PMAI workshop

Predictive and proactive trustable process monitoring

Chiara Ghidini
Process & Data Intelligence
Fondazione Bruno Kessler

The agenda

- From data-centered data to event-entered data
- The multiple dimensions of Predictive Process Monitoring
- Examples of Predictive Process Monitoring works
- The quest for explainable predictions
- Beyond Prediction: Action

The contributors

Chiara
Di Francescomarino

Chiara Ghidini

Williams Rizzi

Massimiliano Ronzani

Giulio Petrucci

Luca Simonetto

Marco Federici

Andrei Buliga

Stefano Branchi

Marlon Dumas

Anna Leontjeva

Irene Teinemaa

Fredrik Milani

Arik Senderovich

Anton Yeshchenko

Kerwin Jorbina

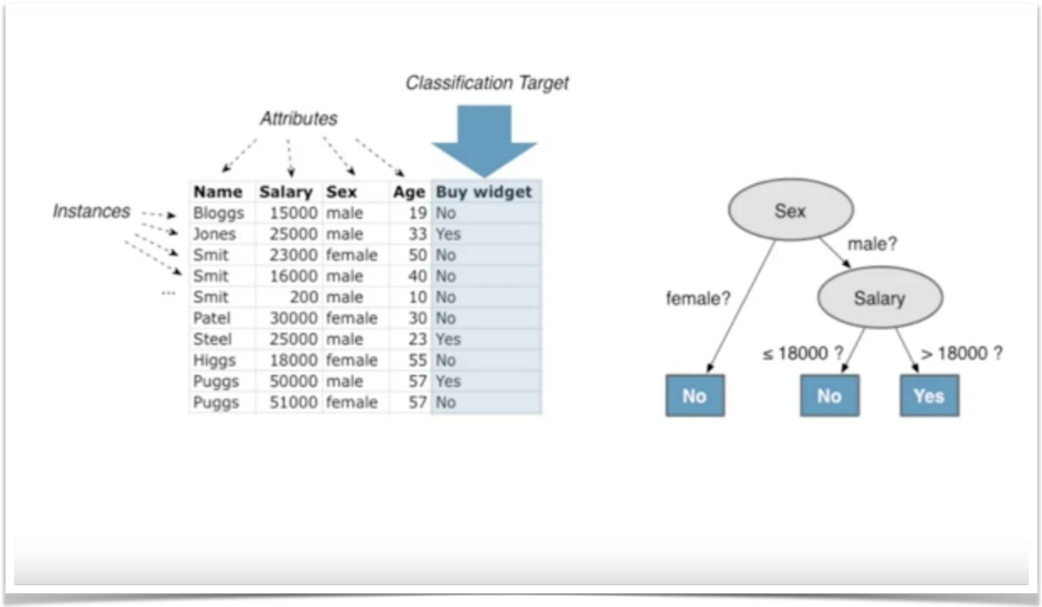
Fabrizio Maggi

David Massimo

Francesco Ricci

The typical data (mining) view

- Data centered around the notion of object
- Objects identified by (unique) IDs



Credits: Anne Rozinat

The typical process (mining) view

Everything starts from an execution trace

Case ID	Case101	Case101	Case101	Case101	Case101	
Event	Park car	Open door	Enter flat	Turn Light on	Turn Radio on	TIME
Timetamp	5:00 pm	5:03 pm	5:10 pm	5:11 pm	5:15 pm	

- Data centered around the notion of ``story”
- Execution traces associated to a unique ID

The typical process (mining) view

Everything starts from an execution trace

Case ID	Case101	Case101	Case101	Case101	Case101	
Event	Park car	Open door	Enter flat	Turn Light on	Turn Radio on	TIME
Payloads	5:00 pm	5:03 pm	5:10 pm	5:11 pm	5:15 pm	
	Car: Tesla		Address: …			

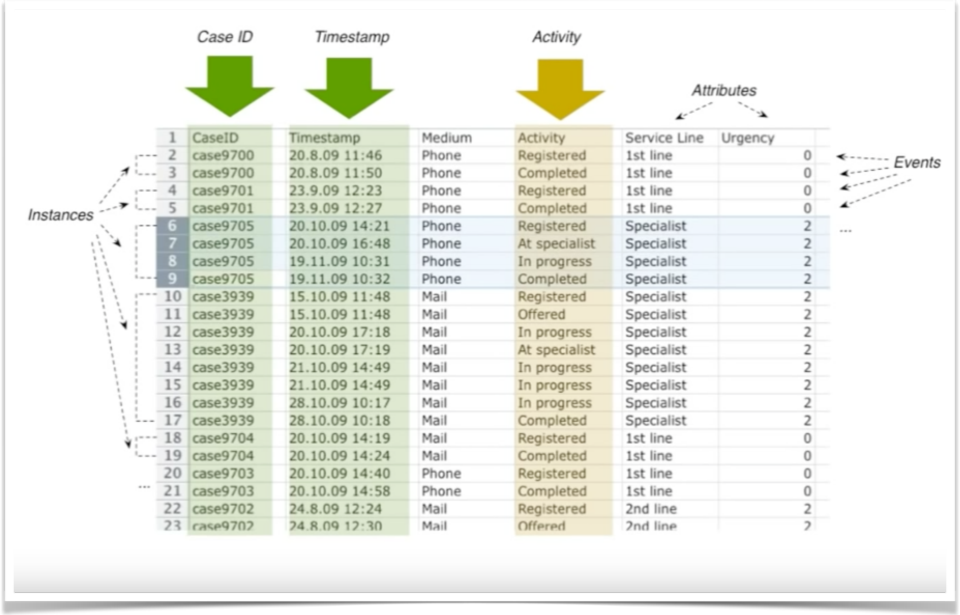
- Data centered around the notion of ``story”
- Execution traces associated to a unique ID

Multi-perspective data!

Longitude: time
Several vertical dimensions:
resources, objects, costs, …

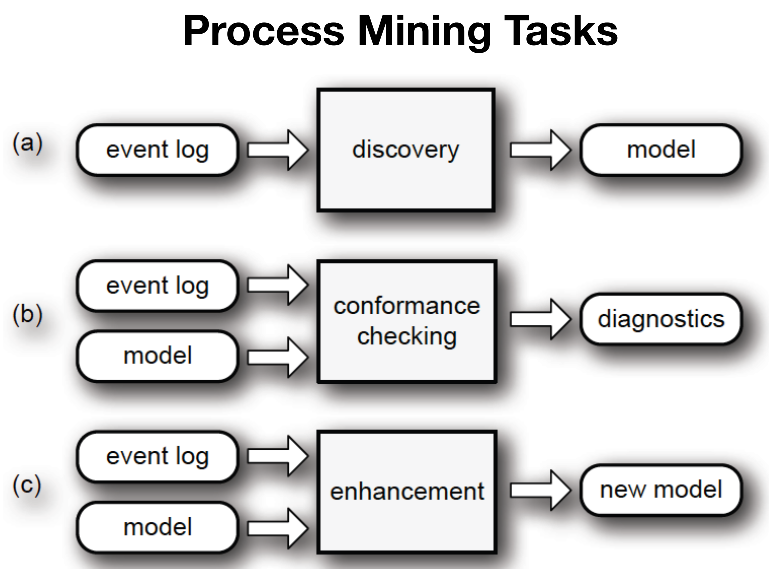
The typical process (mining) view

- Data centered around the notion of ``story”
- Execution traces associated to a unique ID
- Event Log: set of execution traces identified by unique IDs



1	CaseID	Timestamp	Medium	Activity	Service Line	Urgency
2	case9700	20.8.09 11:46	Phone	Registered	1st line	0
3	case9700	20.8.09 11:50	Phone	Completed	1st line	0
4	case9701	23.9.09 12:23	Phone	Registered	1st line	0
5	case9701	23.9.09 12:27	Phone	Completed	1st line	0
6	case9705	20.10.09 14:21	Phone	Registered	Specialist	2
7	case9705	20.10.09 16:48	Phone	At specialist	Specialist	2
8	case9705	19.11.09 10:31	Phone	In progress	Specialist	2
9	case9705	19.11.09 10:32	Phone	Completed	Specialist	2
10	case3939	15.10.09 11:48	Mail	Registered	Specialist	2
11	case3939	15.10.09 11:48	Mail	Offered	Specialist	2
12	case3939	20.10.09 17:18	Mail	In progress	Specialist	2
13	case3939	20.10.09 17:19	Mail	At specialist	Specialist	2
14	case3939	21.10.09 14:49	Mail	In progress	Specialist	2
15	case3939	21.10.09 14:49	Mail	In progress	Specialist	2
16	case3939	28.10.09 10:17	Mail	In progress	Specialist	2
17	case3939	28.10.09 10:18	Mail	Completed	Specialist	2
18	case9704	20.10.09 14:19	Mail	Registered	1st line	0
19	case9704	20.10.09 14:24	Mail	Completed	1st line	0
20	case9703	20.10.09 14:40	Phone	Registered	1st line	0
21	case9703	20.10.09 14:58	Phone	Completed	1st line	0
22	case9702	24.8.09 12:24	Mail	Registered	2nd line	2
23	case9702	24.8.09 12:30	Mail	Offered	2nd line	2

Credits: Anne Rozinat

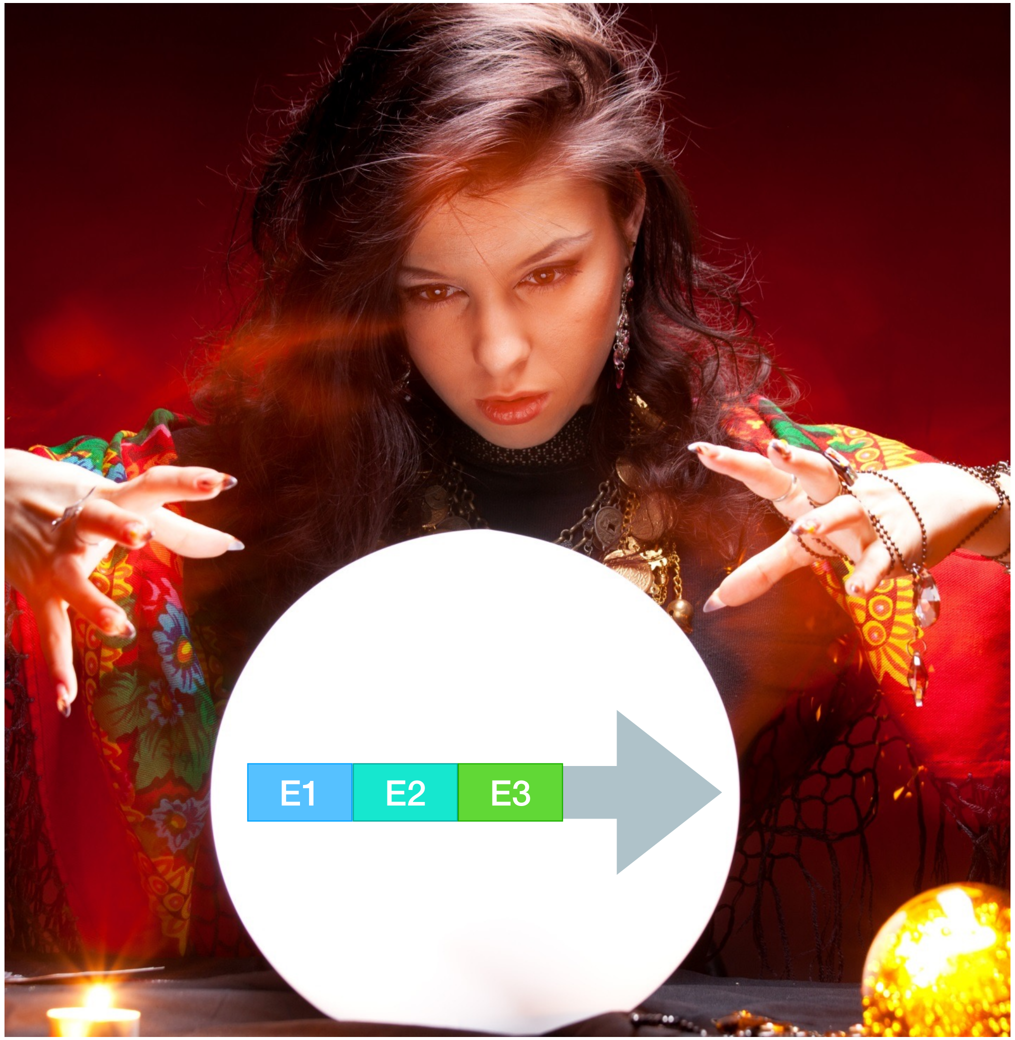


Predictive AI meets Event Logs: Predictive Process Monitoring

- Availability of data represented in a Event Log format (typically XES)
- A bounce of techniques to produce XES event logs from e.g., relational databases

