



A Dashboard for emulating LSTM-based Predictive Process Monitoring and its Qualitative Evaluation

29TH NOVEMBER 2021



Rehan Fazal

Assessment Committee
Dr. D. Fahland
Dr. M.R.V. Chaudron
Dr. R. Medeiros de Carvalho

Agenda



Overview



Backend Design



Frontend Design

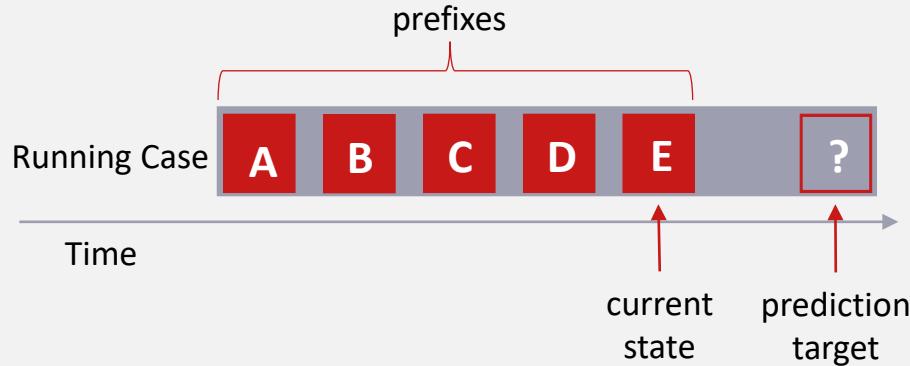
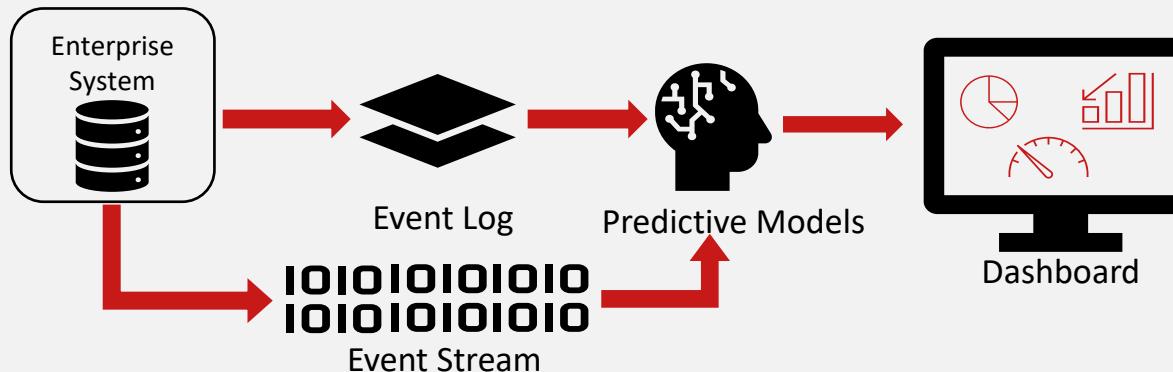


Evaluation



Conclusion

Predictive Process Monitoring “in a Nutshell”



- What is the next activity for this case?
- When is this next activity going to take place?
- What is the outcome of this case?

Motivation

Current Advances

- ✓ Extensive Research in Predictive Models
 - ❑ Next Event in a Case
 - ❑ Completion time of Event
 - ❑ Remaining execution of a trace
- ✓ High Accuracy

Current Gaps

- ✓ Not yet Applicable in Real World
 - ❑ Predictions beyond Business Users' comprehension
 - ❑ Gap in domain knowledge integration
- ✓ No Alternate Recommendations
- ✓ Focus only on most likely outcome
- ✓ No general predictive framework for business users like domain experts

So, what would a domain expert need?

Domain Experts

What they would want?



Who are they?

- Experts in the business domain with valuable functional knowledge
- With limited ML background

- ✓ Trust : accurate predictive mechanism
- ✓ Explainability : leveraging domain knowledge to explain specific cases
 - Satisfaction : ease of the use
 - Effectiveness : aiding in the decision-making
 - Efficiency : make fast decisions
 - Scrutability : permitting user-system communication

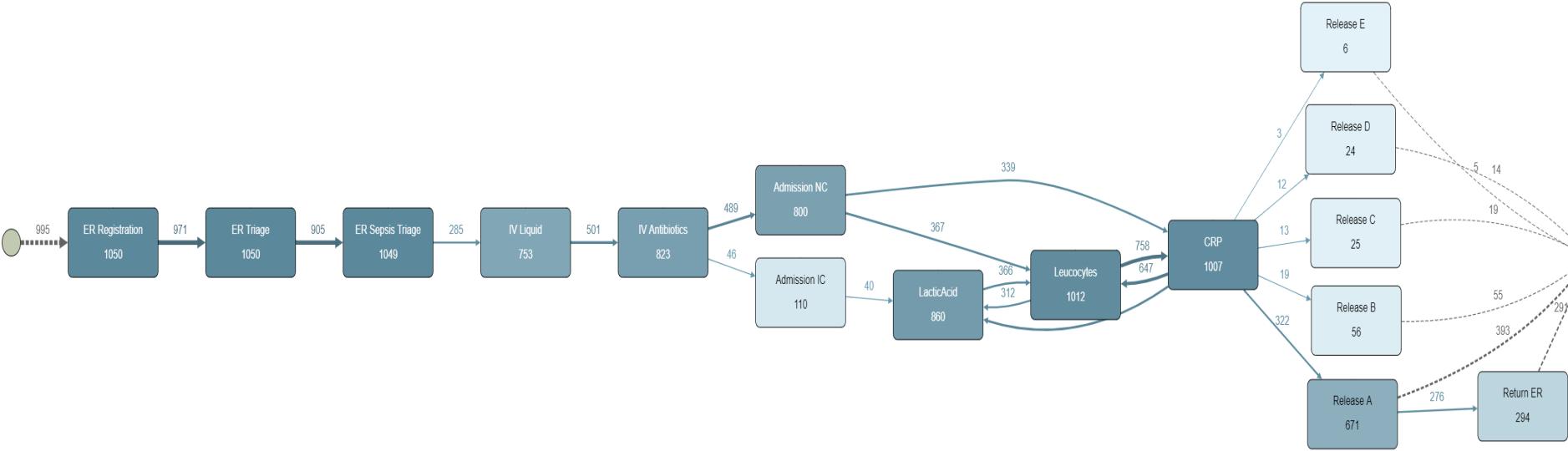
How should such Predictive Process Monitoring system be designed?

