Building robust and defensible ML systems in Fincrime

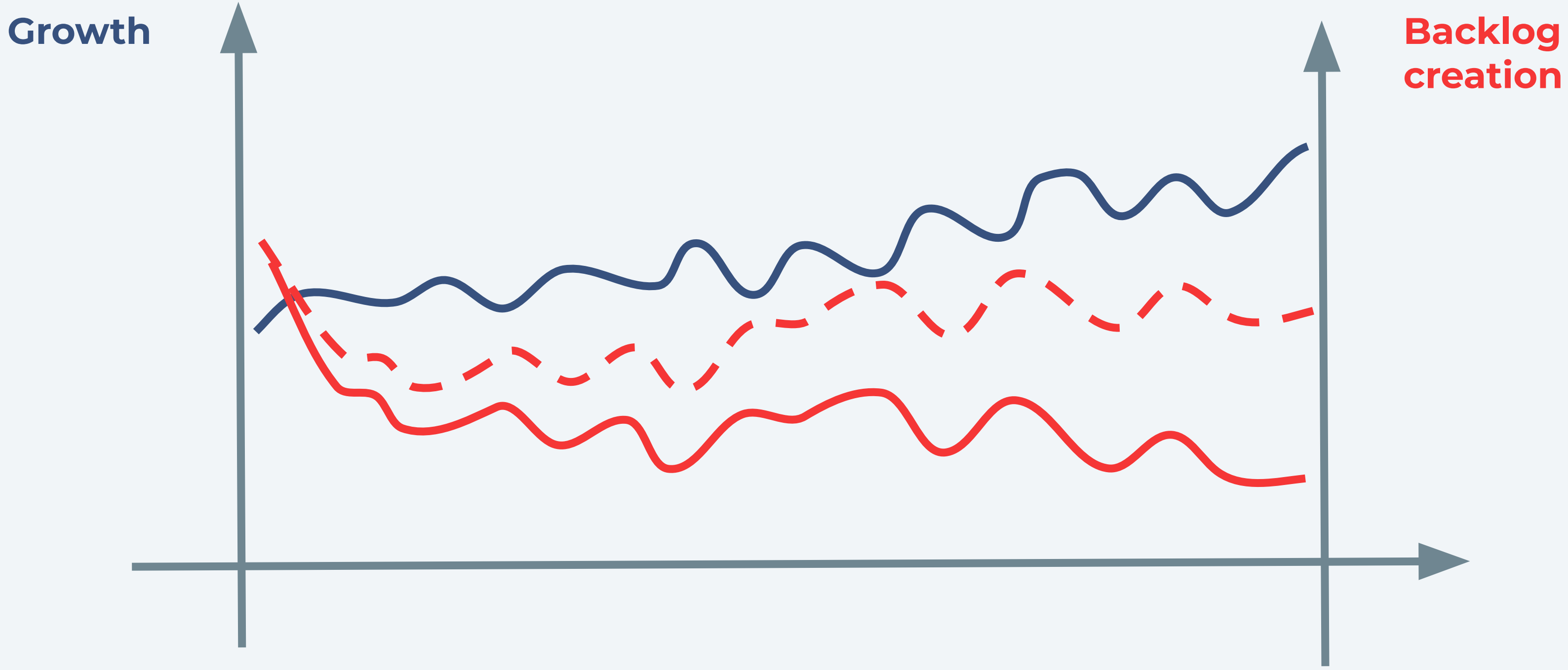
Luca Traverso
Lead Data Scientist @Wise

Why Machine Learning.

