

Decision Camp 2023

From ML integration to Decision Transparency

September, 2023

Guilhem Molines
STSM, Chief Architect – IBM Decision Automation

Your speaker

STSM, Chief Architect for IBM Decision Automation

ODM (formerly, ILOG Jrules)

ADS

Previously:

Consultant (1998-2004) covering US East Coast

Computer Sciences Master, AI major



@GuilhemMolines



[linkedin.com/in/guilhemmolines](https://www.linkedin.com/in/guilhemmolines)

Level Up



Setting the stage

Added value of rules over code:

- Transparency, visibility to the business
- Traceability
- Explainability

This talk starts from a real-life use case, looks at transparency, and proposes ideas for improvements

Financial transaction fraud use case

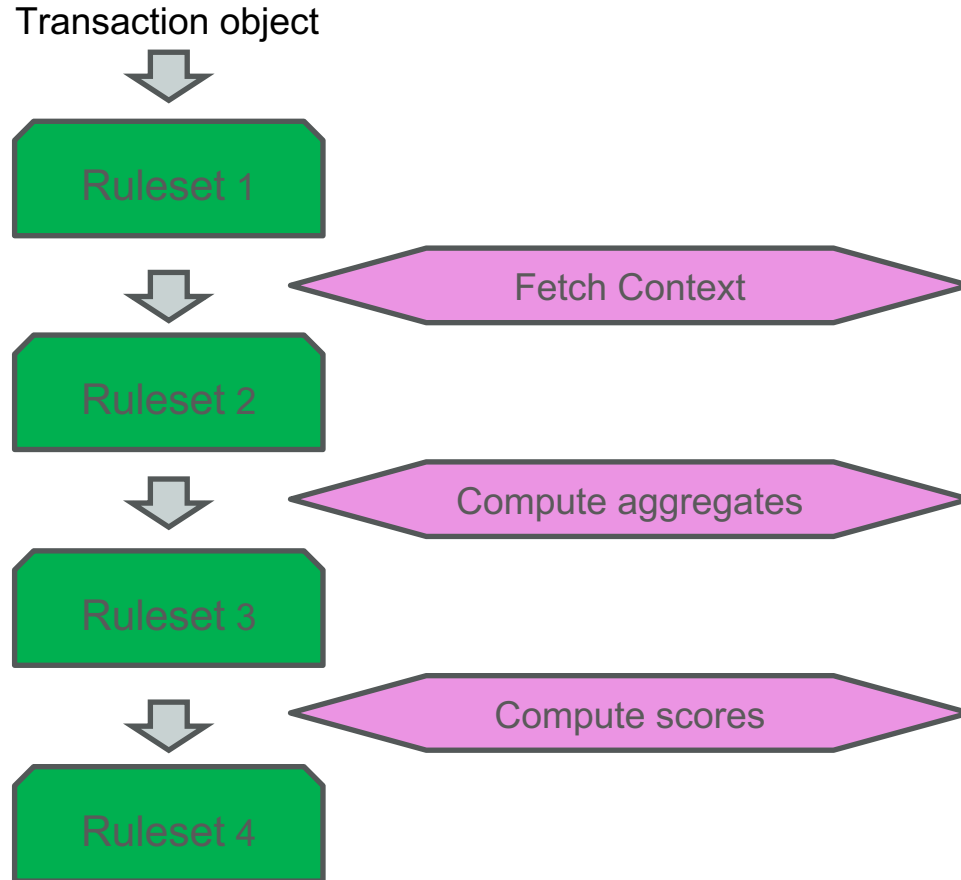
How to detect a fraudulent transaction?

- Transaction: amount, date, currency, merchant,
- Context: contract, countries, thresholds, exchange rate, ...
- Aggregates: last x days total, max spend, sliding averages, ...
- Scores: risk, merchant history, previous fraud, specific patterns, ...