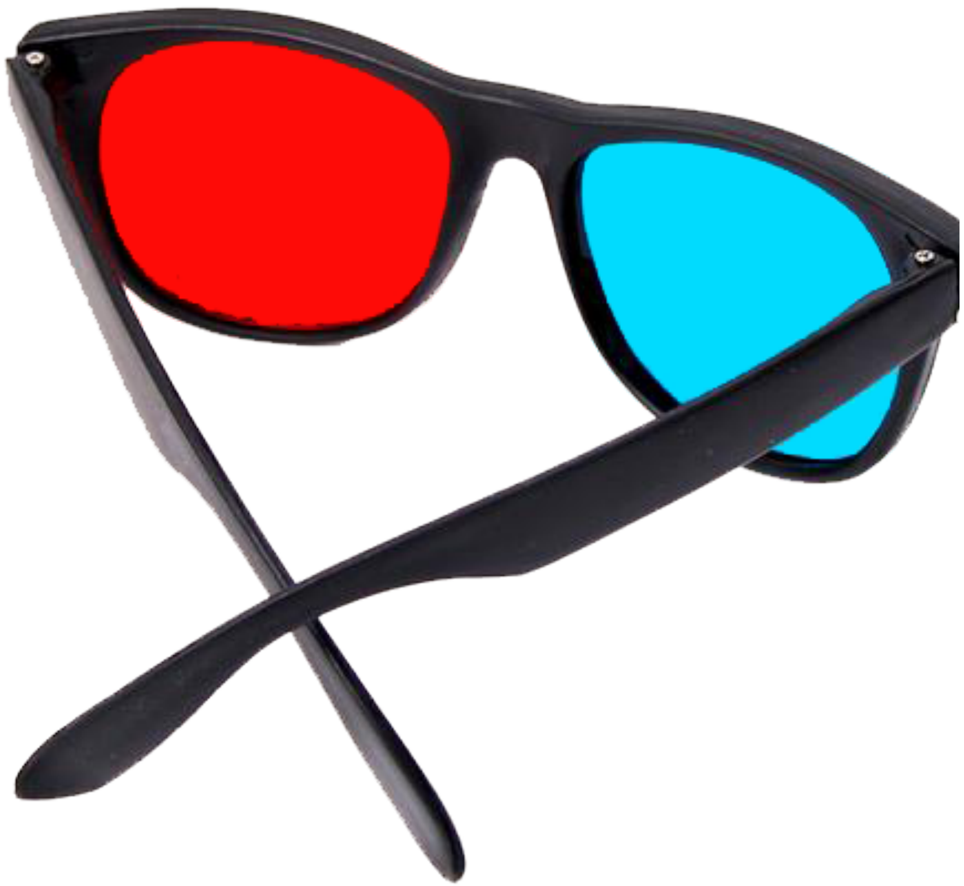
Historical Traces

Ongoing Trace

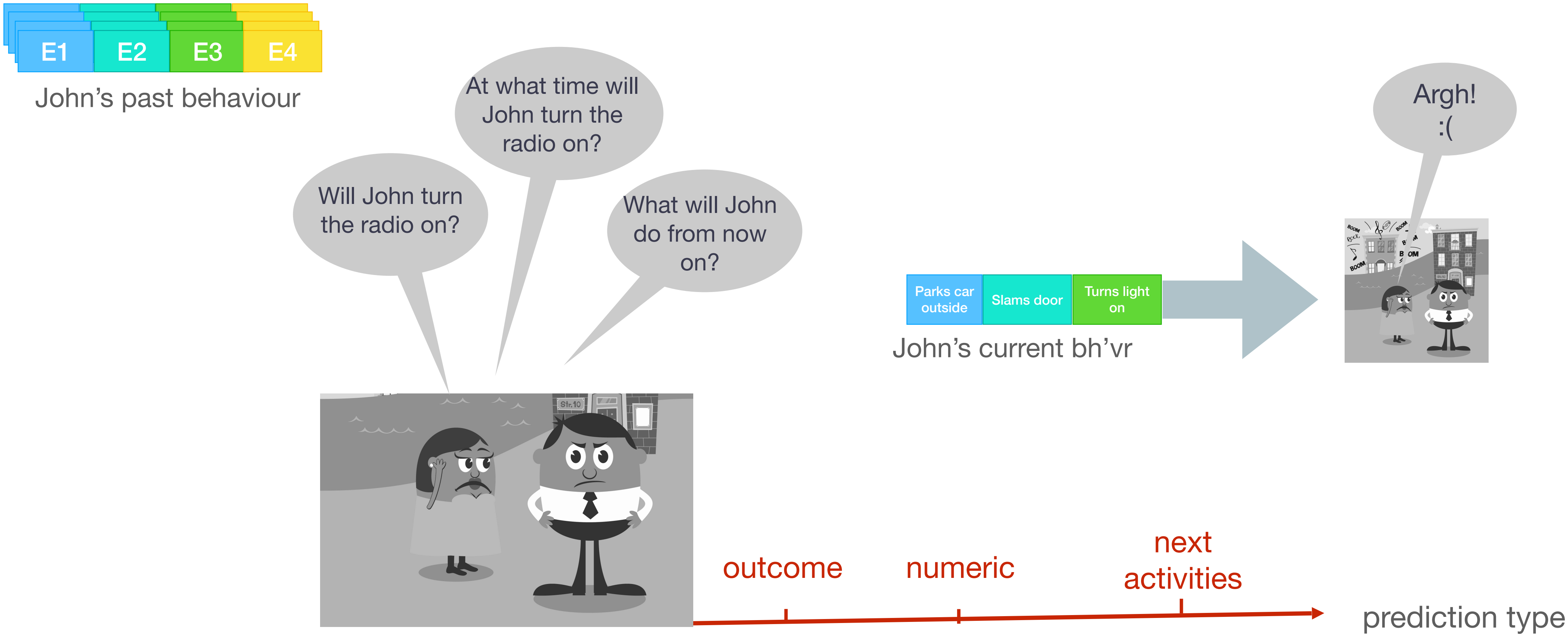


What is Predictive Process Monitoring?

A 3D Perspective



Dimension 1: what to predict



Examples of outcomes

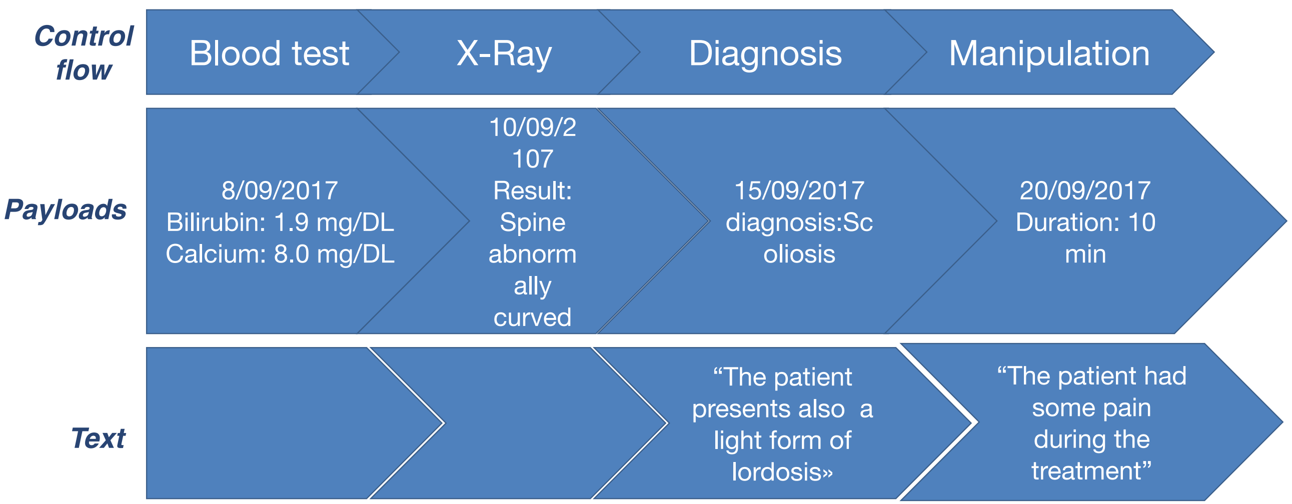
- Fast vs. Slow
- Events happening into traces
  - BPI Challenge 2011 event log about a healthcare process (treatments of patients in a Dutch hospital):

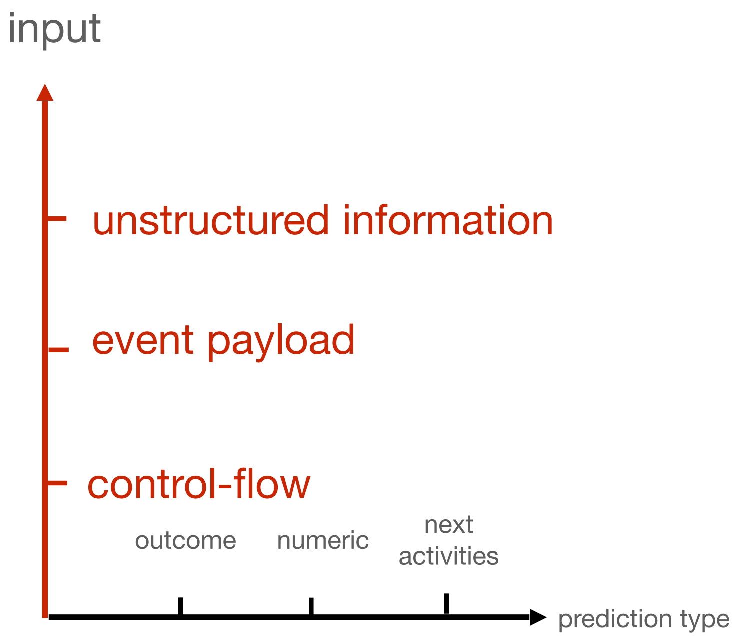
$$\phi_1 = \diamond(\text{“histological examination - big respites”})$$
$$\phi_2 = \diamond(\text{“tutor maker CA - 19.9”}) \vee \diamond(\text{“ca - 125 using meia)”})$$
$$\phi_3 = \Box(\text{“CEA - tutor marker using meta”} \rightarrow \diamond(\text{“squamous cell carcinoma using eia”}))$$
$$\phi_4 = \neg\text{“histological examination - biopsies nno”}\,\mathcal{U}\,\text{“cytology - ectocervix”}$$

- Somehow … everything you can transform into a label.

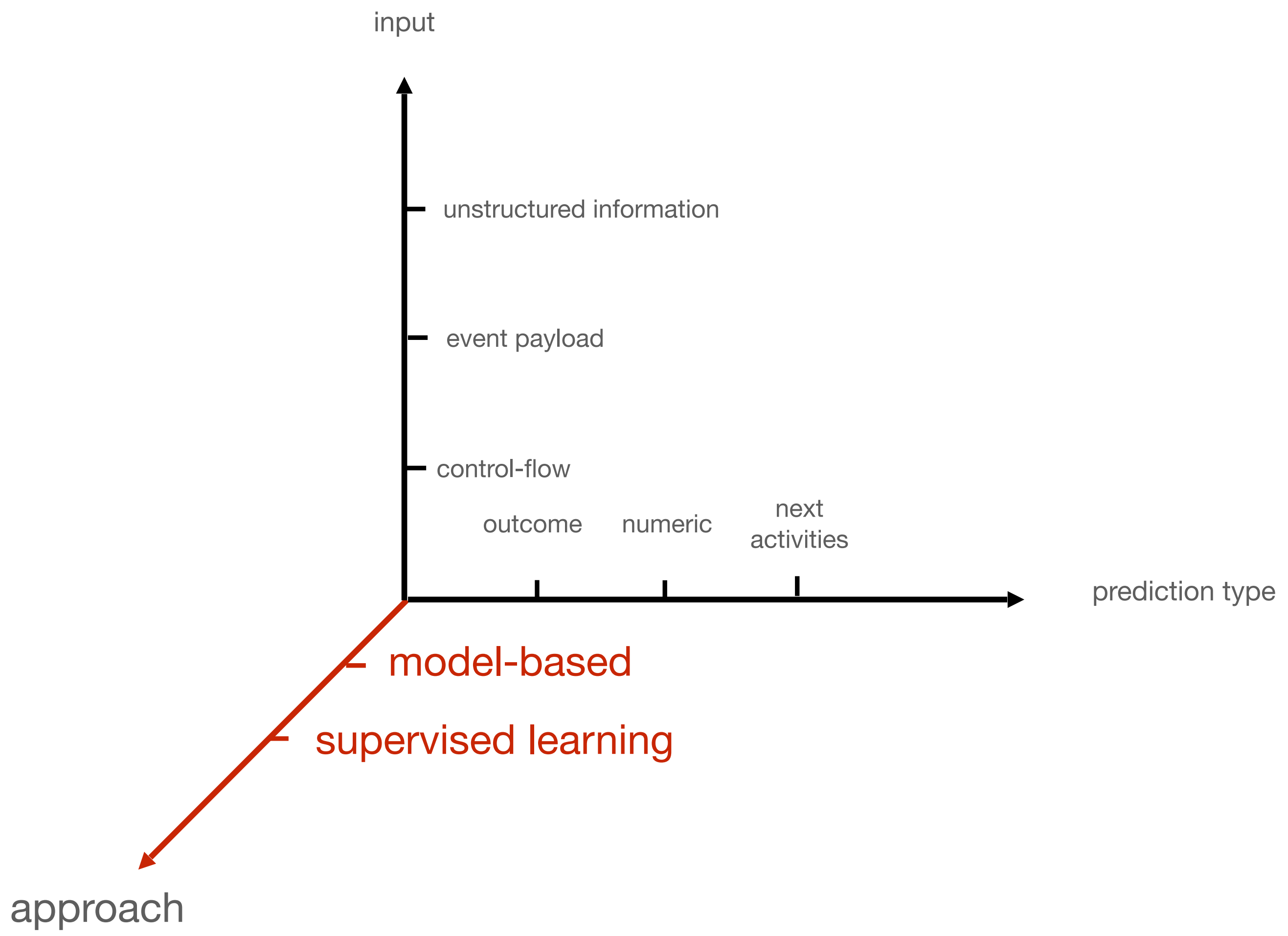
Dimension 2: leveraging what?

