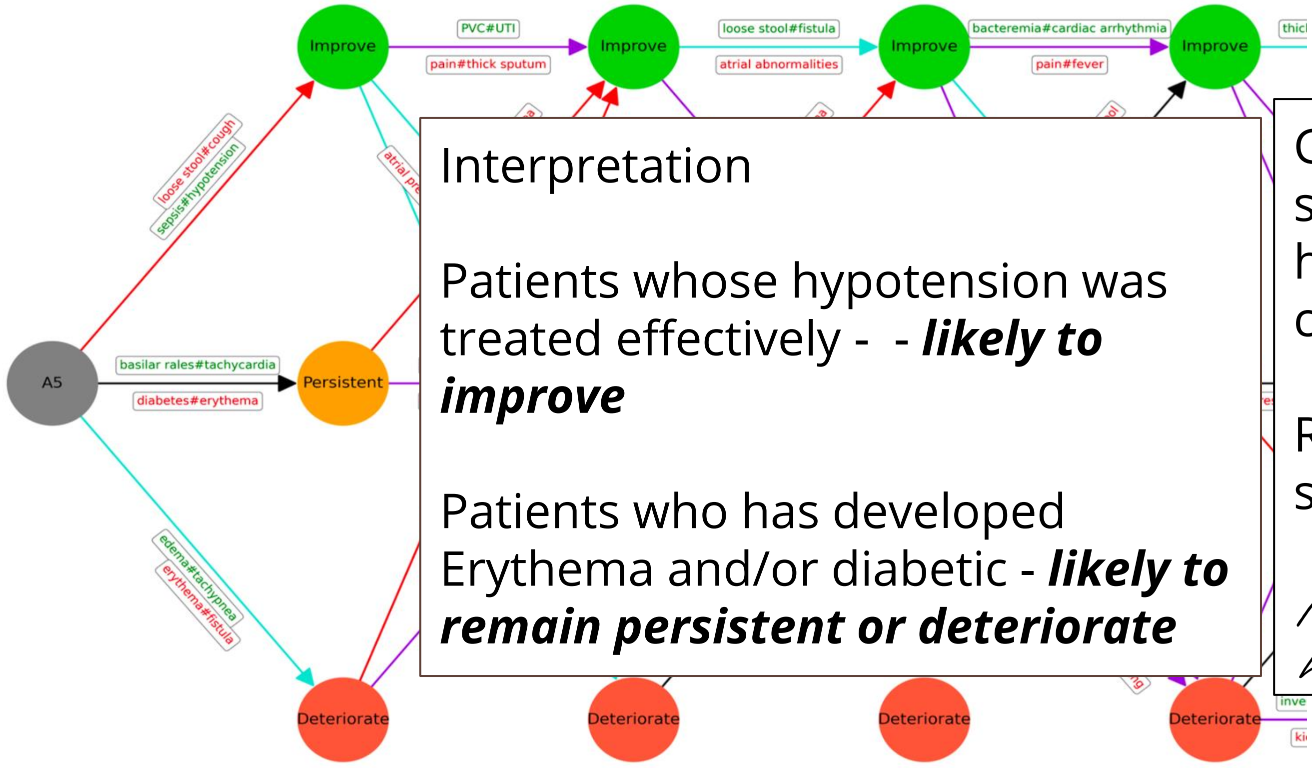
Predicting outcome after 3 days from admission

Model for predicting	**Accuracy of prediction**
BlueBERT representation of admission notes	43%
Autoencoded UML features (500 dim)	70%
Autoencoded UML features (500 dim) + cohort information	75%

Transition probabilities to next stage for each cluster



Zoom in - symptoms responsible for state transitions



Interpretation

Patients whose hypotension was treated effectively - - ***likely to improve***

Patients who has developed Erythema and/or diabetic - ***likely to remain persistent or deteriorate***

Green - symptoms that have effectively cured

Red - New symptoms

Note - only two shown here - there are others

Individual Patients

Psis
No fever
Dyspnea
Hypotension
Pain
Chills
Anemia

Improved

None

Atrial_premature_com
plexes
Sepsis
No melanoma
Weakness
Muscle_weakness
Urinary_tract_infectio
nHematuria
Osteoporosis
No urosepsis
Benign_hematuria
Hypertensive_disease
No cough
No hypotension
Mental_depression
No_dehydration

Improved

Atrial
Arrhythmia

Thick_sputum pulmonary_edema
#cough#nausea#diabetes#anasarca#agitation#yellow_sputum#hemodynamically_stable#anxiety#edema#able_to_communic
ate#dependent_for_dressing#fatigue#pressure_ulcer#premature_ventricular_contractions#body_substance_discharge#dyspne
a#pain#septic_shock#premature_ventricular_contraction#respiratory_distress#clear_sputum#tingling_in_fingers#intraventricul
ar_conduction_defect#feeling_relief#gastrointestinal_leak_nos#chest_pain

