



@twitterhandle

## Automated Process Mining and Learning of Optimal Therapeutic Actions in the Intensive Care Unit

---

Yuval Shahar

Head, Medical Informatics Research Center  
*Ben Gurion University, Beer Sheva, Israel*





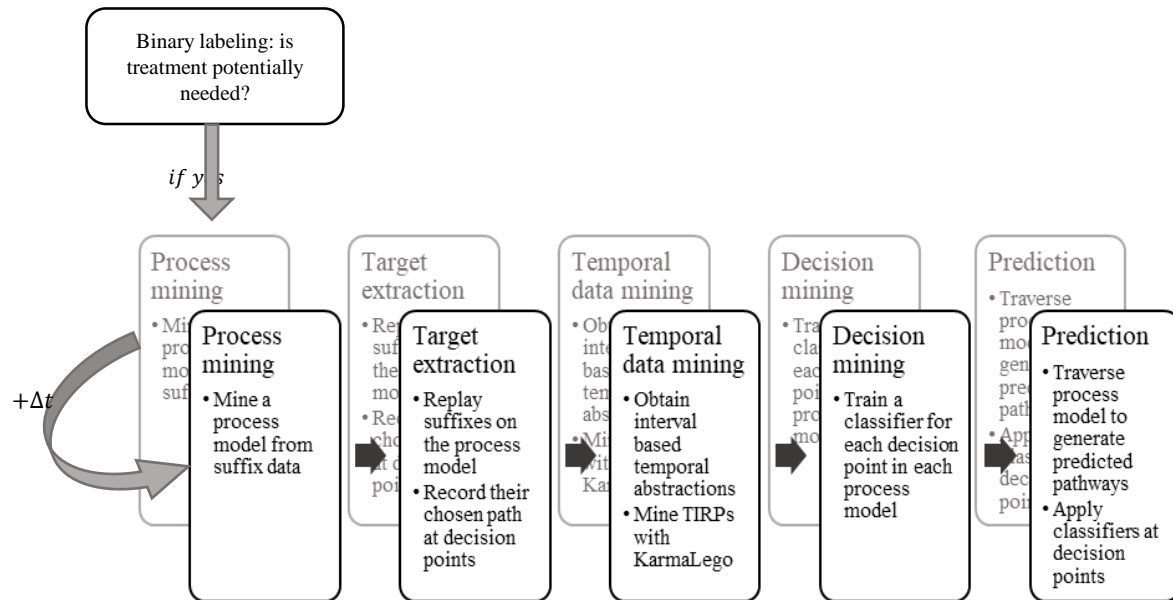
## Introduction

---

- Objective: develop an explainable framework that can -
  - Model existing clinical care pathways in the ICU
  - Given an unknown patient, predict the next actions that would usually be performed in a time window of several hours into the future
- The predictions can support the decisions of less experienced attending clinicians
- The methodology was evaluated on medical data from the hypokalemia and hypoglycemia domains using the MIMIC-IV database



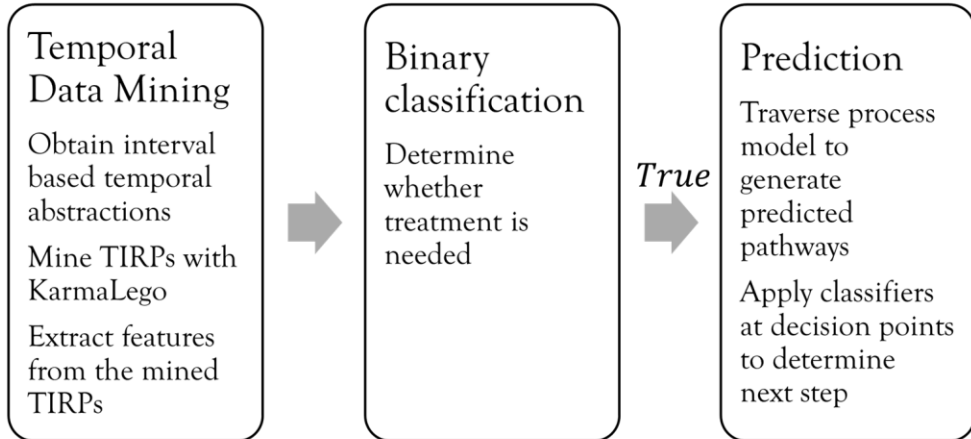
## Methodology Scheme





## Methodology

- Application of the trained framework on a new patient:
  - The data of the new patient is processed into interval-based temporal abstractions and their frequent combinations, *Temporal Interval Relation Patterns* [TIRPs], which are in turn used to construct features (using the TIRPs seen in the training set)
  - Binary classification is applied to determine whether treatment is required
  - If yes, the process model is traversed, with each visited action, represented by a node in the model, being appended to the final predicted sequence