Patient's history (trace)

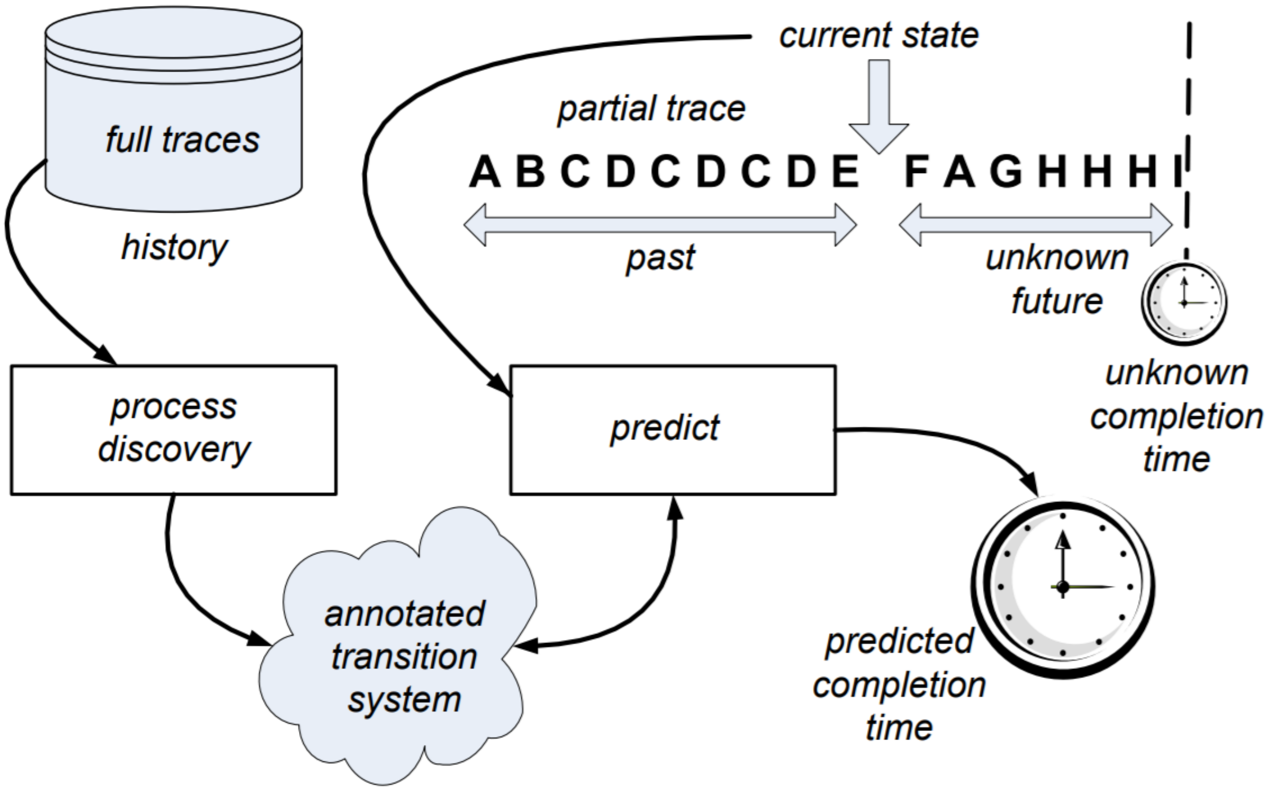




Dimension 3: which technique?

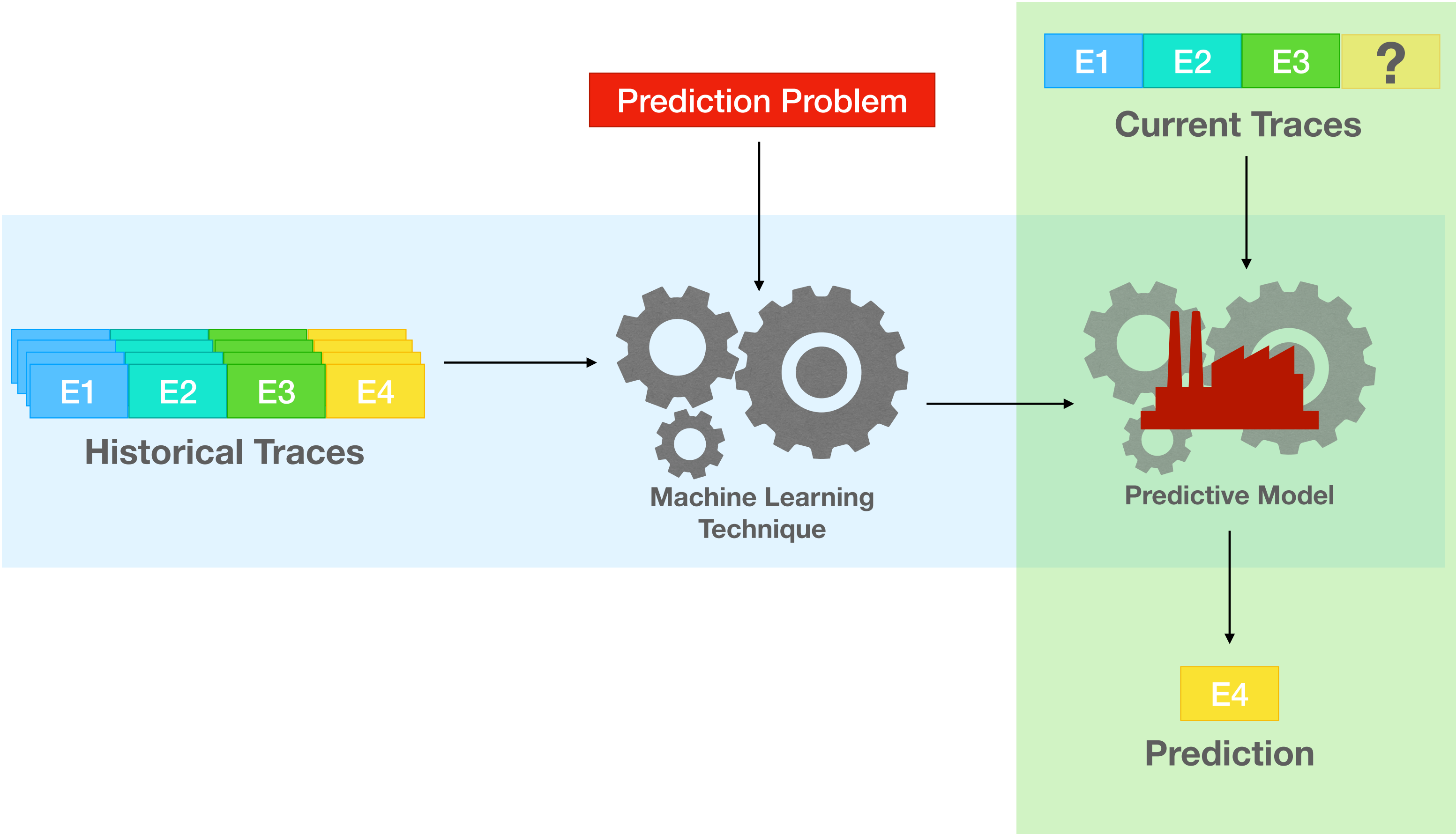


Approach 1: Model-based

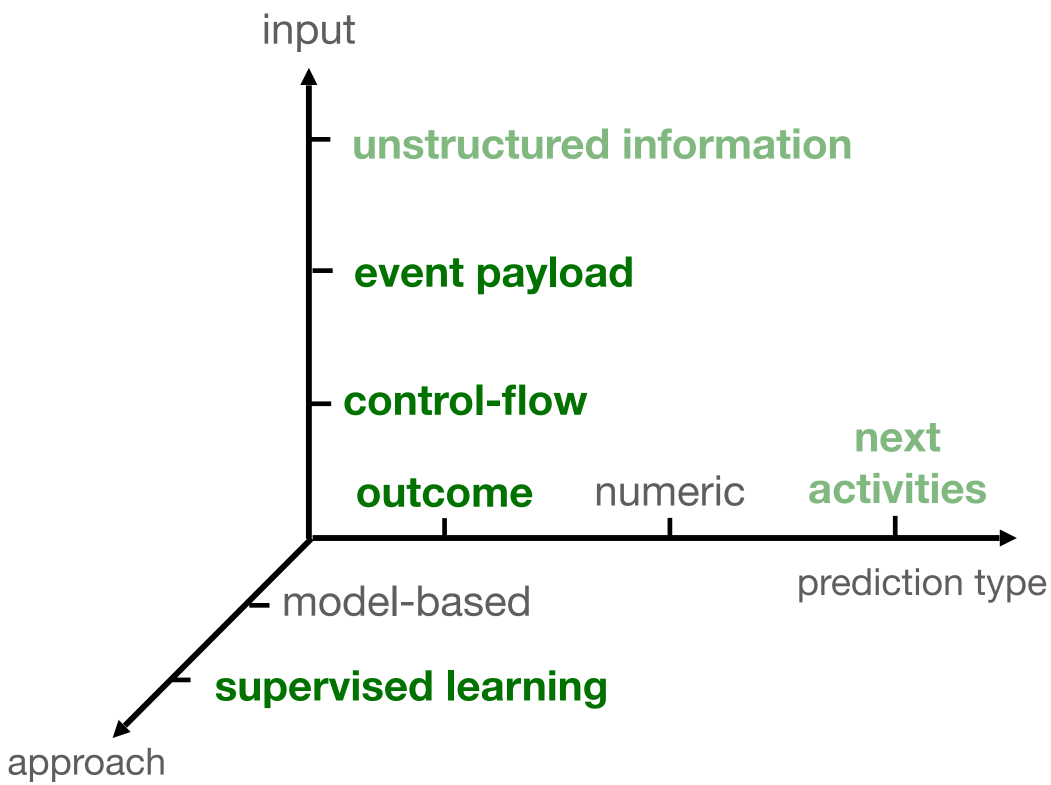


credits to W.M.P. van der Aalst

Approach 2: Supervised learning

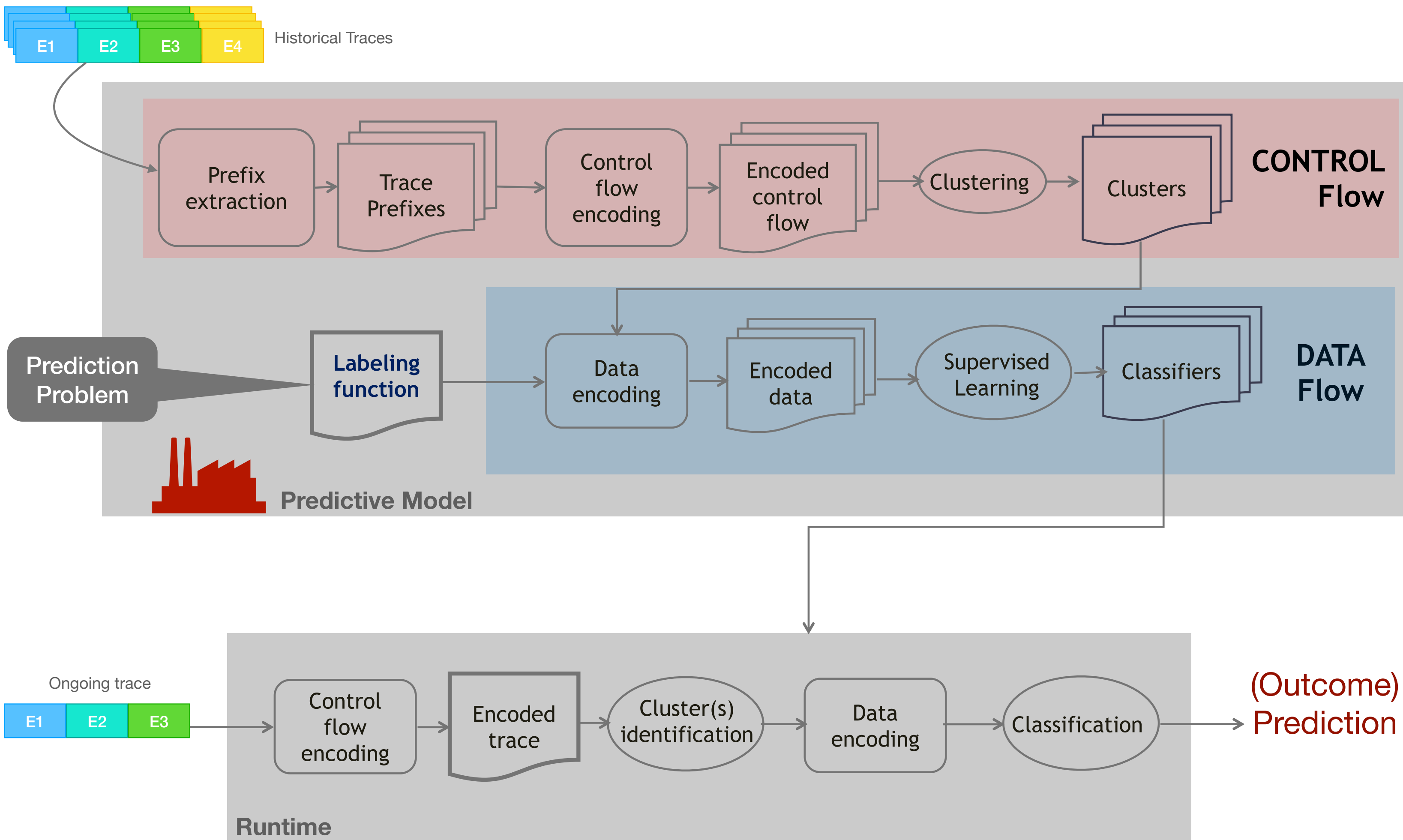


Trento in the picture!