Fincrime Team Objectives

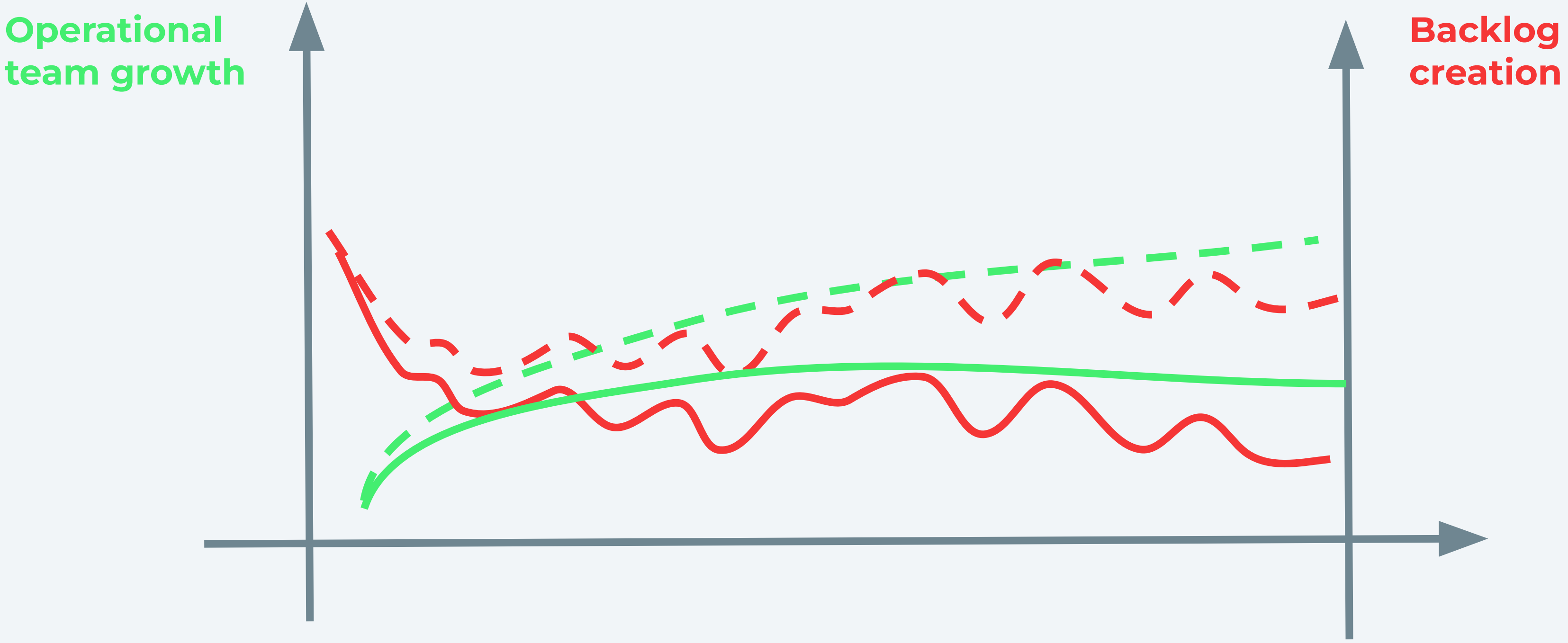
- Prevent onboarding of criminals
- Minimizing the impact on good customers

In fast growing companies



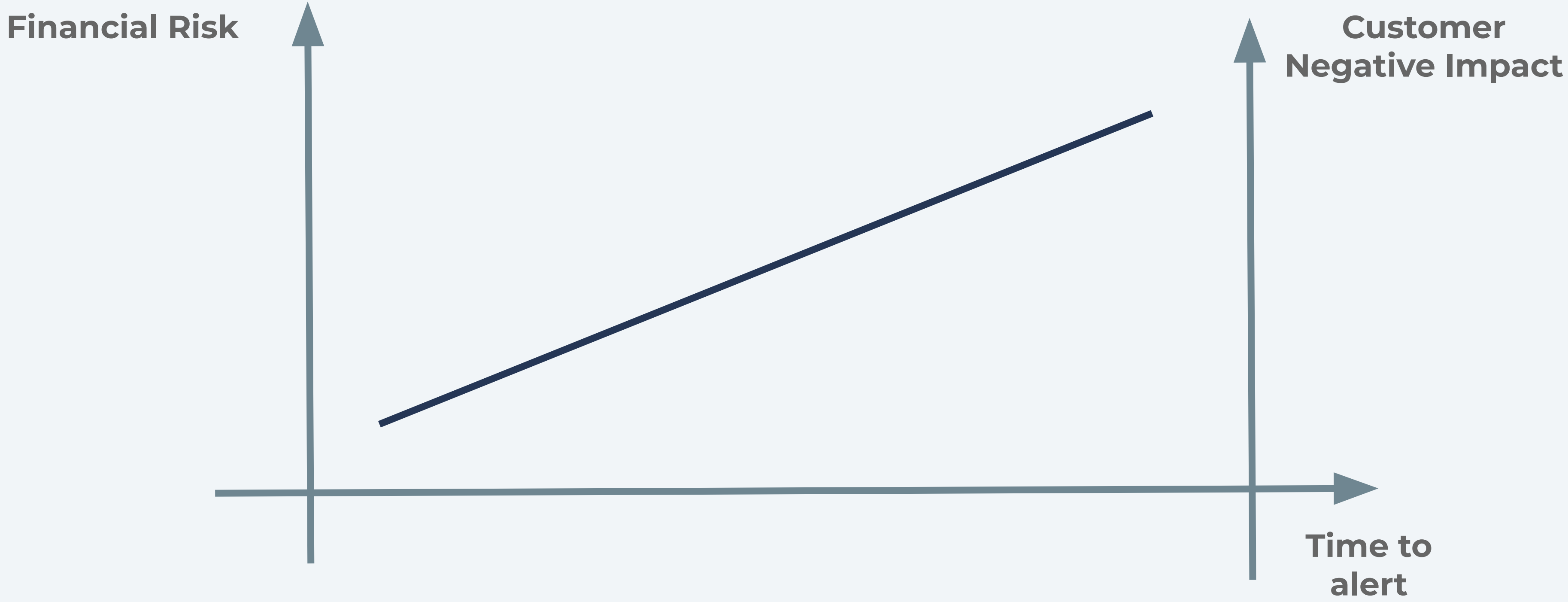
* **Backlog** : alerts / reviews generated