## Binary Labelling

- Label the training data into two classes:
  - **Positive:** patients who got a treatment within the given time window
  - **Negative:** patients who did not get any treatment in the given time window
- This labelling is then used to train a classifier to determine whether treatment is needed
- Only the data labelled by the positive labels are used for the construction of the process model in the next step

suffix

### Positive label

Low Potassium
Urine Output 1.5 mL/kg/hour
Creatinine 1 mg/dL
...
Potassium Chloride
...

### Negative label

Low Potassium
Urine Output 1.5 mL/kg/hour
Creatinine 1 mg/dL
...



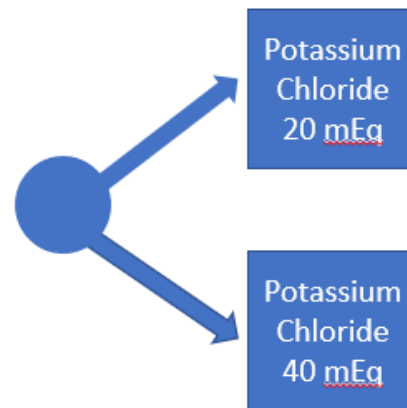
## Process Mining

- Generate a process model from the suffix data [i.e., the data following the detection of a relevant potentially treatable state], representing the typical workflows found in the data, including XOR and AND splits and joins
- Both heuristic miner and IMf are tested, and the better performing one is chosen to create the process model



## Target Extraction

- Each XOR split in the process model defines a classification problem, with outgoing arcs as the different classes
- The label of each sample for every XOR split is determined by replaying the suffix on the model and observing the chosen outgoing arc at each XOR



Example: the decision point creates a classification problem with 2 classes: Potassium Chloride 20 mEq and Potassium Chloride 40 mEq



## Feature Extraction

- Discover abstracted time intervals and mine for TIRPs using temporal data mining techniques
  - Use different values for the vertical support threshold
- TIRPs are aggregated in 3 possible ways:
  - Binary*: Was the TIRP observed for the patient?
  - Horizontal support (HS)*: How many times was each TIRP observed for the given patient?
  - Mean duration (MeanD)* of a TIRP: What was the mean duration of each TIRP for the given patient?
- The mining of the temporal patterns is performed separately for each of the found decision points
- The TIRP features, along with additional patient information, are used as the features for classifying the next action in XOR splits present in the process model.

### Binary

TIRP1	TIRP2	TIRP3	...	TIRP4
1	0	0		1

### Horizontal Support

TIRP1	TIRP2	TIRP3	...	TIRP4
2	0	0		1

### Mean Duration

TIRP1	TIRP2	TIRP3	...	TIRP4
2	0	0		10



## Feature Extraction: Additional Features

- Diagnoses families features  $F_i$  (where  $i$  indicates a diagnoses family according to ICD codes):

$$F_i(p) = \begin{cases} \sum_{d \in D_i(p)} \frac{1}{\text{ranking}(d)^2} & \text{if } D_i(p) \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

where  $D_i(p)$  denotes all the diagnoses of patient  $p$  belonging to diagnoses family  $i$ .