Research Problem

- Predictive Process Monitoring system which can
 - ❖ Integrate domain knowledge in Predictive Model without sacrificing accuracy
 - ❖ Dashboard be designed for Business Users specifically Domain Experts and solving current gaps
 - ❖ Translate and Explain the recommendations to Domain Experts

Real Life Event Log

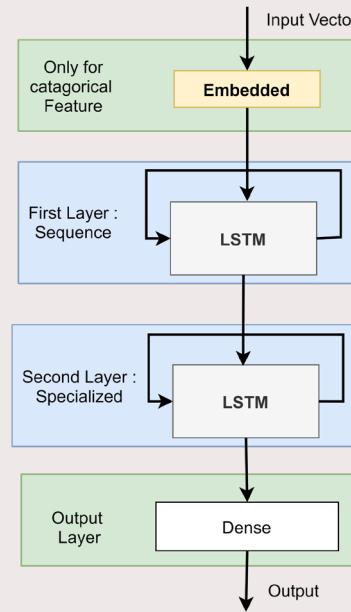
Sepsis Cases : A real-life event log, records of the trajectories of patients with life threatening sepsis symptoms at a Dutch hospital.



Predictive Model – RNN-LSTM

Which Predictive Model?

- ✓ Structured predictor based model → Next Action
- ✓ Recurrent neural networks (RNNs) is an ideal fit for predictive model because of the **sequential structure of business processes**
- ✓ Recent advances in RNN-LSTM model architecture lets to **predict Activity, Role and Time at the same time.**

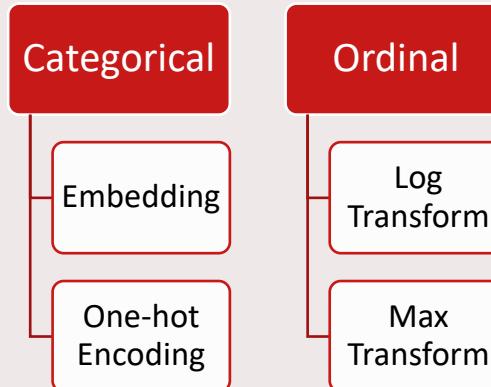


Baseline Architecture

Pre-processing and Feature Engineering



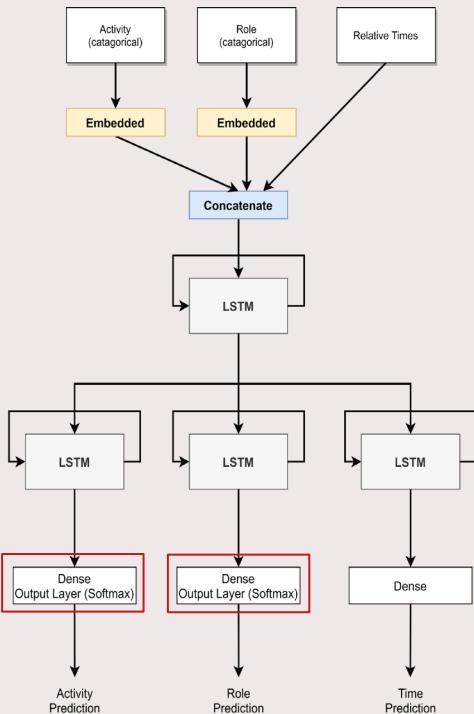
Feature Group	Attribute Name	Category
Mandatory	Activity	Categorical
	Role	Categorical
	Timestamp	Ordinal
Inter-Case	Open Cases	Ordinal
Intra-Case	Leucocytes	Ordinal
	CRP	Ordinal
	Lactic Acid	Ordinal
	Diagnose	Categorical
	Age	Ordinal
Timestamp	Weekday	Ordinal
	Daytime	Ordinal



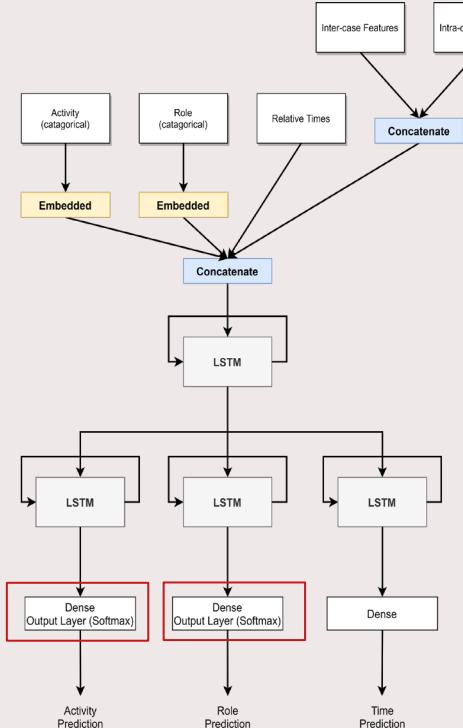
N-Gram Sequence of Size 5

Time Step	Activities
0	[0 0 0 0 0]
1	[0 0 0 0 4]
2	[0 0 0 4 6]
3	[0 0 4 6 5]
4	[0 4 6 5 10]
5	[4 6 5 10 3]

Model Structure and Quantitative Evaluation



Camargo et al. Full shared LSTM Architecture
(Base Model)