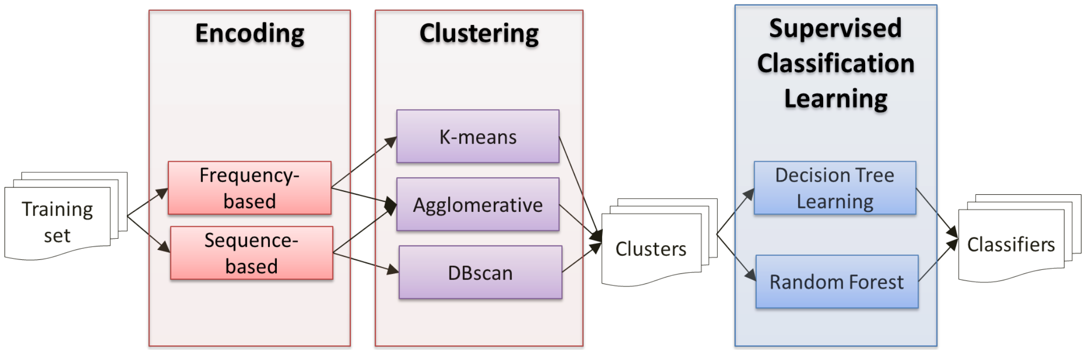
- Fabrizio Maria Maggi, Chiara Di Francescomarino, Marlon Dumas, Chiara Ghidini: **Predictive Monitoring of Business Processes**. CAiSE 2014: 457-472
- Di Francescomarino C., Dumas M., Maggi F.M., Teinemaa I., **Clustering-Based Predictive Process Monitoring**. IEEE Transactions on Services Computing (TSC). To appear.
- Anna Leontjeva, Raffaele Conforti, Chiara Di Francescomarino, Marlon Dumas, Fabrizio Maria Maggi: **Complex Symbolic Sequence Encodings for Predictive Monitoring of Business Processes**. BPM 2015: 297-313
- Irene Teinemaa, Marlon Dumas, Fabrizio Maria Maggi, Chiara Di Francescomarino: **Predictive Business Process Monitoring with Structured and Unstructured Data**. BPM 2016: 401-417

Predict with control + data flow - a first approach



Encoding… clustering… classification…

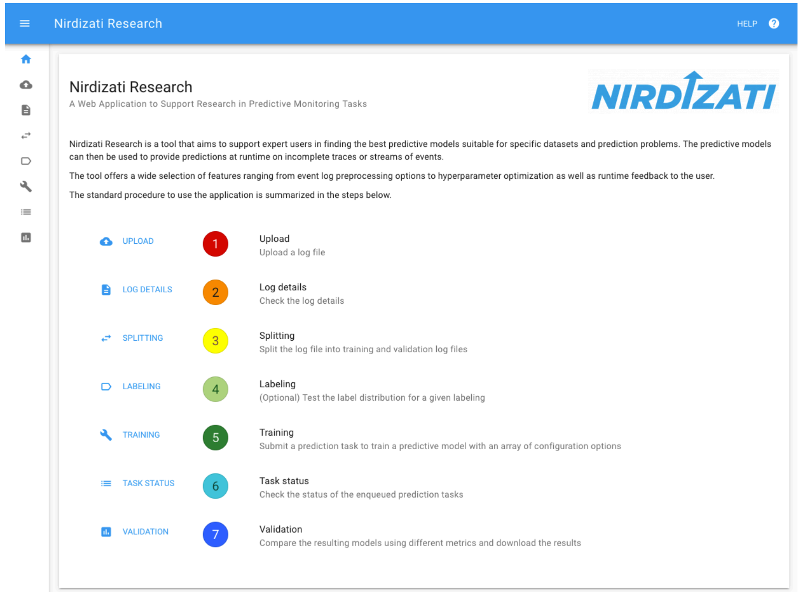
- A single single silver bullet does not exist
- Implementation of a wide set of techniques to train predictive Models and to **compare** them in effective ways
- **Investigation of several types of encodings**
  (and encompass the dichotomy clustering vs classification)



- How to encode traces?
- How to encode sequentiality?
- Can we encode events and payloads together and go beyond clustering and classification?

From research framework to a solid tool

- A single single silver bullet does not exist
- Implementation of a wide set of techniques to train predictive Models and to **compare** them in effective ways
- Hyperparameter optimisation: **Automated solution based on genetic algorithms**



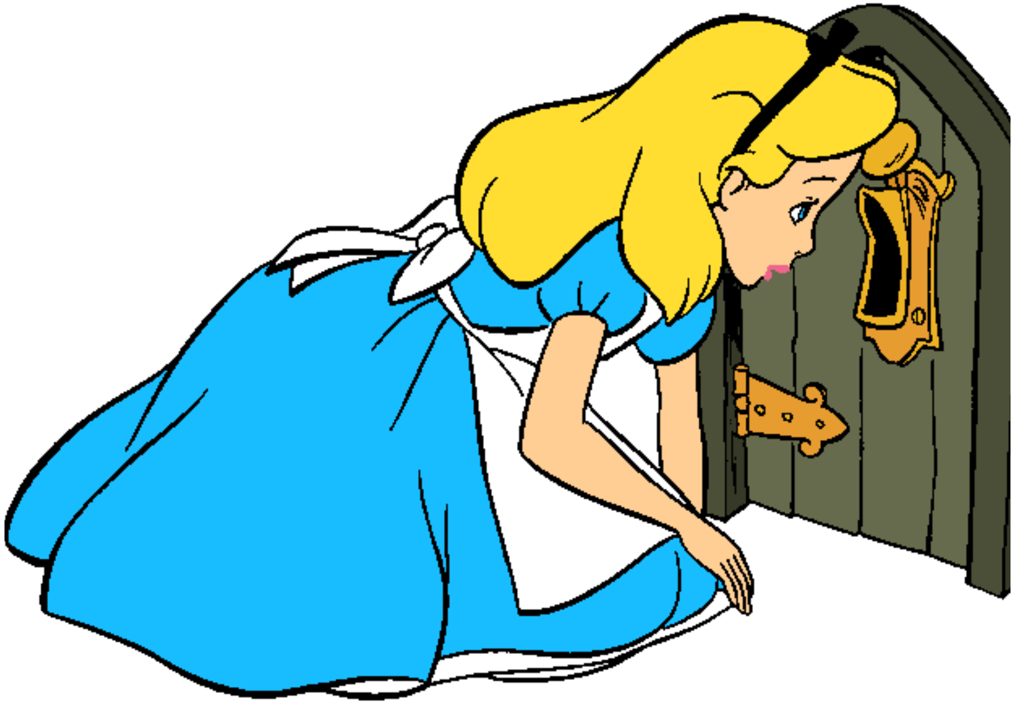
http://research.nirdizati.org

Best demo award @ BPM17

Beyond 3D

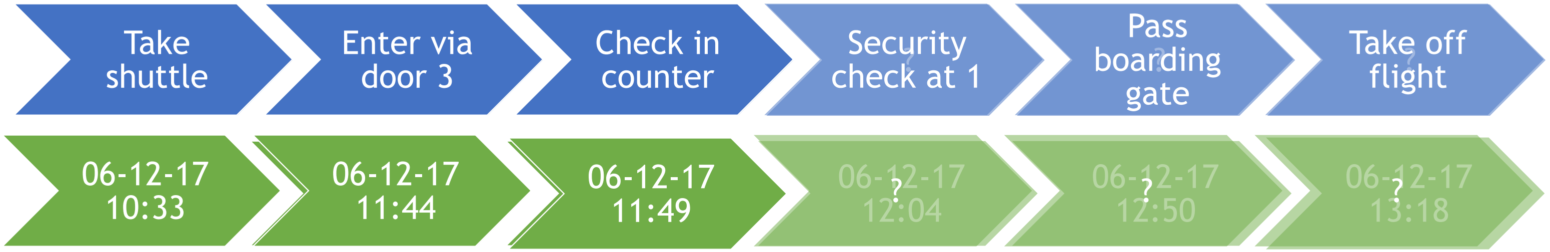
Other dimensions and values

- Predictions with a-priori knowledge
- Inter-case predictions



Chiara Di Francescomarino, Chiara Ghidini, Fabrizio Maria Maggi, Giulio Petrucci, Anton Yeshchenko: **An Eye into the Future: Leveraging A-priori Knowledge in Predictive Business Process Monitoring**. BPM 2017: 252-268

Traveller’s next activities



What if … there is a strike and we know it!

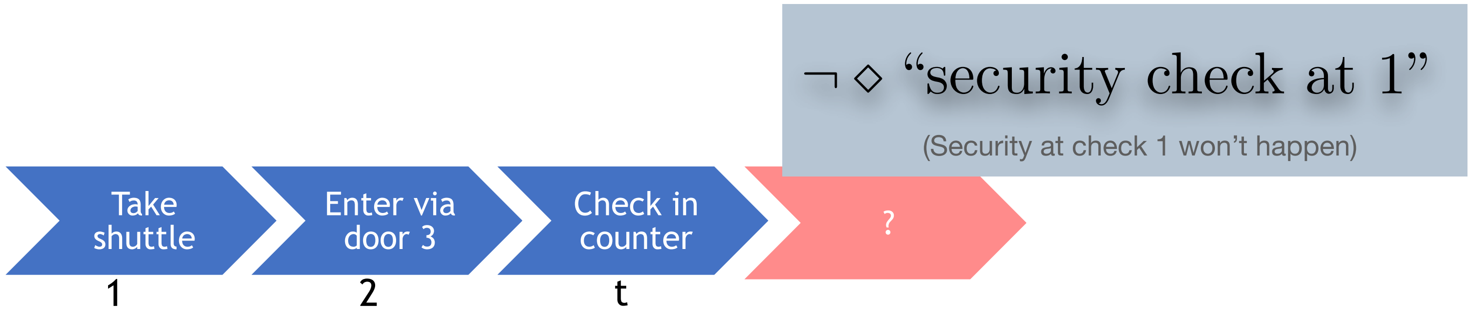
Take shuttle	Enter via door 3	Check in counter	Security check at 1	Pass boarding gate	Take off flight
06-12-17 10:33	06-12-17 11:44	06-12-17 11:49	06-12-17 12:04	06-12-17 12:50	06-12-17 13:18

Can we leverage a-priori knowledge in order to improve the accuracy of the prediction of the next activities?

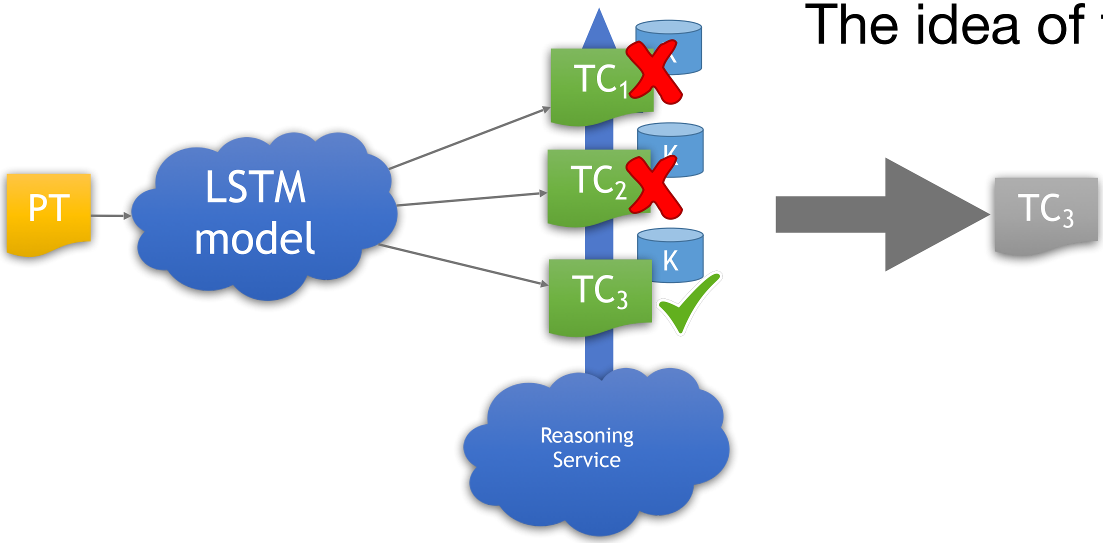
Security check at 3	Late show, no boarding
06-12-17 15:54	06-12-17 18:43

Solution:

Guide prediction algorithms with the a-priori knowledge

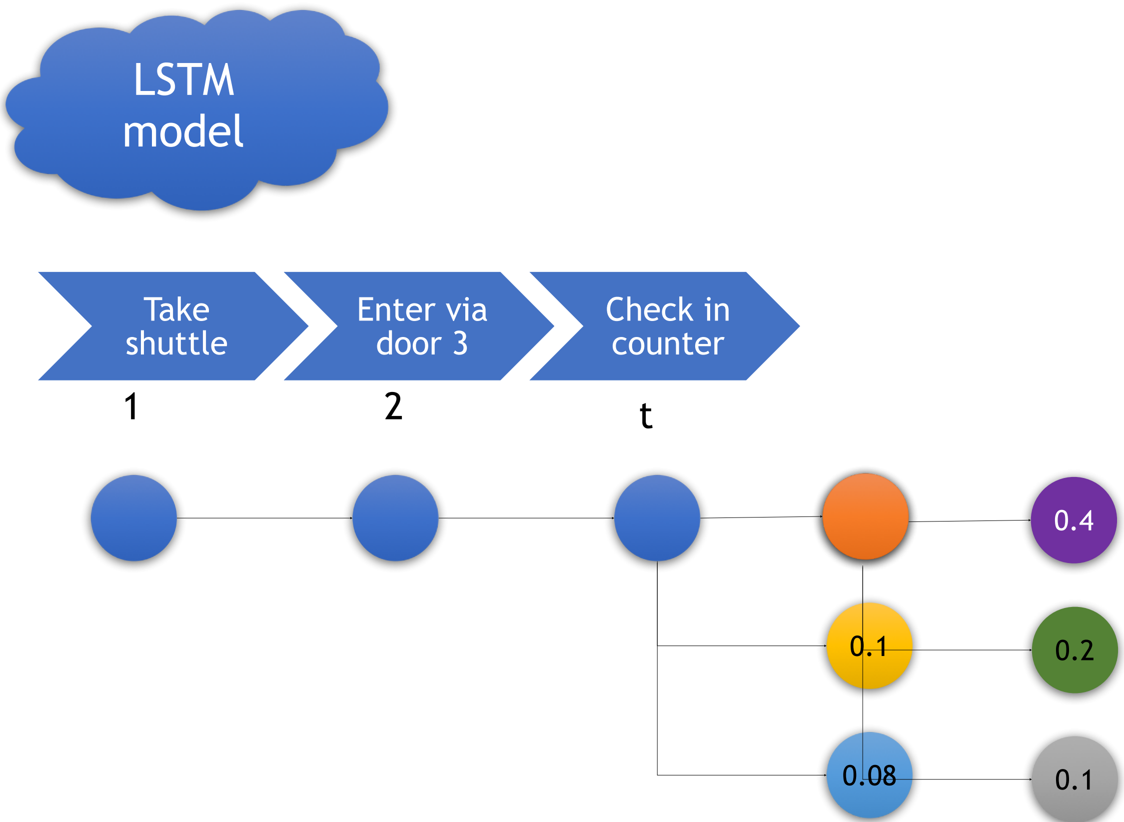


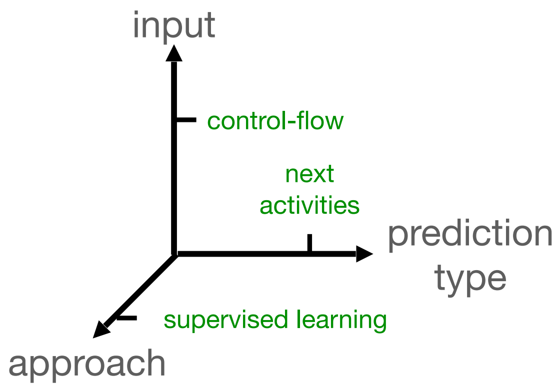
The idea of the solution



Solution:

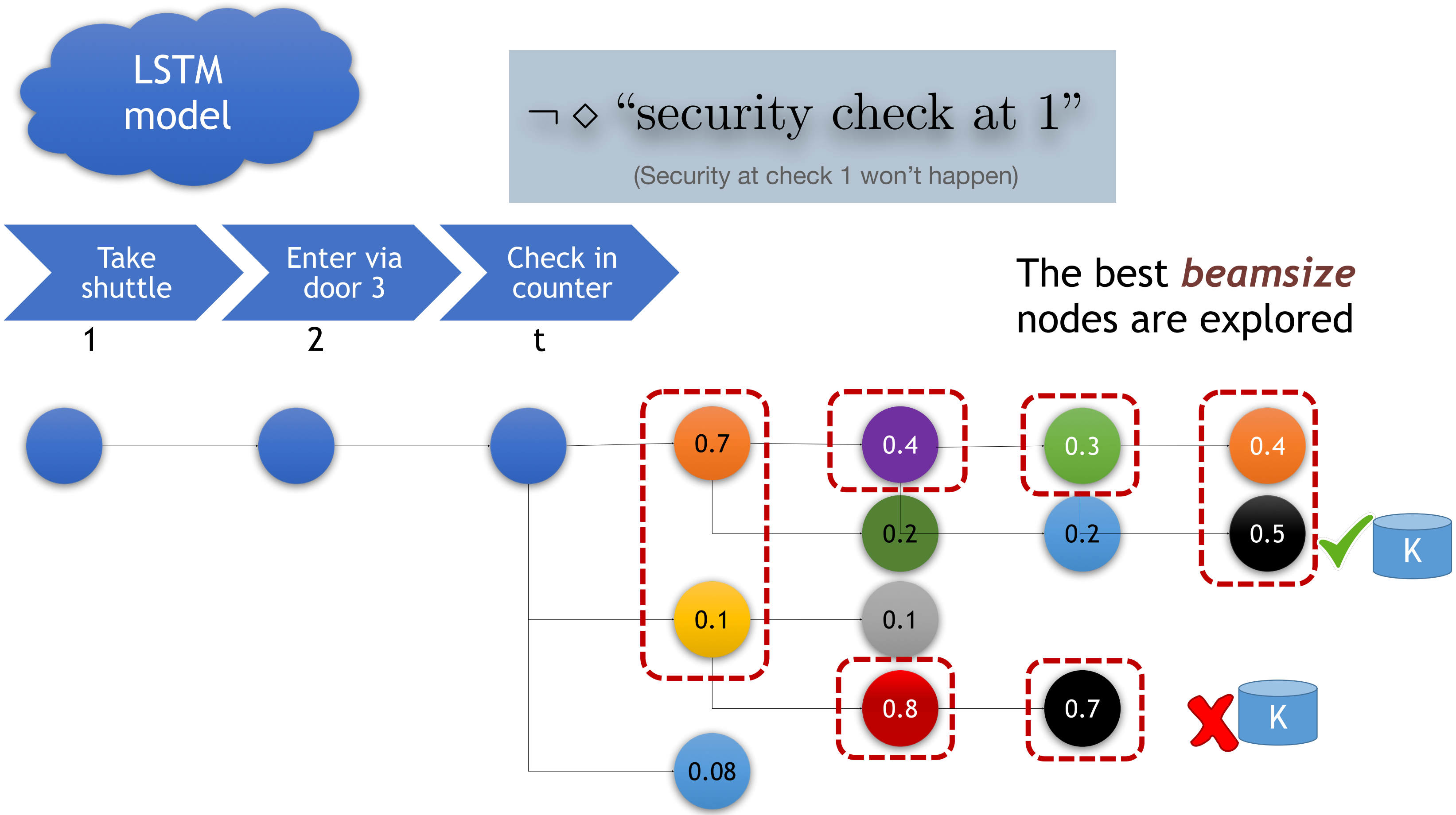
Guide prediction algorithms with the a-priori knowledge





Niek Tax, Ilya Verenich, Marcello La Rosa, Marlon Dumas: **Predictive Business Process Monitoring with LSTM Neural Networks**. CAiSE 2017: 477-492

Beam-search for a-priori knowledge



Evaluation

Log	A-priori Strong	A-priori Weak
EnvLog	$\square(a \rightarrow \Diamond b) \wedge \Diamond a \wedge \square(c \rightarrow \Diamond d) \wedge \Diamond c$	$\Diamond a \wedge \Diamond c$
HelpDesk	$\square(e \rightarrow \Diamond f) \wedge \Diamond e$	$\Diamond e$
BPIC11	$\square(g \rightarrow \Diamond h) \wedge \Diamond g \wedge \square(i \rightarrow \Diamond l) \wedge \Diamond i \wedge \square(m \rightarrow \Diamond n) \wedge \Diamond m$	$\Diamond i \wedge \Diamond h \wedge \Diamond o$
BPIC12	$\square(p \rightarrow \Diamond q) \wedge \Diamond p$	$\Diamond p$
BPIC13	$\square(r \rightarrow \Diamond s) \wedge \Diamond r \wedge \square(t \rightarrow \Diamond r) \wedge \Diamond t$	$\Diamond s \wedge \Diamond r$
BPIC17	$\square(u \rightarrow \Diamond v) \wedge \Diamond u$	$\Diamond u$

- A priori knowledge does help in producing more accurate predictions (unless the log is highly sparse)
- Strong a-priori more useful than weak a-priori

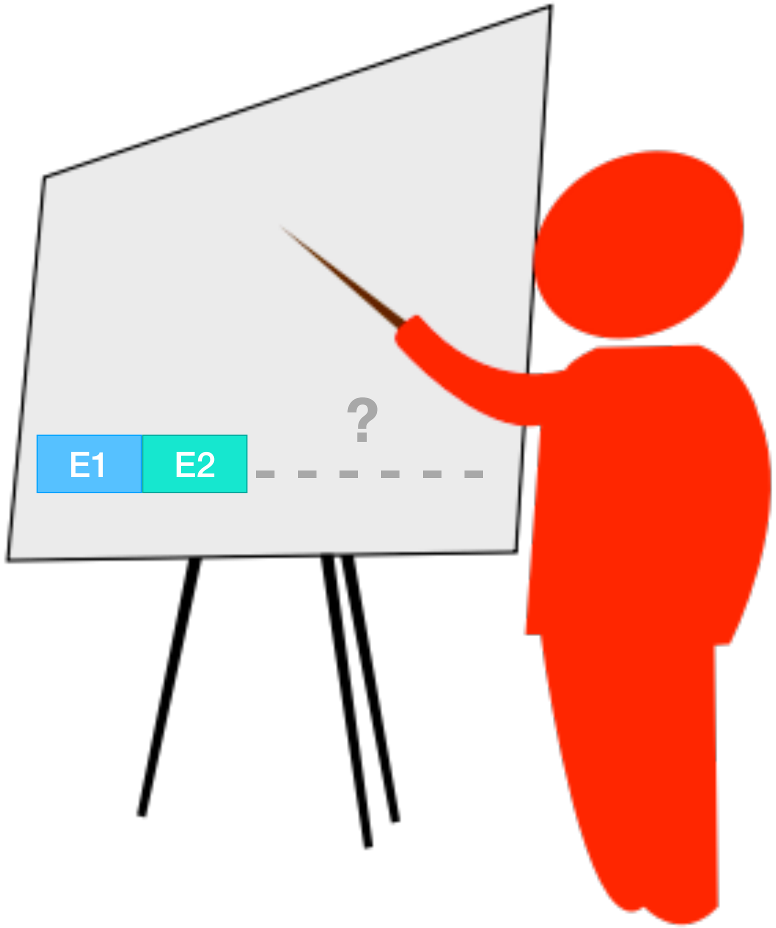
Leverage a-priori knowledge on data payloads

Doctor Charlie is on sick leave

$\neg \diamond \text{“}Manipulation''[Resource == Charles]$

Beam-search with data payload: How to combine events and data payloads?

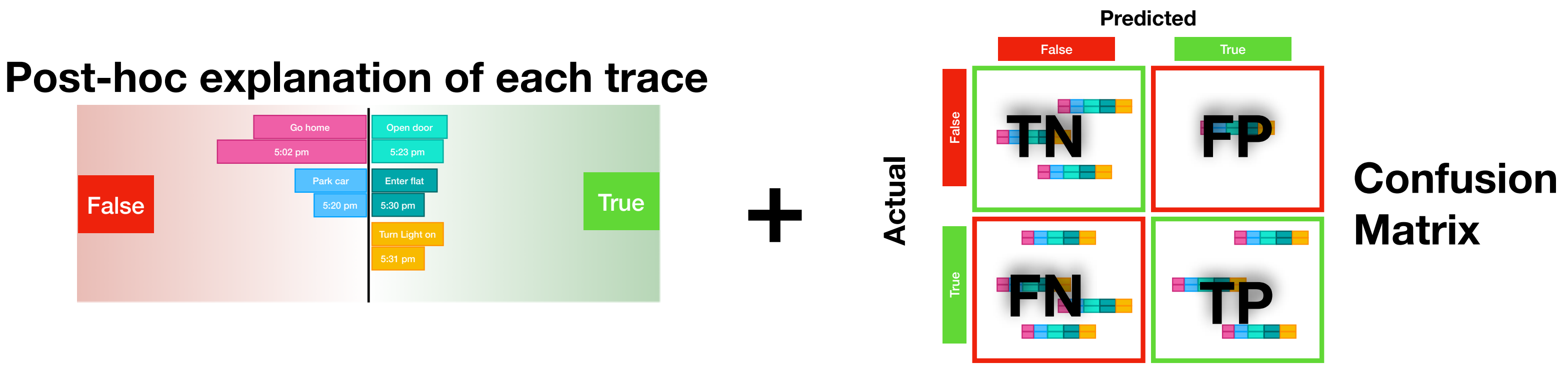
Let me explain!

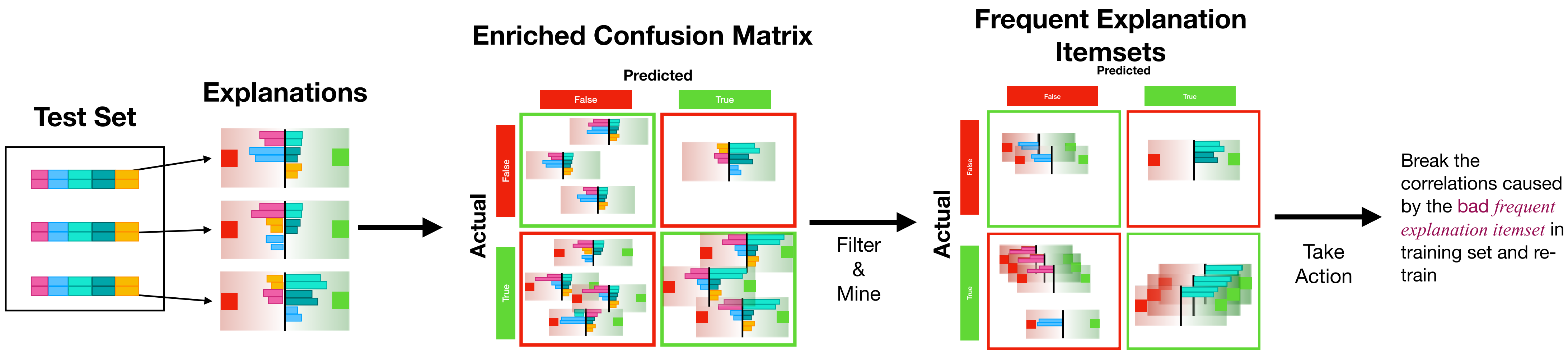


Why did we get into explanations?