- Total amount of medication given within intervals of 2 hours, up to 12 hours into the past (before  $t_0$ )

Gender	Age	Diagnoses count	<u>diagnosis Diseases of the circulatory system</u>	...	<u>diagnosis Diseases of the digestive system</u>	...
1	80	7	1.3		0.1	



## Decision Mining

- Train a classifier for each decision point (XOR split) in the process models
  - Use features mined from the prefixes and targets obtained from the process model.
- We used two classifiers:
  - XGBoost
  - Logistic regression (baseline)



## Prediction

- Process and transform new prefix
  - Use only features seen in the training set
- Apply the binary classifier to the prefix data
- If the binary prediction is to administer treatment, Then traverse the process model:
  - Start from the source place
  - If current place has one outgoing arc, the next transition is appended to output
  - If current place is a decision point, the corresponding classifier is applied
- Traversal is performed according to the token replay formal definitions
  - Flow of tokens between places and transitions
  - Only enabled transitions are fired



## Evaluation Data

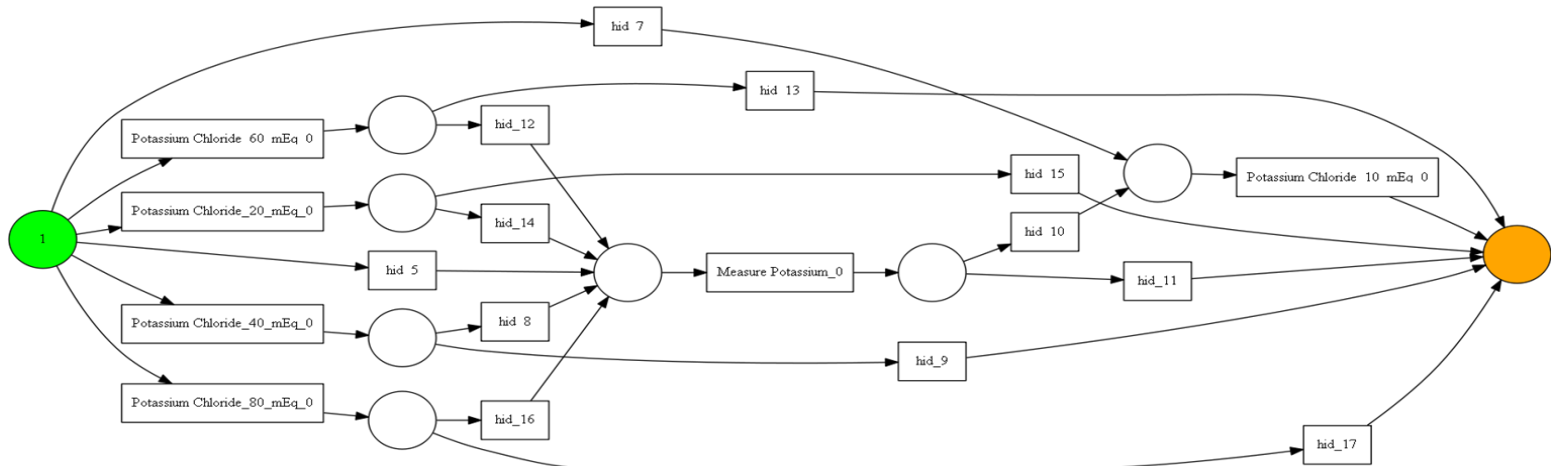
---

- Hypokalemia data:
  - Low Potassium levels in the bloodstream
  - Affects about 20% of people admitted to hospital
  - Treated by taking Potassium replacements, mainly Potassium Chloride
  - Diagnosed through a blood test
- Hypoglycemia Data
  - Low levels of glucose in the blood
  - Treated by taking Dextrose
  - Diagnosed through a blood test



## Results

- Process model example: hypokalemia (first time window)





## Hypokalemia Results

Decision point	VS threshold	Feature type	AUC	Recall	Specificity
Measure Potassium_0	0.4	MeanD	0.691	0.857	0.417
Potassium Chloride_10_mEq_0	0.4	MeanD	0.710	0.798	0.667
Potassium Chloride_20_mEq_0	0.1	MeanD	0.665	0.779	0.410
Potassium Chloride_40_mEq_0	0.1	HS	0.641	0.563	0.664
Potassium Chloride_60_mEq_0	0.1	Binary	0.639	0.487	0.766
Potassium Chloride_80_mEq_0	0.1	Binary	0.663	0.487	0.766
source0	0.1	Binary	0.537	0.234	0.799
binary	0.1	Binary	0.715	0.600	0.708

Classification results in the decision points of the first time window data of Hypokalemia.