Improved LSTM Architecture

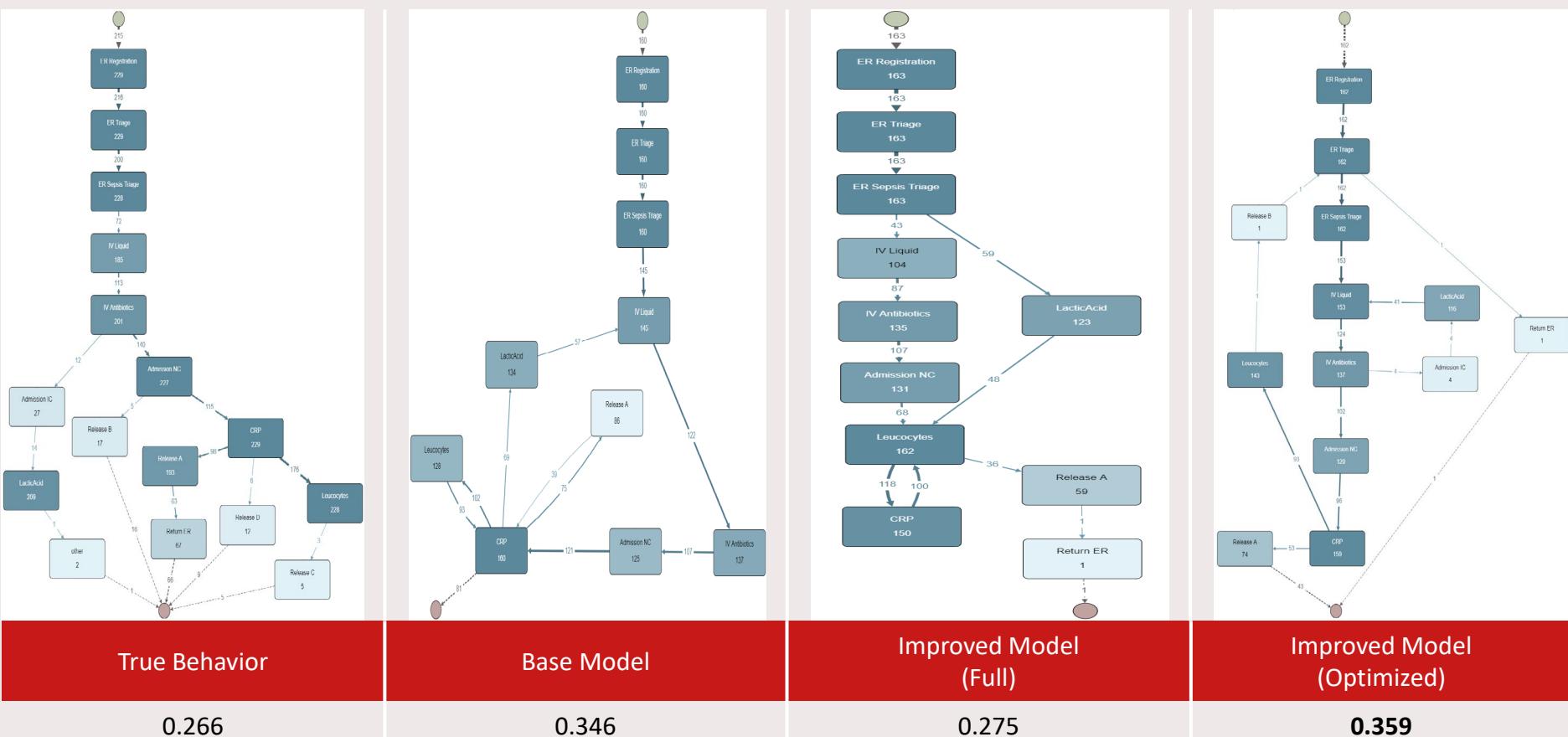
Measure Used

- Accuracy
- Damerau-Levenshtein (DL)
- Mean Absolute Error metric (MAE)

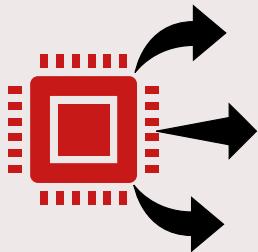
Measure	Base Model	Improved Model (Full)	Improved Model (Optimized)*
Activity Accuracy	0.5392	0.5211	0.5763
Role Accuracy	0.7937	0.8046	0.7990
DL Activity	0.3601	0.2636	0.4706
DL Role	0.8564	0.8247	0.8558
MAE (Days)	2.8624	1.6792	2.6583

* Model encoded with the Open Cases, Leucocytes, and Age

Model Qualitative Evaluation



Dashboard Functional Requirements



Multi-Predictive Recommendation

- ✓ Confidence of Recommendation
- ✓ Capping of Recommendation

```
if |activity| > |role|:  
    x <- |role|  
else:  
    x <- | activity |  
Number of Recommendations <- x
```



Batch Processing

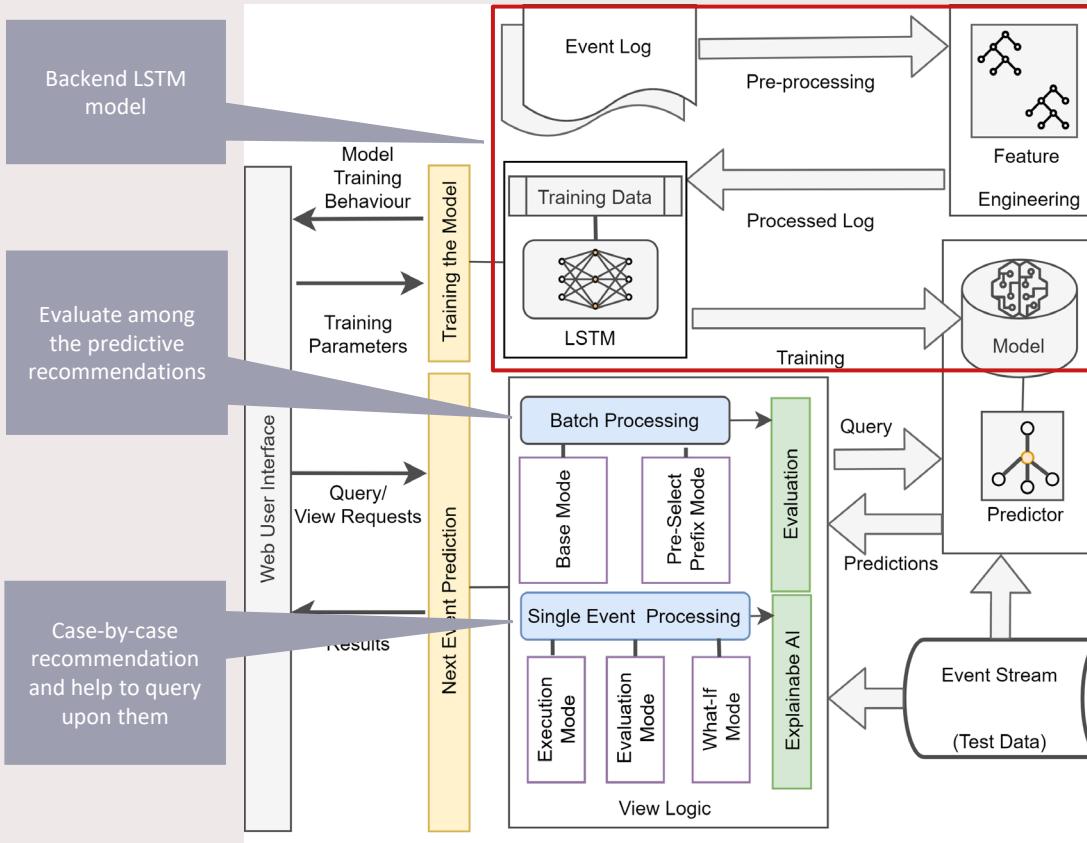
- ✓ Processing of entire test event log
- ✓ Evaluation Metric
 - Log level Similarity
 - Event level Similarity