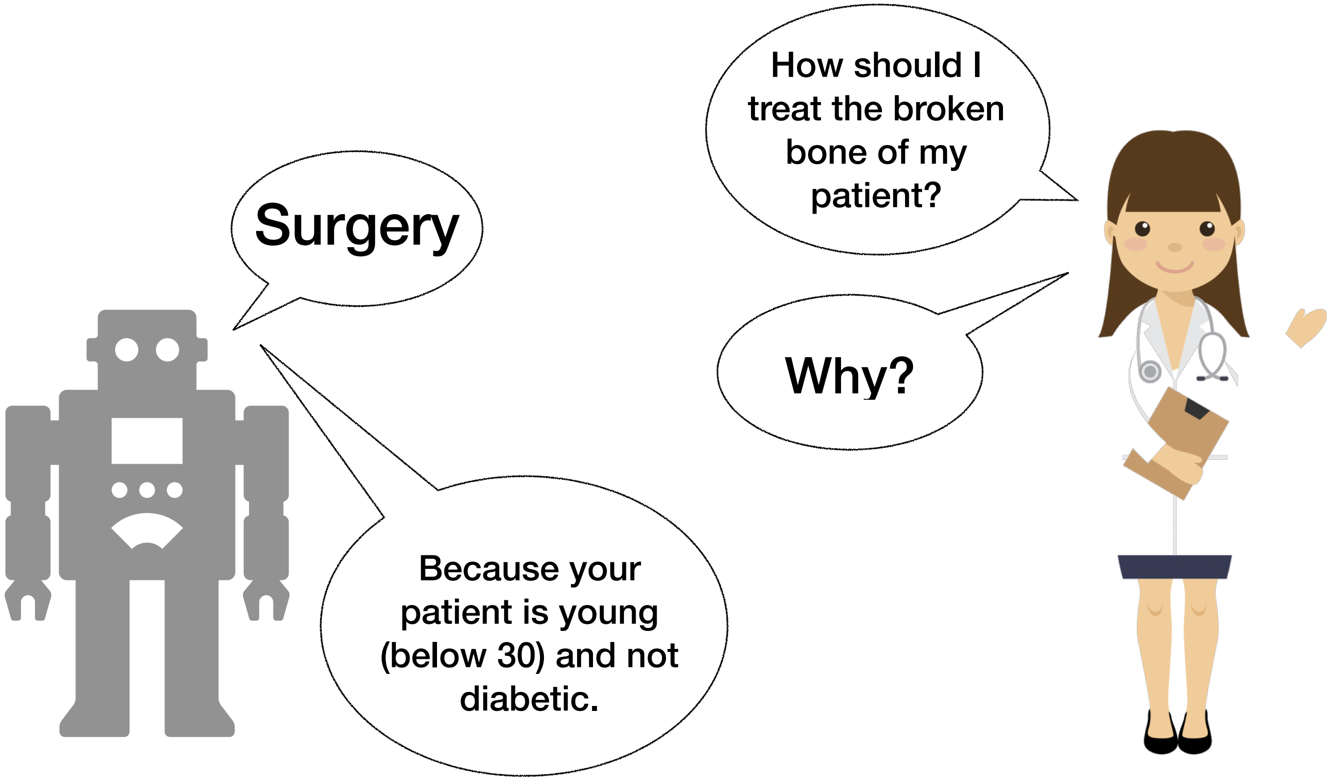
1. Understand what makes the predictive model return wrong predictions
2. Leverage this information to enhance the predictive model

The Ingredients and the Recipe

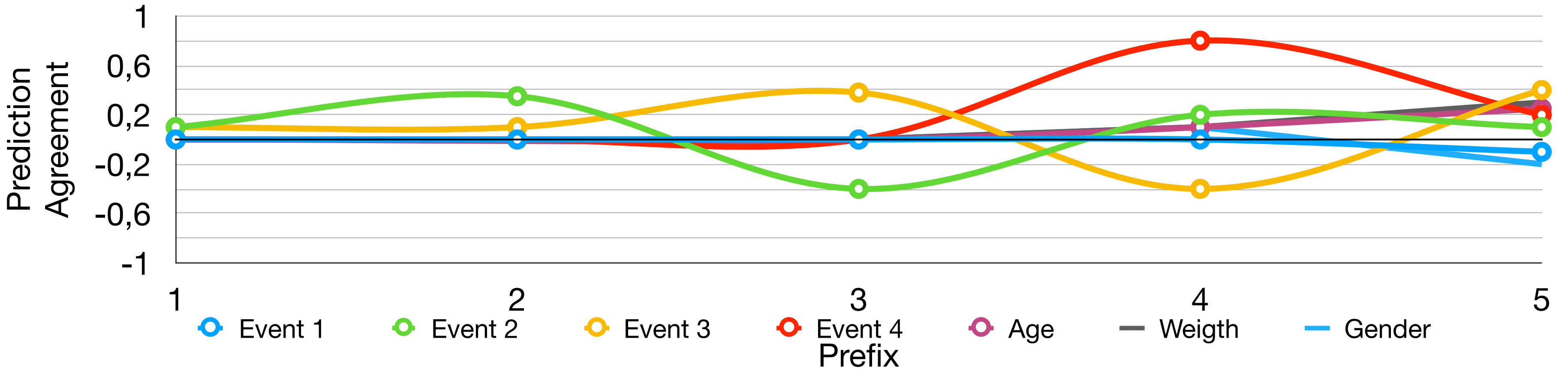


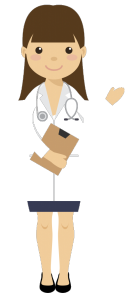


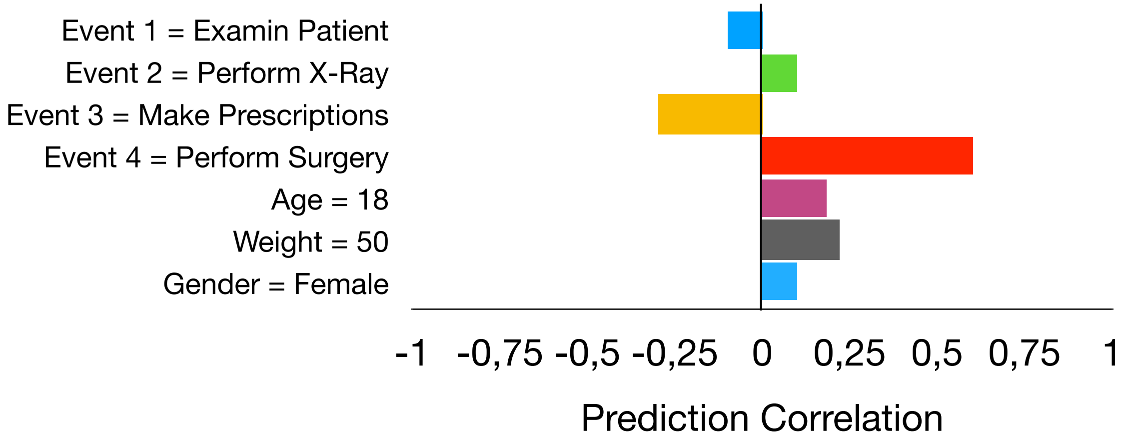
Empowering Users

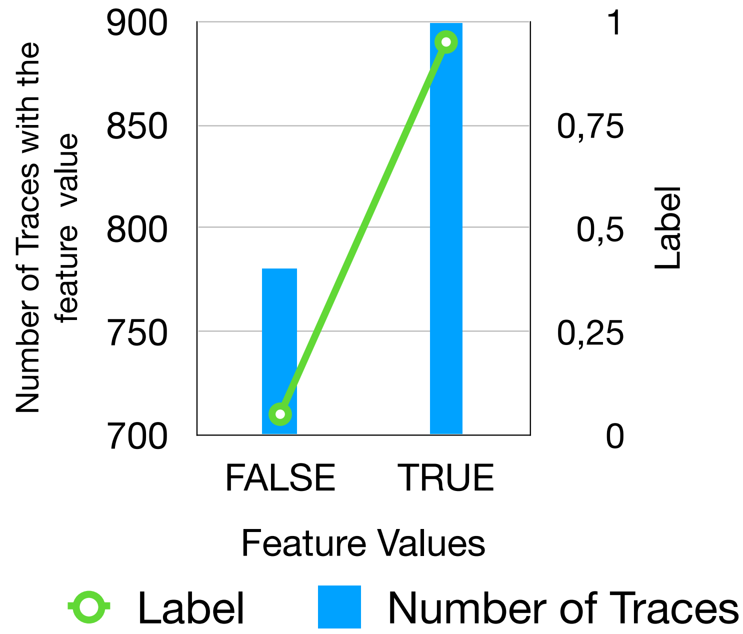


Understand the Prediction

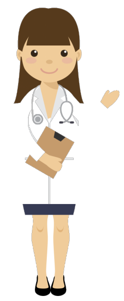






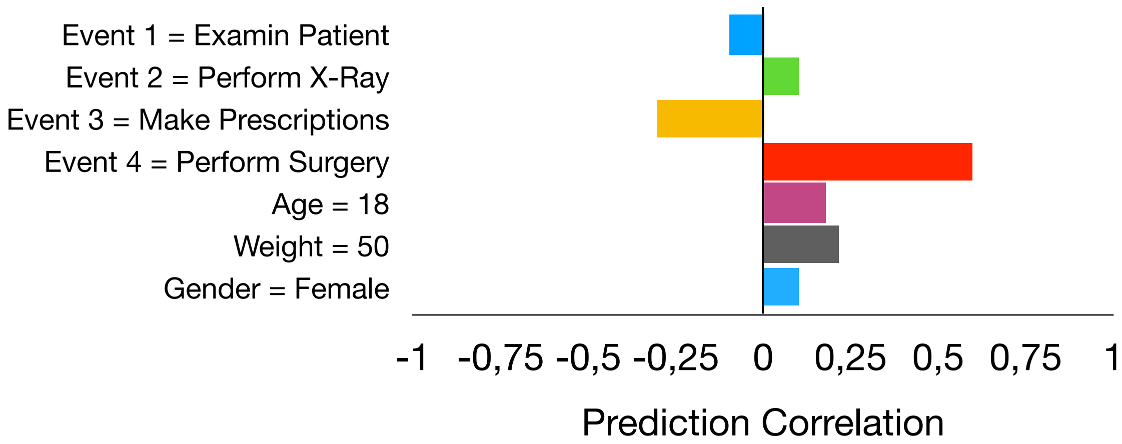


Understand the Prediction



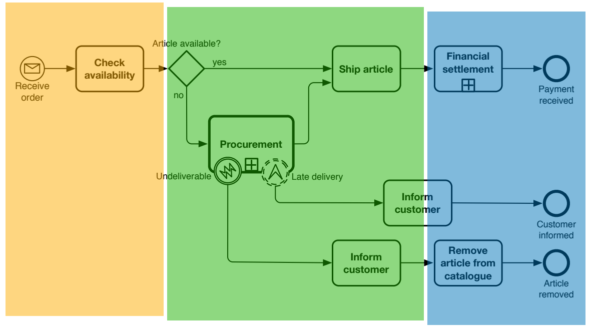
- Focus on factors influencing a prediction, without bothering about the process
- Focus on strong vs weak influences

Static

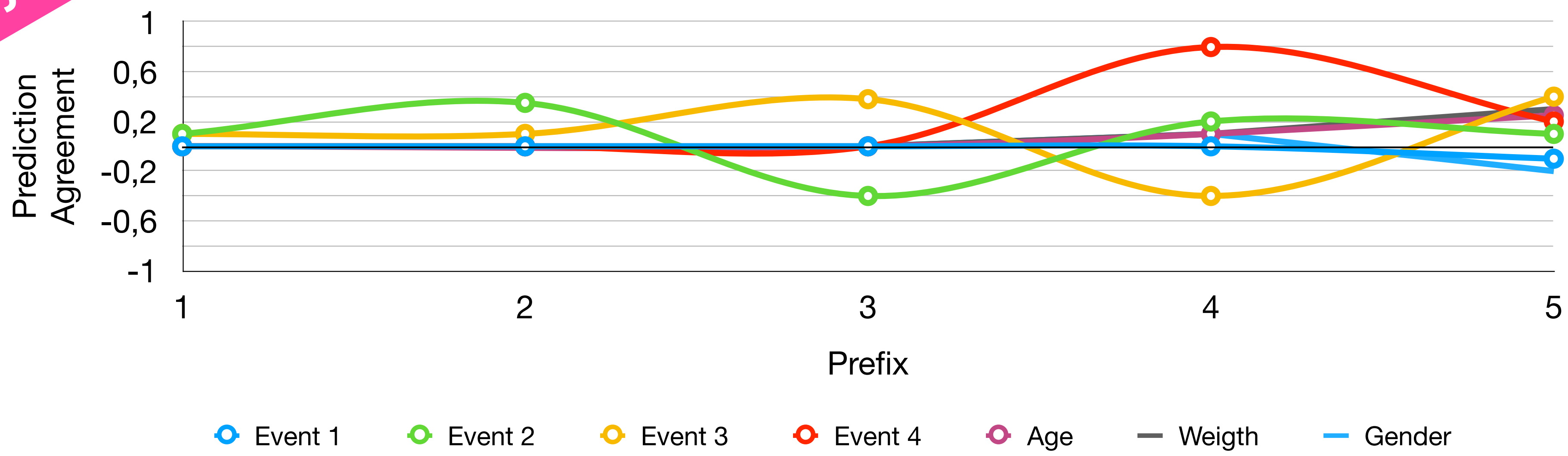


Understand the Prediction

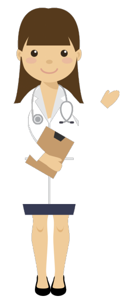
- Focus on the importance of features over time