Decision point	VS threshold	Feature type	AUC	Recall	Specificity
Measure Potassium_0	0.5	MeanD	0.526	0.280	0.790
Potassium Chloride_10_mEq_0	0.5	Binary	0.601	0.587	0.619
Potassium Chloride_20_mEq_0	0.2	Binary	0.586	0.727	0.411
Potassium Chloride_40_mEq_0	0.3	Binary	0.587	0.762	0.393
Potassium Chloride_60_mEq_0	0.1	MeanD	---	0.000	1.000
MeasurePotassium1MeasurePotassium0	0.1	MeanD	---	0.631	0.798
Source0	0.4	MeanD	0.625	0.481	0.650
Binary	0.1	HS	0.691	0.600	0.712

Classification results in the decision points of the second time window data of Hypokalemia.



## Hypoglycemia Results

Decision point	VS threshold	Feature type	AUC	Recall	Specificity
Dextrose_10_gr_0	0.4	MeanD	0.671	0.433	0.859
Dextrose_25_gr_0	0.4	Binary	0.715	0.655	0.704
Dextrose_50_gr_0	0.2	Binary	0.612	0.601	0.662
Dextrose_5_gr_0	0.4	HS	0.769	0.659	0.815
Dextrose_above 50_gr_0	0.1	HS	0.715	0.729	0.609
Measure Glucose_0	0.3	HS	0.546	0.454	0.780
source0	0.1	MeanD	0.656	0.392	0.795
MeasureGlucose1MeasureGlucose0	0.1	Binary	0.663	0.454	0.780
binary	0.1	MeanD	0.972	0.907	0.924

Classification results in the decision points of the first time window data of Hypoglycemia

Decision point	VS threshold	Feature type	AUC	Recall	Specificity
Dextrose_10_gr_0	0.1	HS	0.761	0.744	0.856
Dextrose_25_gr_0	0.2	HS	0.708	0.576	0.793
Dextrose_50_gr_0	0.1	HS	0.584	0.865	0.288
Dextrose_5_gr_0	0.1	MeanD	0.703	0.860	0.421
Dextrose_above 50_gr_0	0.4	MeanD	0.707	0.511	0.822
Measure Glucose_0	0.2	MeanD	0.618	0.637	0.570
Source0	0.1	Binary	0.602	0.337	0.796
MeasureGlucose1MeasureGlucose0	0.2	HS	0.617	0.367	0.875
Binary	0.1	Binary	0.963	0.926	0.856

Classification results in the decision points of the second time window data of Hypoglycemia



## Jaccard Index

Domain	Time window	Jaccard index
Hypokalemia	first	0.363
	second	0.283
Hypoglycemia	first	0.370
	second	0.318



## Discussion

---

- High variance in decision points - both in data, performance and chosen configuration
- The heuristic miner performed better than the IMf
- Best performance: first time window of hypokalemia, followed by the first time window of hypoglycemia
- Fitness values were lower than precision values.
  - Limiting unseen behavior allowed by the model is 'easier' than fully modeling all seen behavior.
- Overall high performance in the process models
  - Both fitness and precision scored close to 1



## Discussion: Classification

- The optimal configuration (vertical support threshold, pattern aggregation) varies between decision points
- Recall-specificity more balanced in the hypoglycemia domain
- General results in all time windows:
  - AUC range: between 0.5 and 0.75
  - Recall range: between 0.2 and 0.8
  - Specificity range: between 0.4 and 0.8
- Relatively low performance of the source0 decision points.
  - Has structural significance as the first decision point of the model



## Contribution

---

- Integration & comparison between process mining and temporal data mining
  - Both use temporal data as input
  - Temporal data mining highlights some local, low-resolution behavior, more detailed and expressive
  - Process mining focuses on "the big picture" and allows to zoom out and assess larger pathways
  - Maximally beneficial when used together, as demonstrated in this work
- Representation of high-resolution longitudinal medical records in a process model
- Dataset construction based on process models & classifiers training using those datasets
- Ability to predict all of events expected within a future time window
  - Not limited to one-step predictions
  - Can predict AND relationships
- The use of time windows allow more generalizations by using different data
  - For example, predictions for non-immediate time windows
- Sequential classification prevents false positives, by first determining whether *any* therapy is needed



## Questions