Thick_sputum
pulmonary_edema
#cough
#nausea
#diabetes
#anasarca
#agitation
#yellow_sputum
#hemodynamically_stable
#anxiety
#edema
#able_to_communicate
#dependent_for_dressing
#fatigue
#pressure_ulcer
#premature_ventricular_contractions
#body_substance_discharge
#dyspnea
#pain
#premature_ventricular_contraction
#respiratory_distress
#tingling_in_fingers
#intraventricular_conduction_defect

Persistent

cardiac problem
premature ventricular contractions
#mediastinitis
#thick sputum
#empyema
#ulceration
#bacteremia
#aortic coarctation
#anxiety
#premature ventricular contraction
#infection
#pleural effusion (disorder)
#candidemia
#deep vein thrombosis
#edema
#candidiasis
#tachycardia
#hypotension
#yellow sputum
#hernia
#diabetes

Shapley Values for Building a Clinical Decision Support System (Bienefeld et al., 2023)

Predict the onset of ***Delayed Cerebral Ischemia*** in patients with aneurysmal subarachnoid hemorrhage (aSAH)

Predicted score - 0.8

Based on static and dynamic features

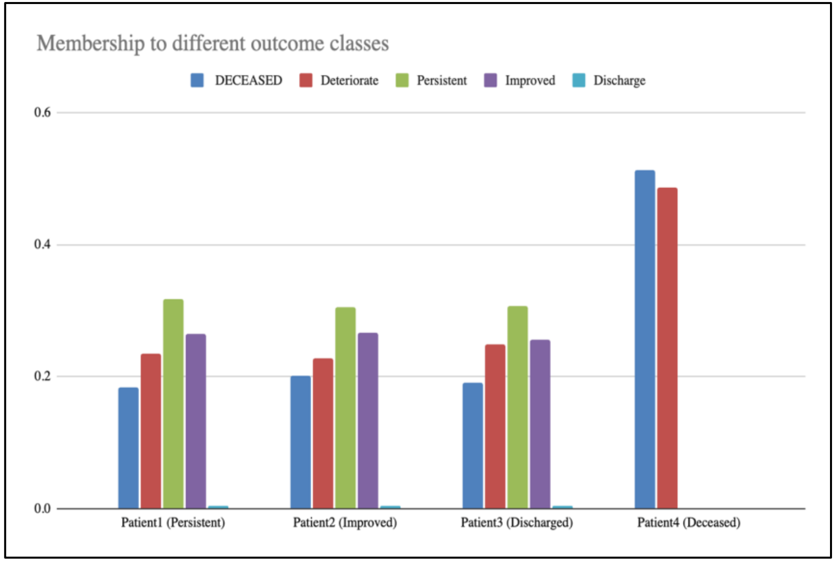
Pink - high risk

Blue - low risk

What next?

Can rough-fuzzy ideas be used for explanations?

Persistant	Improved	Discharged	Deceased
able_to_commu	asleep	abdomen_distended	abnormal_mental_state
ache	cough	abnormal_t-wave	agitation
alcohol_abuse	cyst	abrasion	anuria
anxiety	dyspnea	agitation	chronic_diarrhea
atrial_premature	erythema	atrial_abnormalities_on_ele	chronic_kidney_diseases
brown_urine	fever	atrial_premature_complexe	chronic_multifocal_osteomyelitis
clammy_skin	hypotension	brown_urine	diabetes
diabetes	incisional_pain	comfort	diarrhea
discomfort	lung_consolidati	complete_pharyngeal_conti	exanthema
feeling_relief	pain	confusion	fatigue
hemodynamicall	sepsis	cough	hemorrhage
hypotension	septicemia	crohn_disease-like_reaction	hypotension
insomnia	tachycardia	diabetes	hypovolemia
muscle_cramp		dyspnea	keratoconjunctivitis_sicca
neg_fever		edema	kidney_diseases
neg_nausea		erythema	kidney_failure
oliguria		femoral_fractures	mental_orientation
pain		fracture	multiple_lesions
pallor_of_skin		grimaces	neg_delirium
premature_ventricular_contractic		hemodynamically_stable	neg_ischemia
sepsis		hypotension	neg_nausea
spasm		loose_stool	neg_necrosis
urethral_spasm		mouth_breathing	neg_vomiting
urinary_tract_infection		neg_dysphagia	peripheral_vascular_diseases
		neg_fever	rash
		nonverbal	renal_failure



T-norm and Implicator functions can help in quantifying the dependencies between sets of attributes

To be explored ..

Contribution of combination of features to classes

For each member - predict outcome possibilities based on presence of features

Track changes in symptoms and dynamically predict next outcome

Thank You!