Some dynamics

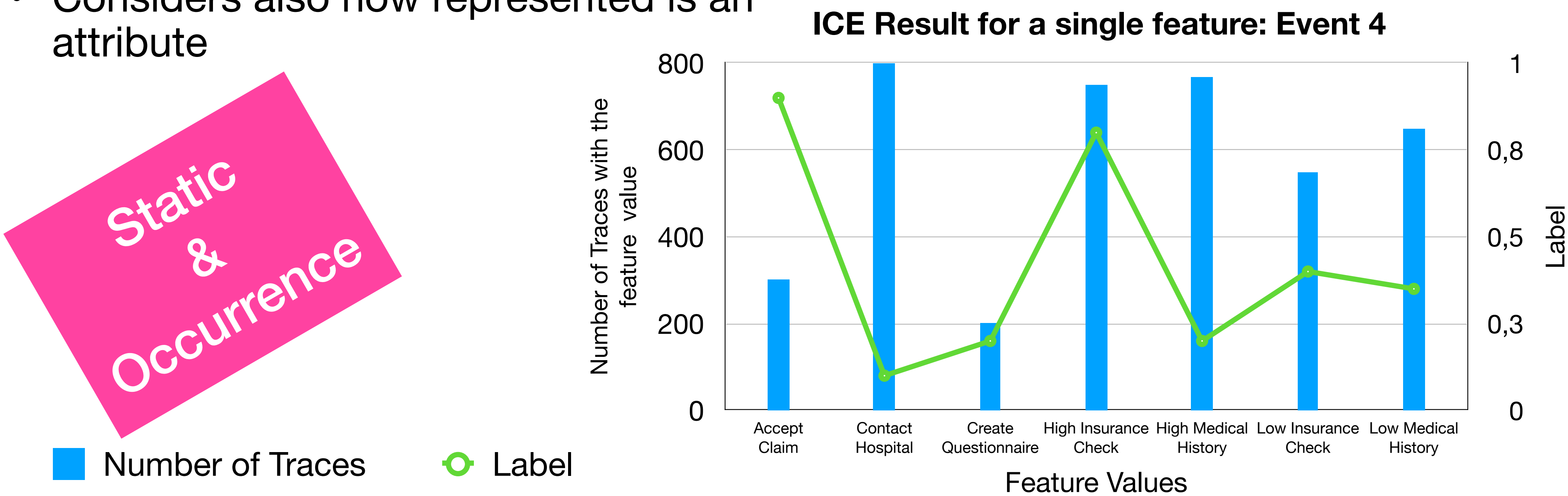


Understand the Prediction

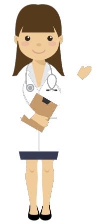


- Focus on dataset characteristics and their relation the specific values of the labels

- Considers also how represented is an attribute

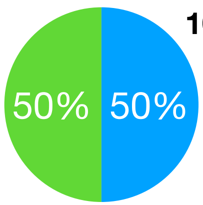


User Evaluation



USER
EVALUATION

1 hour



10 BPM Experts

- BPM
- BPM + ML

RESEARCH QUESTIONS

1. How do users **make sense** of explanation plots in PPM?
2. How can explanation plots support users in **decision making** tasks in PPM?
3. How can PPM explanation plots be improved?

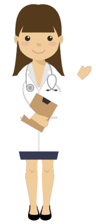
TASKS

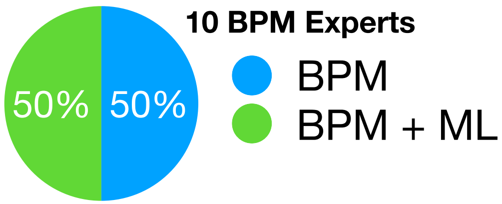
1. Event-level decision making
2. Case-level decision making
3. Process-level decision making

DOMAINS

1. Medical Domain
2. Banking Domain

Results —> next things to do!





- BPM + ML experts have more troubles in understanding the plots and making use of the plots than BPM experts
- Most BPM experts prefer to use the information contained in the plot to foster the focus of a more in-depth study of the data instead of the actual decision making
- All BPM experts experience some difficulties in autonomously spotting what can and what can not be changed in the process to satisfy the decision making task

Williams Rizzi, Marco Comuzzi, Chiara Di Francescomarino, Chiara Ghidini, Suhwan Lee, Fabrizio Maria Maggi, Alexander Nolte.
Explainability in Predictive Process Monitoring: can the explanation be useful for decision making tasks? Submitted for Journal publication.
In Collaboration with the University of Tartu (UT) and the Ulsan National Institute of Science and Technology (UNIST) in South Korea.

Learn to take action

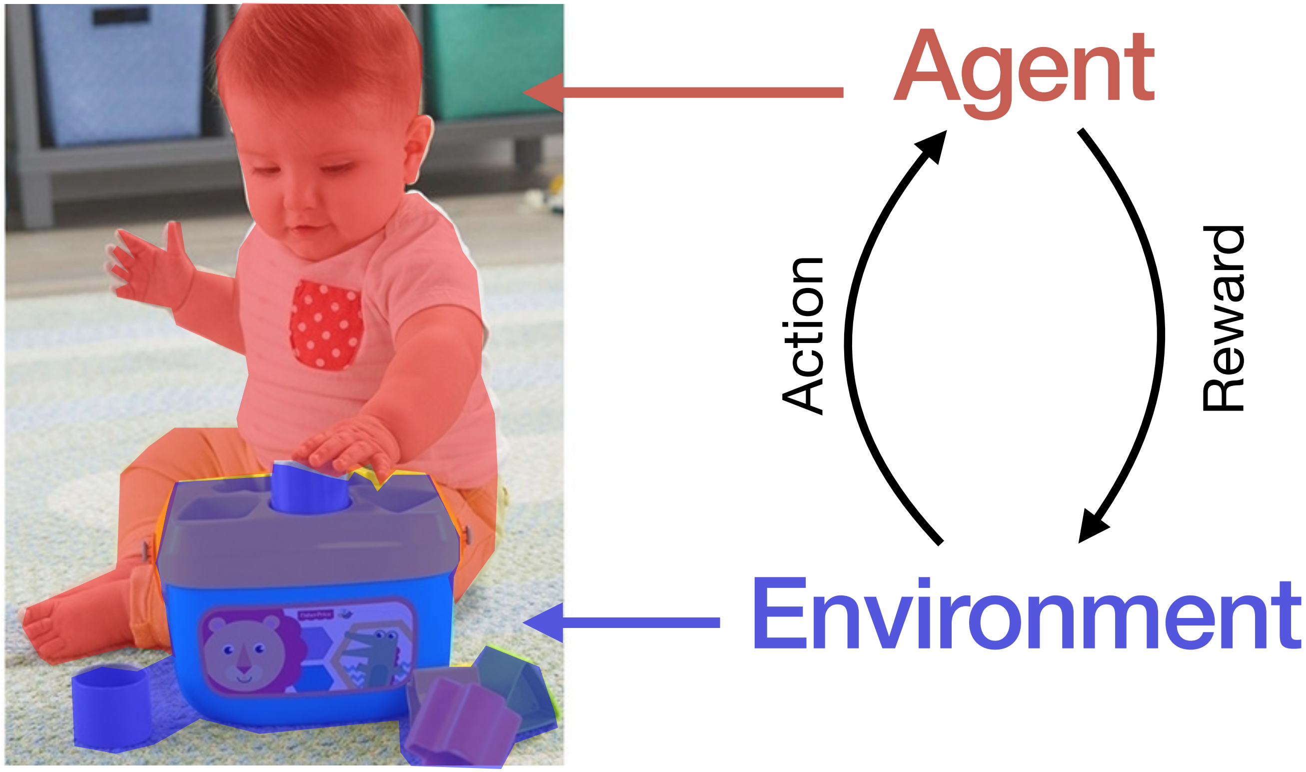
- Learn to recommend via Reinforcement Learning
- Suggest what to do via counterfactual explanation

Stefano Branchi, Chiara Di Francescomarino, Chiara Ghidini, David Massimo, Francesco Ricci, Massimiliano Ronzani
Learning to Act: a reinforcement Learning approach to recommend the best next activities. To appear in BPM Forum 2022.
In Collaboration with the University of Bolzano. (also in Arxiv)

Reinforcement learning

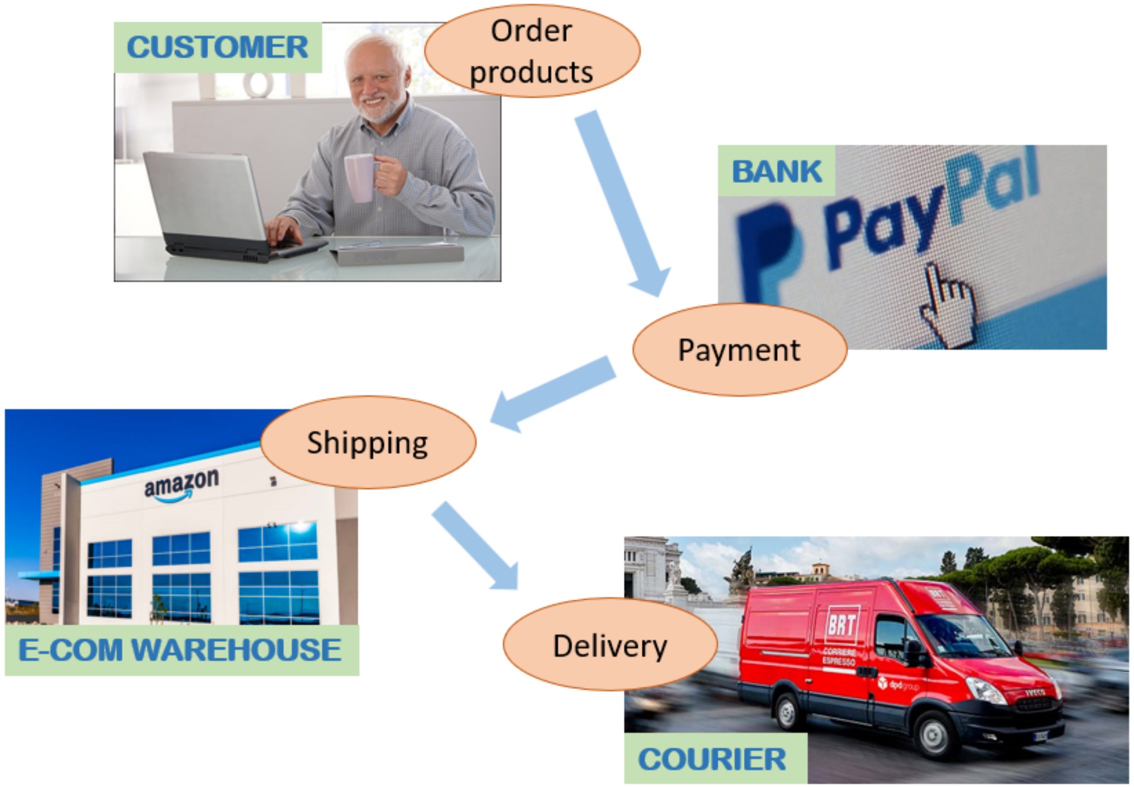
Learning trough experience (self-play)

The agent wants
to “win the game”
(maximize a KPI)

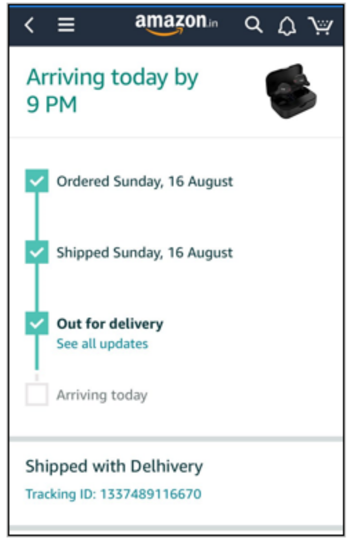


Multi-actor processes (e-commerce)