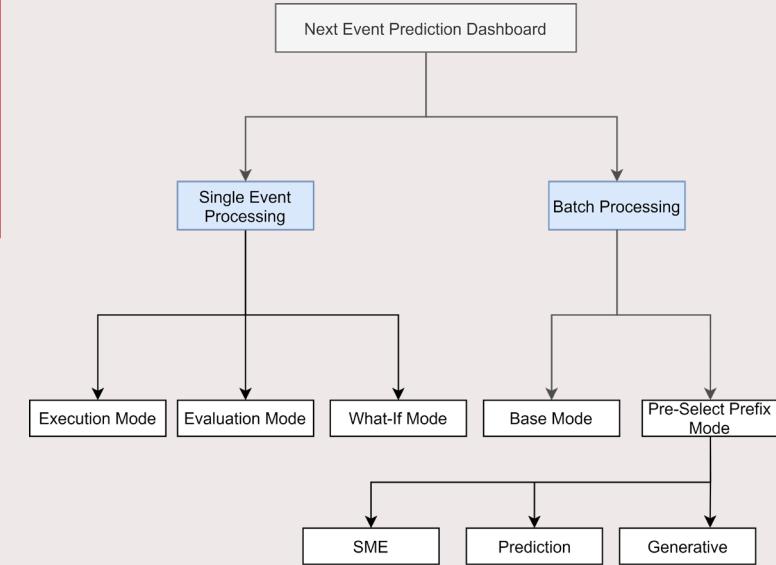
Single Event Processing

- ✓ Case-by-case recommendation for running case
- ✓ Support for query over recommendations
- ❑ Historical actions taken so far
- ❑ Contextual Information