Data distance: from immediate to more and more expensive to compute

Data complexity: from simple to requiring deep expertise

Approach



Issues

Each ruleset is its own model (ruleflow or decision model)

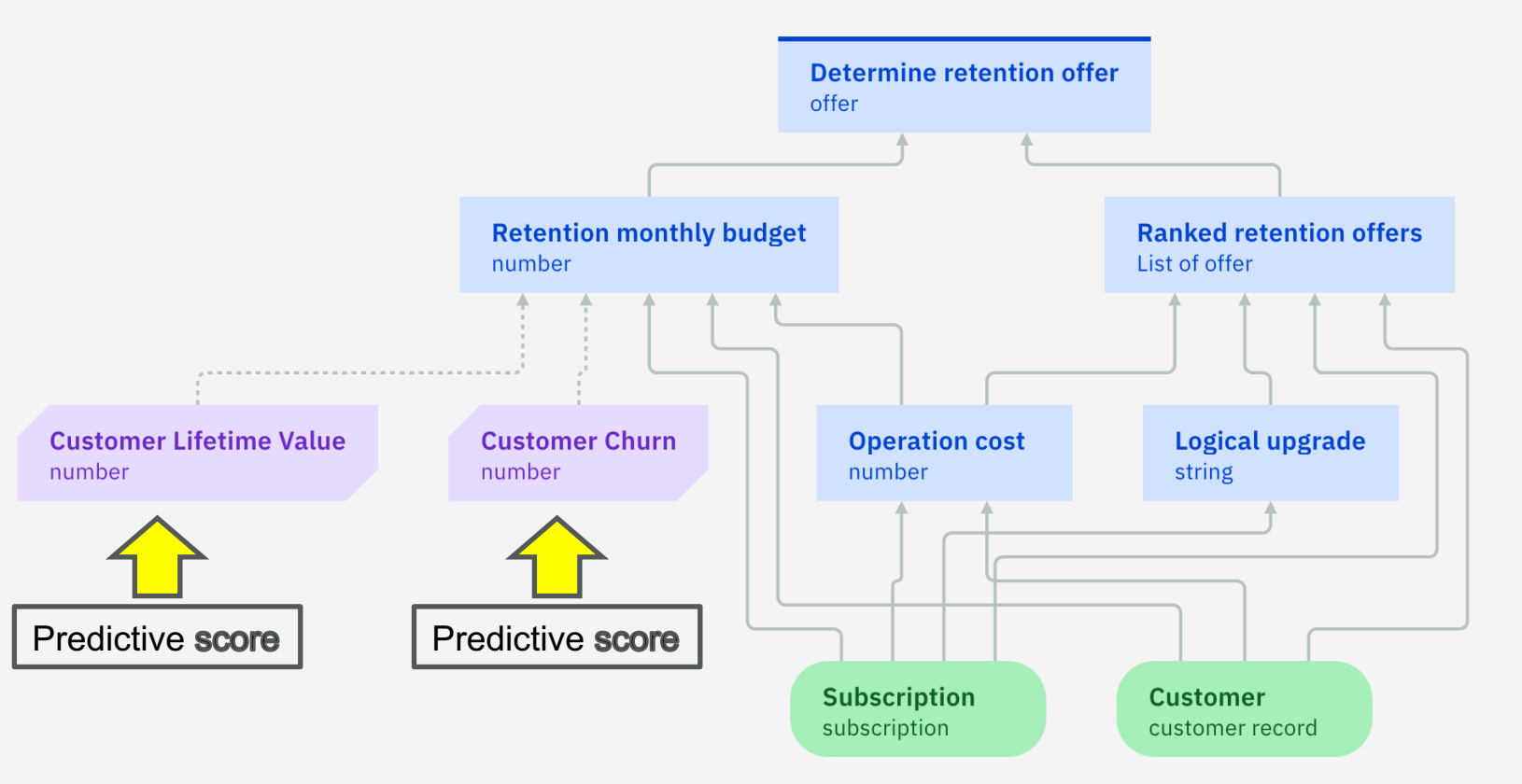
Orchestration external to rule engine

→ No global vision of the decision

Idea #1

- Need a global decision model
- Need sub-models and composition
- Need a representation of the Predictive Score