Every actor has their own interest: maximize a certain KPI



LOG



Main idea

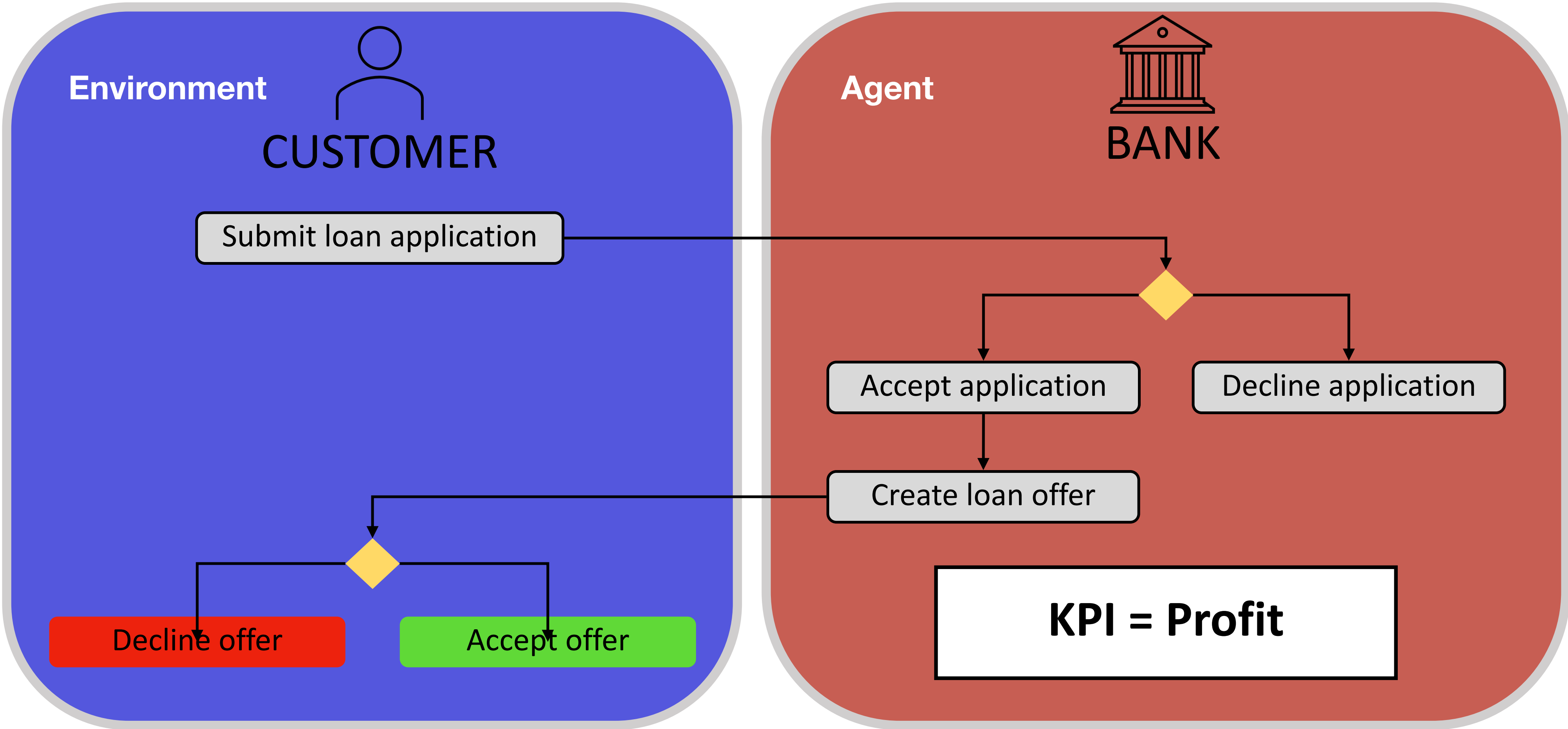
Focus: one actor → AGENT

All the other actors → ENVIRONMENT

WHAT: Recommend best activities (action) to maximize Agent's KPI (**BPM**)

HOW: Find optimal policy π^* of a pertinent MDP (**RL**)

Case study: loan process



BPM: From the Event Log to the MDP

TRACE of EVENT LOG

1. Loan application submitted
   - 24/03/2022 19:00
   - amount = 100.000 $
2. Loan application accepted
   - 26/03/2022 15:00
3. Create loan offer
   - 28/03/2022 09:30
   - offer amount = 80.000 $
   - interest rate = 15%
4. Assess application
   - 02/04/2022 16:00
   - outcome = accepted

+

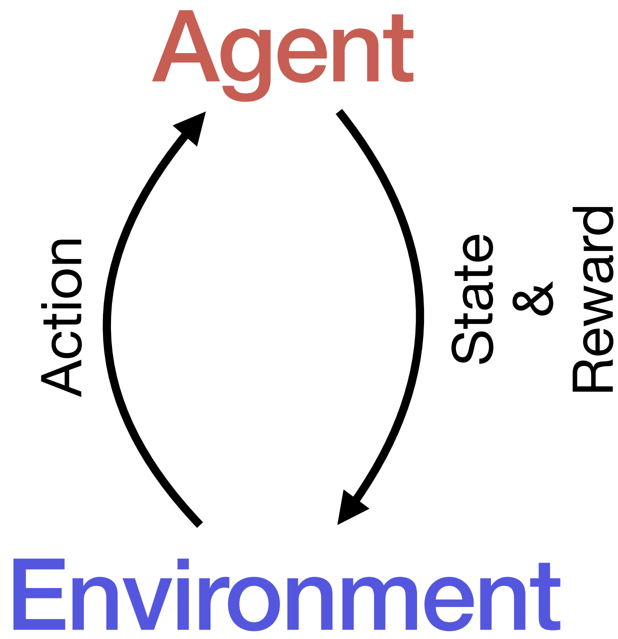
DOMAIN KNOWLEDGE

- Ownership of each activity (agent/environment)
- Features important for the decision making (attributes or computed)

BPM: From the Event Log to the MDP

TRACE of EVENT LOG

1. Loan application submitted (Customer)
   - 24/03/2022 19:00
   - amount = 100.000 $
2. Loan application accepted (Bank)
   - 26/03/2022 15:00
3. Create loan offer (Bank)
   - 28/03/2022 09:30
   - offer amount = 80.000 $
   - interest rate = 15%
4. Assess application (Customer)
   - 02/04/2022 16:00
   - outcome = accepted



BPM: From the Event Log to the MDP

The MDB State

⟨ Last activity , Historical features , Environment features ⟩

- last activity: by agent or environment (e.g., *create offer, accept offer*)
- historical features: information about past activities (e.g., how many offer the bank sent to the customer in the past)
- environment features: depending only on environment (e.g., the loan amount requested by the customer)

BPM: From the Event Log to the MDP

MDP: action, stochastic & reward

- **Action**: AGENT activities (to recommend)
- **Stochastic**: ENVIRONMENT response (mined from the event log)
- **Reward**: (KPI) is the profit of the bank (€)
  - Positive: loan interest (if **customer** *accepts offer*)
  - Negative: **bank** operating costs (too many employees activities ~ salary, ...)

RL: Find optimal policy π^*

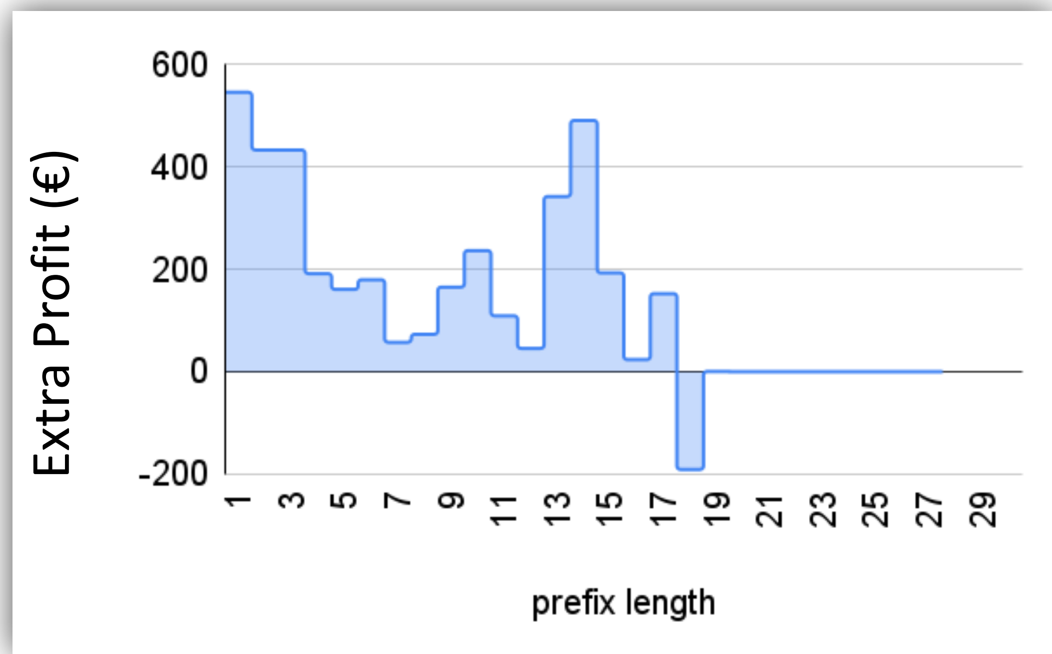
- Let the MDP play with a simulated environment

  (due to a scarcity of data … would be nice to have a big enough event log so
  as to use real customers data)

- Optimal policy: keep to make offers to the client!

Evaluation: how good is π^* and when in the process

- **RQ1**: How good is the recommendation given by π^*?
- **RQ2**: How does it perform at different points in the execution?
- Evaluation: Actual vs Recommended Analysis of incomplete trace executions in test log



Discussion

The dependence on domain knowledge

- Deep or cluster approach for the management of richer states space which **encode a multidimensional history**
- Automatise the activity ownership annotation (agent/environment)

What is optimal, for **whom** and **when**

- We are missing an important part of the story, that is when the customer is not able to repay the loan to the bank —> when do we consider the game over?
- What are the constraints we need to take into account (e.g, personnel costs or welfare in engaging with client, ..)
- [a crude experiment in best action for minimizing time into hospital did actually learn to immediately send patients at home —> multidimensional knowledge is **crucial** in providing decision support in BPM]

Learn to take action

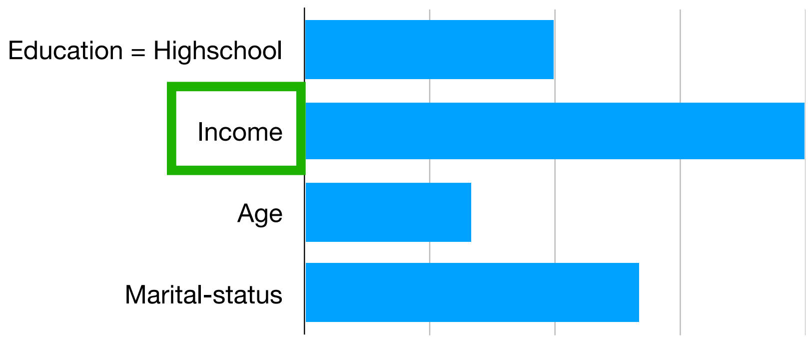
- Learn to recommend via Reinforcement Learning
- Suggest what to change via counterfactual explanation

Andrei Buliga, Chiara Di Francescomarino, Chiara Ghidini. **How to change the outcome of my process? Diverse Counterfactual Explanations for Predictive Process Monitoring.** Submitted,

Explanations beyond feature importance

What if, instead of providing the most important features, we provide the needed change in the input to reach a desired outcome?

Feature importance



Feature Importance Score

Counterfactual explanations

Exploring “what-if” scenarios Watcher et al. (2017)

If your income was $5,000 higher, you would been granted the loan

Desirable properties for counterfactual explanations in PPM

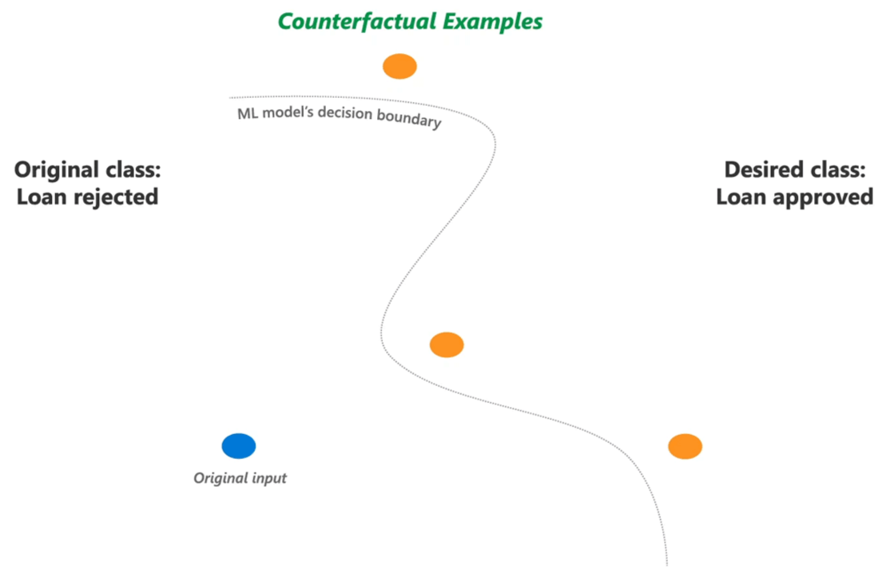
Actionability: the changes recommended by the counterfactuals should be feasible in real-life

Feasibility:

- **Proximity**
- **Plausibility**
- **Sparsity**
- **Control flow constraints**
- **Feature constraints**

+

Diversity



Techniques and challenges

Techniques:

- Case-based methods → Looking in the event log at the closest instances w.r.t distance that the predictive model predicts as the opposite class
- Exogenous (synthetic) methods → Generating synthetic examples through an optimisation process in order to reach the closest counterfactual examples
- Hybrid methods → Utilising an optimisation technique while utilising existing trace executions to guide the optimising process

Challenges:

- Ensuring validity of counterfactual explanations
- Evaluation of counterfactual explanations
- Visualisation of counterfactual examples

Summing up

- Predictive AI meets Event Logs: Predictive Process Monitoring
- Going beyond prediction is challenging but lots of fun
  - Explain multidimensional data
  - Recommend with multidimensional data
  - Keep in mind the impact of what we do

Thanks!