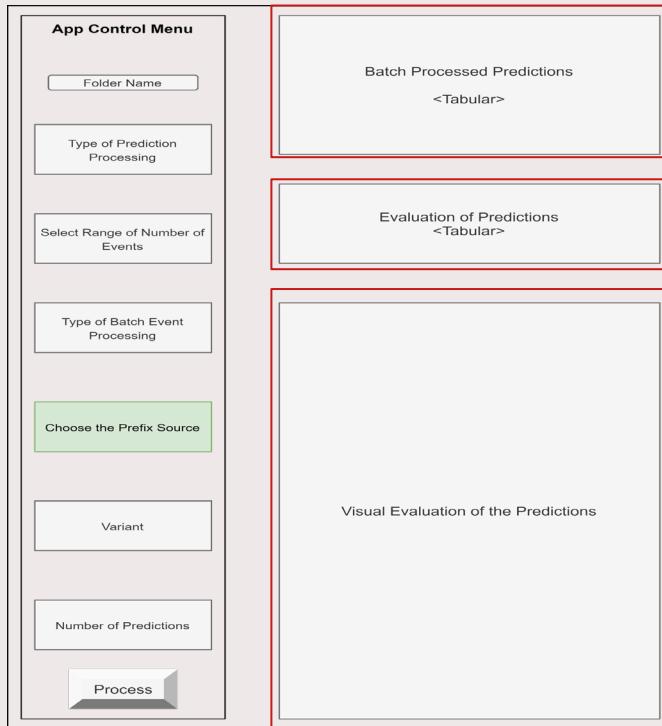
Dashboard Architecture Overview



Dashboard Navigation



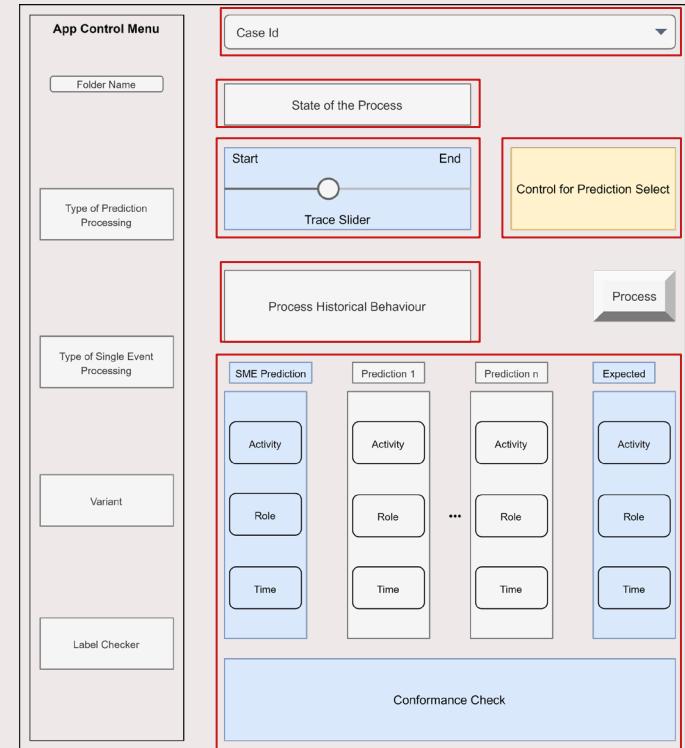
Conceptual UI Design



Batch processing Conceptual UI

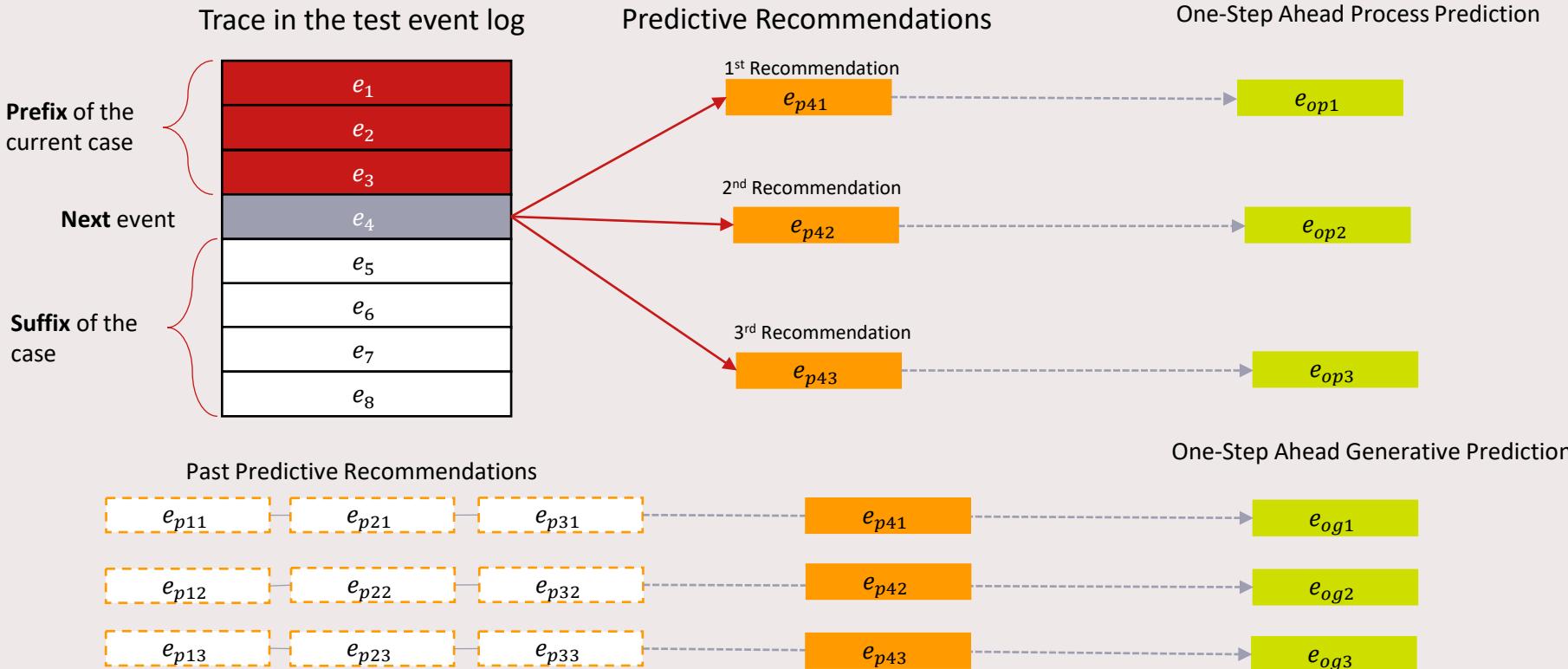


Satisfaction



Single Event processing Conceptual UI

Single Event Processing – Execution Mode Design



Single Event Processing – Execution Mode UI

Next Event Prediction Dashboard

Select Case ID: DHA

State of the Process

Open_Cases	COP	Age	weekday	LacticAcid	Leucocytes	Diagnose
91,0000	0,0000	65,0000	5	0,0600	0,0000	Other

Process

Process Historical Behaviour

Activity	Role	Time
ER Registration	Role 2	0,0000
ER Triage	Role 5	751,0000

Predictive Recommendations

Predicted Time Duration of Predictions

Expected	2
63,562,0000	

1st Prediction

Activity

Predicted	Confidence
ER Sepsis Triage	79,1746

2nd Prediction

Activity

Predicted	Confidence
Leucocytes	8,4191

3rd Prediction

Activity

Predicted	Confidence
CR	7,1879

Role

Predicted	Confidence
Role 2	74,9313

Role

Predicted	Confidence
Role 4	24,3813

Role

Predicted	Confidence
Role 1	0,0392

Label

Not Decided

Label

Not Decided

Label

Not Decided

One Step Ahead Predictions

Prediction deals with taking all the process executed so far with the respective prediction as input to the model. Generative deals with what would have happened if the respective prediction has been selected continuously.

1st Prediction Historical Behaviour

ACTIVITY	ROLE	TIME
ER Registration	Role 2	0,0000
ER Triage	Role 5	751,0000
ER Sepsis Triage	Role 2	63,562,0000

1st Prediction

Activity

Predicted	Confidence
Leucocytes	45,0281

Role

Predicted	Confidence
Role 4	0,04912

Time

Predicted	Time
Predicted	0,200,0000

Label

Regular

2nd Prediction Historical Behaviour

ACTIVITY	ROLE	TIME
ER Registration	Role 2	0,0000
ER Triage	Role 5	751,0000
Leucocytes	Role 4	63,562,0000

2nd Prediction

Activity

Predicted	Confidence
Leucocytes	45,0195

Role

Predicted	Confidence
Role 4	0,04919

Time

Predicted	Time
Predicted	0,200,0000

Label

Regular

3rd Prediction Historical Behaviour

ACTIVITY	ROLE	TIME
ER Registration	Role 2	0,0000
ER Triage	Role 5	751,0000
CR	Role 1	63,562,0000

3rd Prediction

Activity

Predicted	Confidence
Leucocytes	50,0394

Role

Predicted	Confidence
Role 4	0,05027

Time

Predicted	Time
Predicted	100,0000

Label

Regular

1st Generative Historical Behaviour

ACTIVITY	ROLE	TIME
ER Registration	Role 2	100,021,00000
ER Triage	Role 5	63,778,00000
ER Sepsis Triage	Role 2	63,562,00000

1st Prediction

Activity

Predicted	Confidence
Leucocytes	45,0151

Role

Predicted	Confidence
Role 4	0,04919

Time

Predicted	Time
Predicted	0,200,0000

Label

Regular

2nd Generative Historical Behaviour

ACTIVITY	ROLE	TIME
ER Registration	Role 2	100,021,00000
ER Triage	Role 5	63,778,00000
Leucocytes	Role 4	63,562,00000

2nd Prediction

Activity

Predicted	Confidence
ER Sepsis Triage	52,2326

Role

Predicted	Confidence
Role 2	0,04926

Time

Predicted	Time
Predicted	41,667,00000

Label

Regular

3rd Generative Historical Behaviour

ACTIVITY	ROLE	TIME
CR	Role 5	100,021,00000
IV Liquid	Role 2	63,778,00000
CR	Role 3	63,562,00000