Example



Issues

Predictive models are black boxes

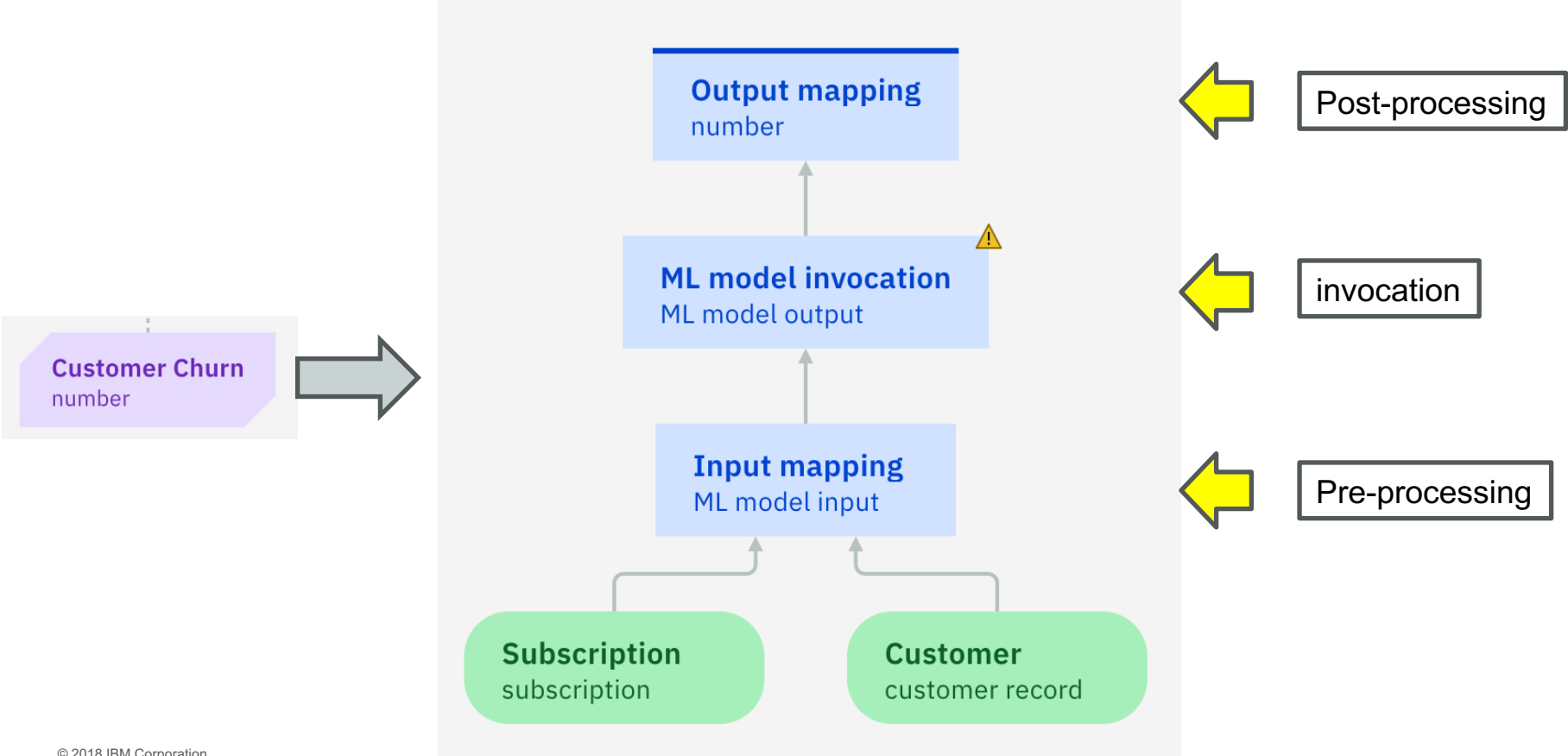
Decision Model points to an ML Model, but difficult to know which one, is it relevant?