3rd Prediction

Activity

Predicted	Confidence
CR Sepsis Triage	71,0100

Role

Predicted	Confidence
Role 2	0,05002

Time

Predicted	Time
Predicted	79,518,00000

Label

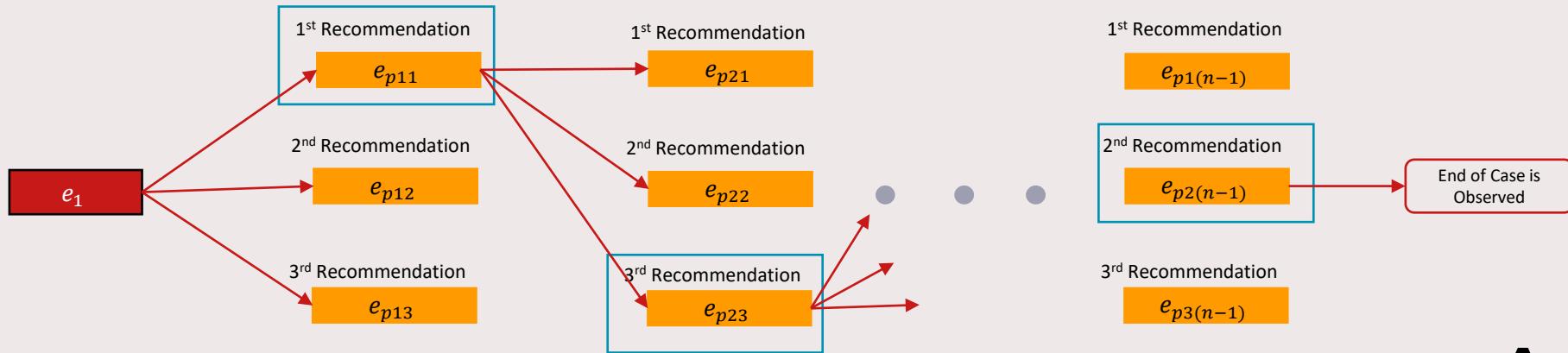
Regular

One-Step Ahead Process Prediction

One-Step Ahead Generative Prediction

Single Event Processing – What-If Mode Design

- ❖ Predictive Recommendations are constrained → Prefixes are grouped
- ❖ Classical process mining What-If technique → Limited to Checking the performance



Scrutability

Single Event Processing – What-If Mode UI

Next Event Prediction Dashboard

Select Case ID
FG

State of the Process

weekday	CRP	LacticAcid	Leucocytes	open_cases	Diagnose	Age
0	0.0000	0.0000	0.0000	62.0000	H	85.0000

What-IF Prediction Choose Box

Choose the Prediction according to which System generates the next prediction, **SME** (Subject Matter Expert): decision solely based on users instinct and knowledge of business process about the process, **Prediction n**: decision solely based on the respective ranked confidence of the process

SME Prediction 1 Prediction 2 Prediction 3 Prediction 4 Prediction 5

Process

Process Historical Behaviour

Activity	Role	Time	Choice
0 ER Registration	Role 2	108,024.0000	Prediction 1

Predicted Time Duration of Predictions

Expected	231,030.0000
----------	--------------

Predictive Recommendations

1st Prediction	2nd Prediction	3rd Prediction	4th Prediction	5th Prediction
Activity Predicted: 1 ER Registration Confidence: 51.3465	Activity Predicted: 1 Leucocytes Confidence: 25.2964	Activity Predicted: 1 CRP Confidence: 12.9250	Activity Predicted: 1 LacticAcid Confidence: 5.9508	Activity Predicted: 1 ER Triage Confidence: 2.4617
Role Predicted: Role 2 Confidence: 71.3370	Role Predicted: Role 4 Confidence: 27.1493	Role Predicted: Role 5 Confidence: 1.3919	Role Predicted: Role 1 Confidence: 0.0538	Role Predicted: Role 3 Confidence: 0.0258
Label Not Decided	Label Not Decided	Label Not Decided	Label Not Decided	Label Not Decided

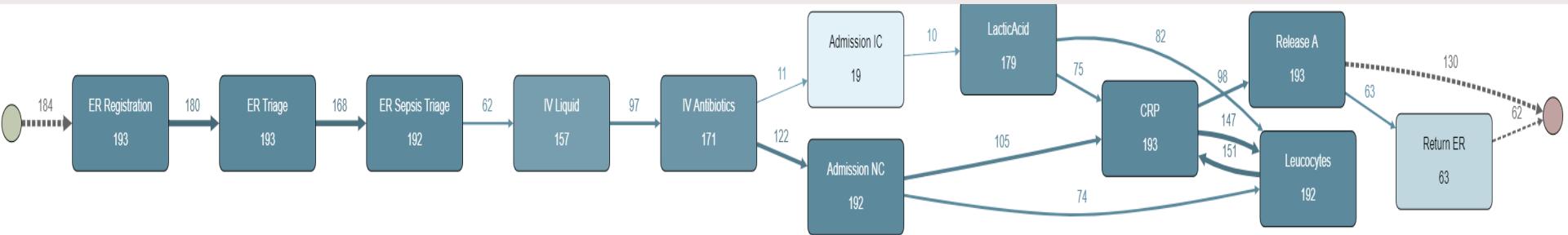
TU/e

Single Event Processing – Execution Mode Objective

- ✓ If the dashboard's predicted recommendation includes the user's next event, we may infer that the dashboard was helpful in this case.
- ✓ If the user's next event isn't on the dashboard then the dashboard recommendations weren't useful.

This will atleast ensure dashboard was able to give basic help and advice throughout the process execution.

Most likely Process Behaviour of Sepsis Cases looks like :



Single Event Processing – Execution Mode Demonstration

Prefix	Predictive Recommendation				One-Step ahead Process Outcome		One-Step ahead Generative Outcome	
	1 st Prediction		2 nd Prediction		1 st Prediction	2 nd Prediction	1 st Prediction	2 nd Prediction
	Pred.	Conf.	Pred.	Conf.				
Start	ER Registration	95.6911	Leucocytes	1.6362	ER Triage	ER Registration	ER Triage	ER Registration



Effectiveness



Efficiency

Quantitative Evaluation among Predictive Recommendations

Similarity Metric Used

Event Level Similarity

Damerau-Levenshtein (DL)

Mean Absolute Error metric (MAE)

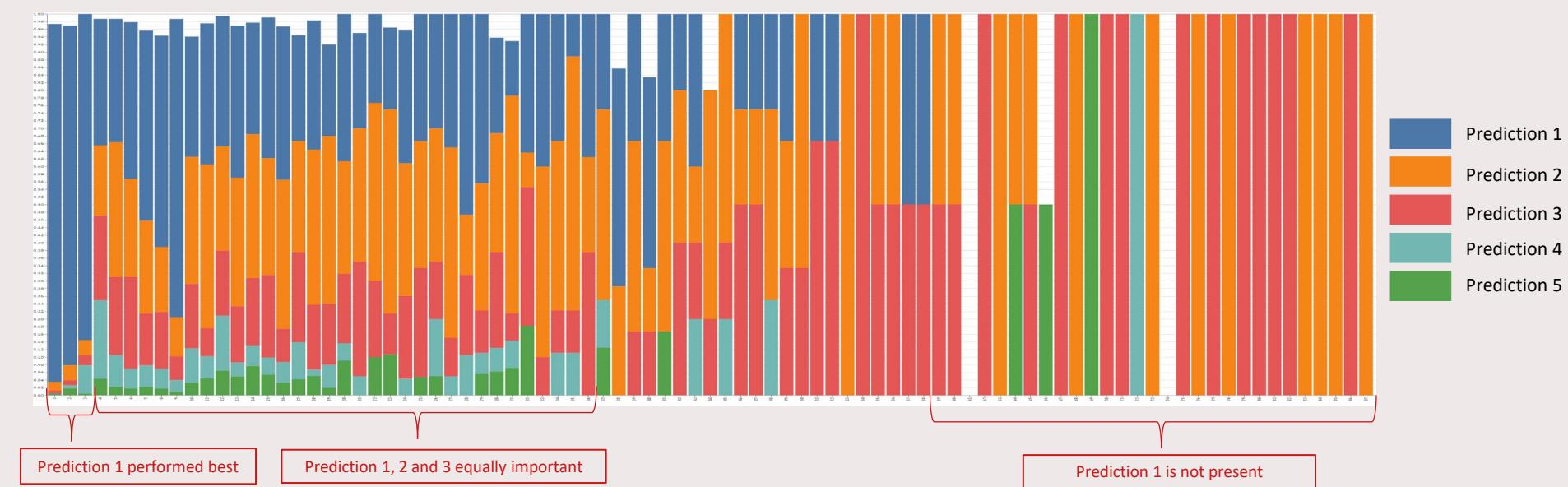
Log Level Similarity :

Control-Flow Log Similarity (CFLS) : DL + Hungarian Algorithm

Event Log Similarity (ELS) : Business Process Trace Distance (BPTD) + Hungarian Algorithm

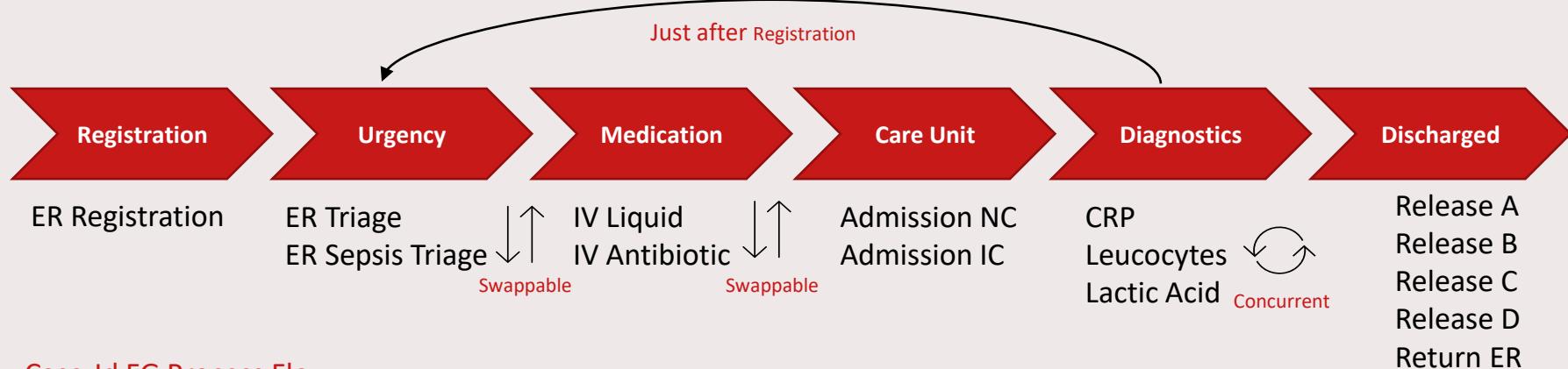
Pred. Number	Activity Similarity (DL)	Role Similarity (DL)	Control-Flow Log Similarity	Event Log Similarity
Prediction 1	0.2473	0.8226	0.5933	0.5621
Prediction 2	0.3716	0.1212	0.4934	0.4779
Prediction 3	0.2329	0.0416	0.4154	0.3945
Prediction 4	0.0509	0.0082	0.3346	0.3228
Prediction 5	0.0477	0.0004	0.3212	0.3150

Similarity comparison among Predictive Recommendations for each Event Number

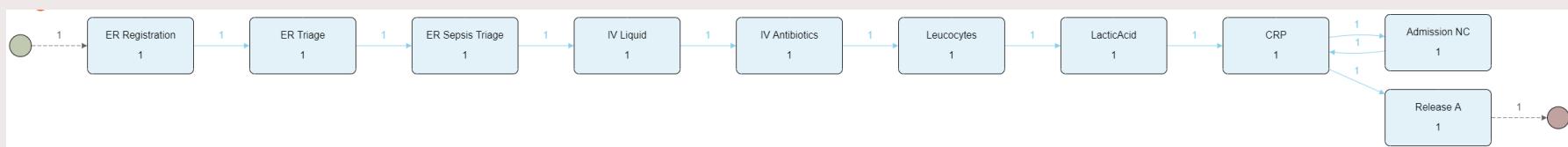


Single Event Processing – What-If Mode Objective

A-priori Knowledge about the Process



Case-Id FG Process Flow



- ✓ If changing the order of sequence, where diagnostic actions are performed first followed by analysing the urgency.
 - ☐ If it causes the case to converge to Release A, it will be considered successful

Single Event Processing – What-If Mode Demonstration

Case ID FG Activity	Predictive Recommendation									
	1 st Prediction		2 nd Prediction		3 rd Prediction		4 th Prediction		5 th Prediction	
	Pred.	Conf.	Pred.	Conf.	Pred.	Conf.	Pred.	Conf.	Pred.	Conf.
Start	ER Registration	95.6911	Leucocytes	1.6362	CRP	0.9738	LacticAcid	0.3258	ER Sepsis Triage	0.3225

Single Event Processing – Other Features

Evaluation Mode

Requirements

- ✓ Simulate until end of case over executed cases
 - Allow altering prefixes : **Pre-select Prefix**
- ✓ Let's to select pre-select prefix for each case
- ✓ Analyze how case will end with and without assistance
- ✓ Does conformance check on the pre-select prefix



Trust on the Model



Debug

Label Checker

Requirements

Labelling with Deviant and Regular

The runtime labelling can be considered to use as KPI's for the decision-making.