→ Decision is now clearer globally, except for the ML part

Idea #2

- Wrap the ML model

Example



Example (continued)

Post-processing

```
if
    'ML model invocation' is in error
then
    set decision to 0.5 ;
```

Pre-processing

```
set decision to the probability t of 'ML model invocation' ;
```

rate plan ∨

	rate ... ↑↓	ML rate plan	↑↓
1	Basic		0
2	Essenti		1
3	Gold		2
4	Premiu		3

Issues

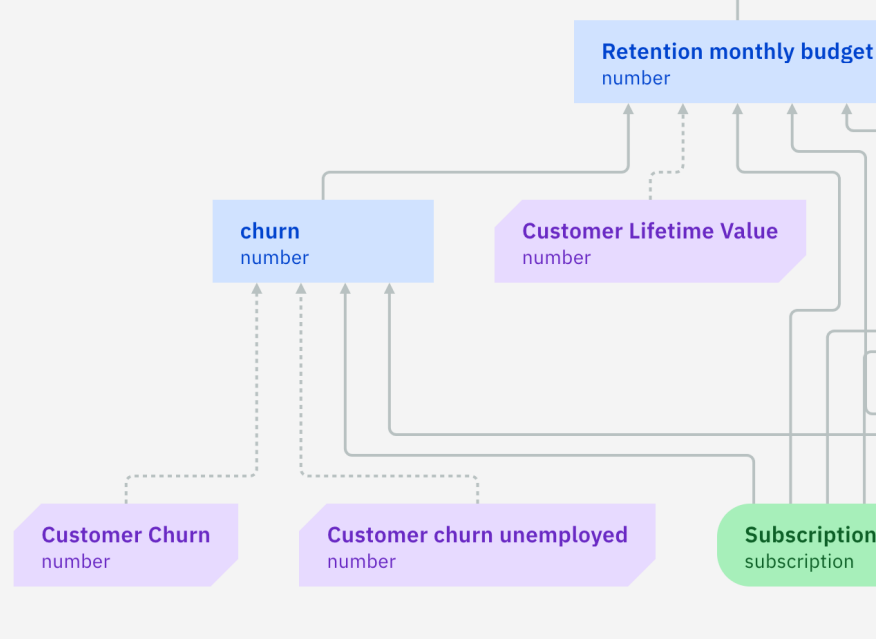
How the model is used is now clearer

But we don't know if using a relevant model

Idea #3

- ML model selection based on processed data

Example



```
if
  Customer is unemployed
then
  set decision to the customer churn unemployed computed from
    Customer being Customer ,
    Subscription being Subscription ;
else
  set decision to the customer churn computed from
    Customer being Customer ,
    Subscription being Subscription ;
```

Issues

Now we know which model is used

The choice of the model is made based on data passed to it and on rules

→ The model is still a black box

Idea #4

- Rule induction

Rule induction

What about we could rewrite the model as a set of rules / tables?

<https://pages.github.ibm.com/Dublin-Research-Lab/aimee-pages/>

- *Interpretable surrogate*: a compact representation of the model behaviour that users can assess and interact with. It reflects the model's decision logic in the form of Boolean decision rules. Boolean rules are used as the building blocks of the AIMEE. Users can also provide feedback in the form of Boolean rules.

References:

- Alkan, Öznur, Dennis Wei, Massimiliano Mattetti, Rahul Nair, Elizabeth M. Daly, and Diptikalyan Saha. "FROTE: Feedback Rule-Driven Oversampling for Editing Models." arXiv preprint arXiv:2201.01070 (2022).
- Nair, Rahul, Massimiliano Mattetti, Elizabeth Daly, Dennis Wei, Ozgur Alkan, and Yunfeng Zhang. "What Changed? Interpretable Model Comparison." IJCAI, 2021.
- Daly, Elizabeth M., Massimiliano Mattetti, Öznur Alkan, and Rahul Nair. "User Driven Model Adjustment via Boolean Rule Explanations." In Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35, no. 7, pp. 5896-5904. 2021. Harvard
- Owen Cornec, Rahul Nair, Öznur Alkan, Dennis Wei, and Elizabeth M. Daly, "AIMEE: Interactive model maintenance with rule-based surrogates." NeurIPS, 2021.

AIMEE in action

Demo

Conclusion

AIMEE brings some transparency

Helps with comparing two versions of same model after retraining

Still not perfect set of rules, requires human interpretation



Thank you