For Sepsis Cases

- ✓ Patient is re-admitted to the emergency department (Return ER).
- ✓ Patient is admitted to intensive care (Admission IC).
- ✓ Patient is released from the hospital for a Reason besides Release A.

Logic

1. User Set after how many events the check should start
2. At run-time dashboard start checking the label selected

Answering the Research Questions

Integrate domain knowledge in Predictive Model without sacrificing accuracy

Answered by integrating inter and intra case features in the LSTM models

Dashboard be designed for Business Users specifically Domain Experts and solving current gaps

Answered by designing the dashboard which provide contextual, historical, alternative recommendations, providing reliability of the recommendation, and user defined labels.

Translate and Explain the recommendations to Domain Expert

Answered by demonstrating the Single Event processing mode from the user perspective with the business knowledge comprehended from just process model

Conclusion

- ❑ A framework for the predictive process monitoring dashboard for domain experts, which could explain the prediction outcome and provide an alternative recommendation.
 - ✓ Provide enough information contextually and about process state.
 - ✓ Let's the user to build trust on the model by replaying on executed case by changing the prefix
 - Helps to determine how the model will behave
 - ✓ Let's the user to query over the model by choosing the predictions
 - ✓ Capable of including the other event log.
 - ✓ Framework is also capable to be deployed on running case.

Limitations and Future Research

- Demand for more accurate predictive model which can incorporate contextual information.
- Labelling of recommendations are user defined, although it is programmatic. It can be enhanced using supervised machine learning models.
- Dashboard design is only capable of handing one case at a time, which is not the case in the real business setting. Thus the dashboard design needs a mechanism to save the case state and execute on them as it progress.
- What-if mode recommendations are grouped which can be made modular.
- Multi model recommendation might make the framework to be used by users with little experience with business knowledge.
- Predictive technique algorithm could be improved for more swiftness in the performance.
- Evaluation of the dashboard could be done on a event log whose business knowledge is well documented.

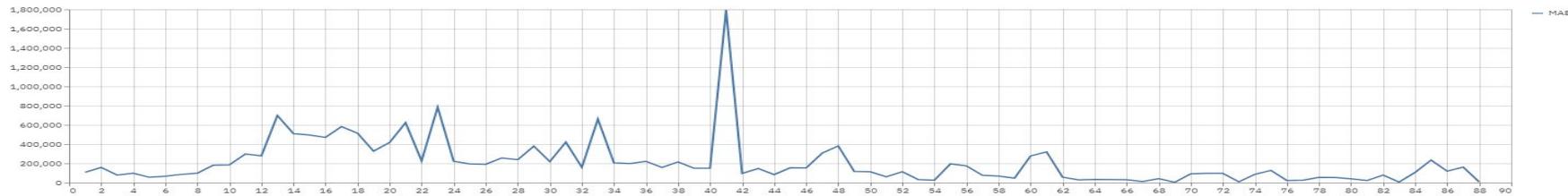


Dashboard Hosted Online

Questions?

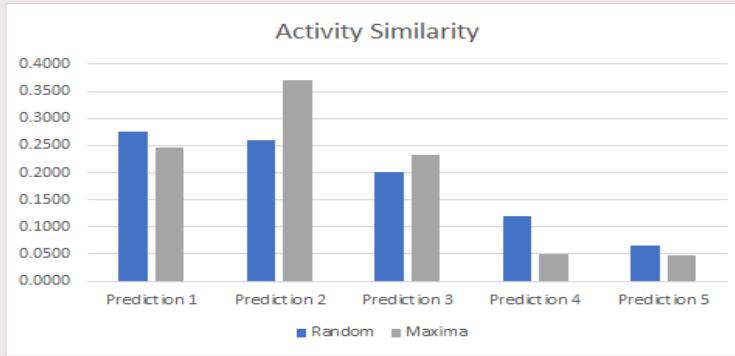
Thank You

MAE Replaying Historical Events

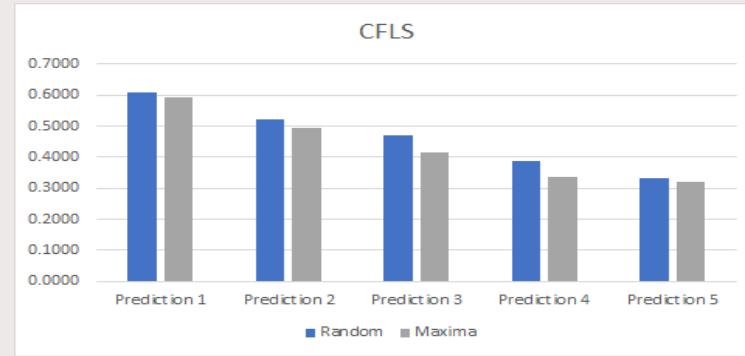


Predictive Recommendations Random vs Maxima

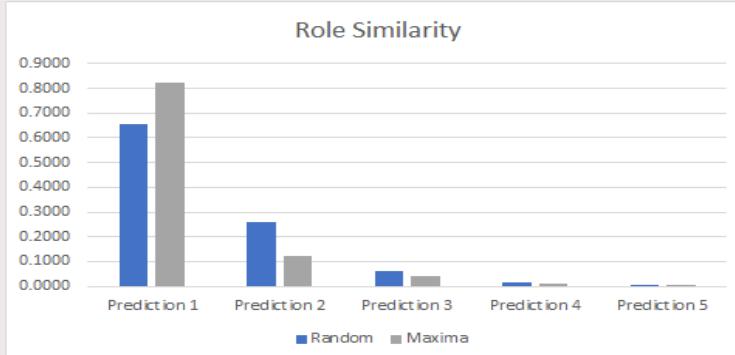
Event Level Similarity



Log Level Similarity



Role Similarity



ELS