In fast growing companies



* Backlog : alerts / reviews generated

In fast growing companies



Machine learning system

Gives us :

- ❖ **Stable backlog creation**
- ❖ **Reduce risk exposure**

Needs to be :

- ❖ **Robust : it adapts to changing conditions**
- ❖ **Defensible : understandable and credible to a third party**

How do we do that ?

ML system components

1. Feature creation & labelling
2. Automatic model retraining (*closed-loop learning*)
3. Model governance
4. Model interpretability

ML system components

1. **Feature creation & labelling**
2. Automatic model retraining (*closed-loop learning*)
3. Model governance
4. Model interpretability

Features creation

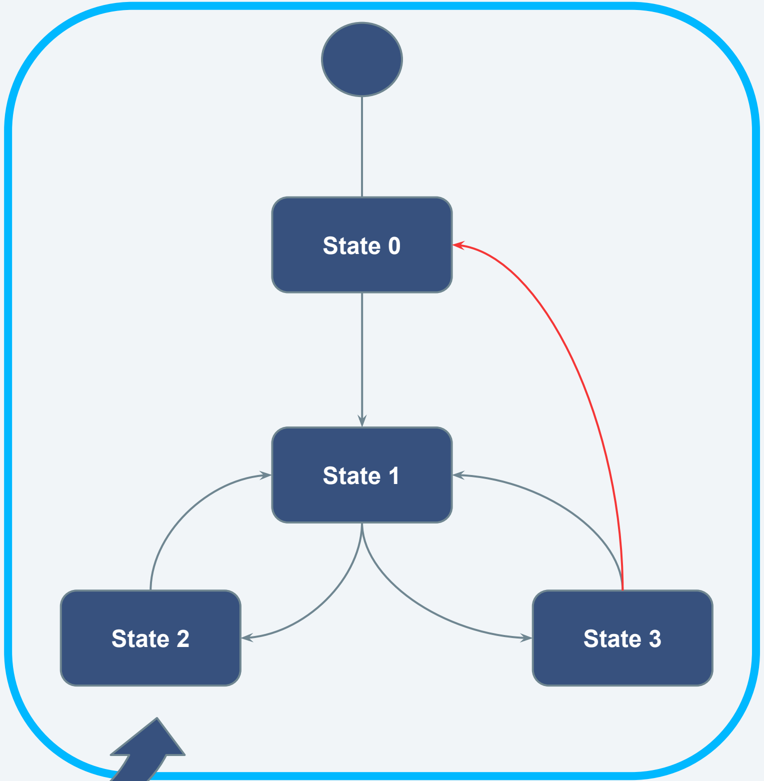
Feature : The numerical representation of an observed behaviour

Defensible :

If features can be explained model results can be better interpreted / explained

Labelling

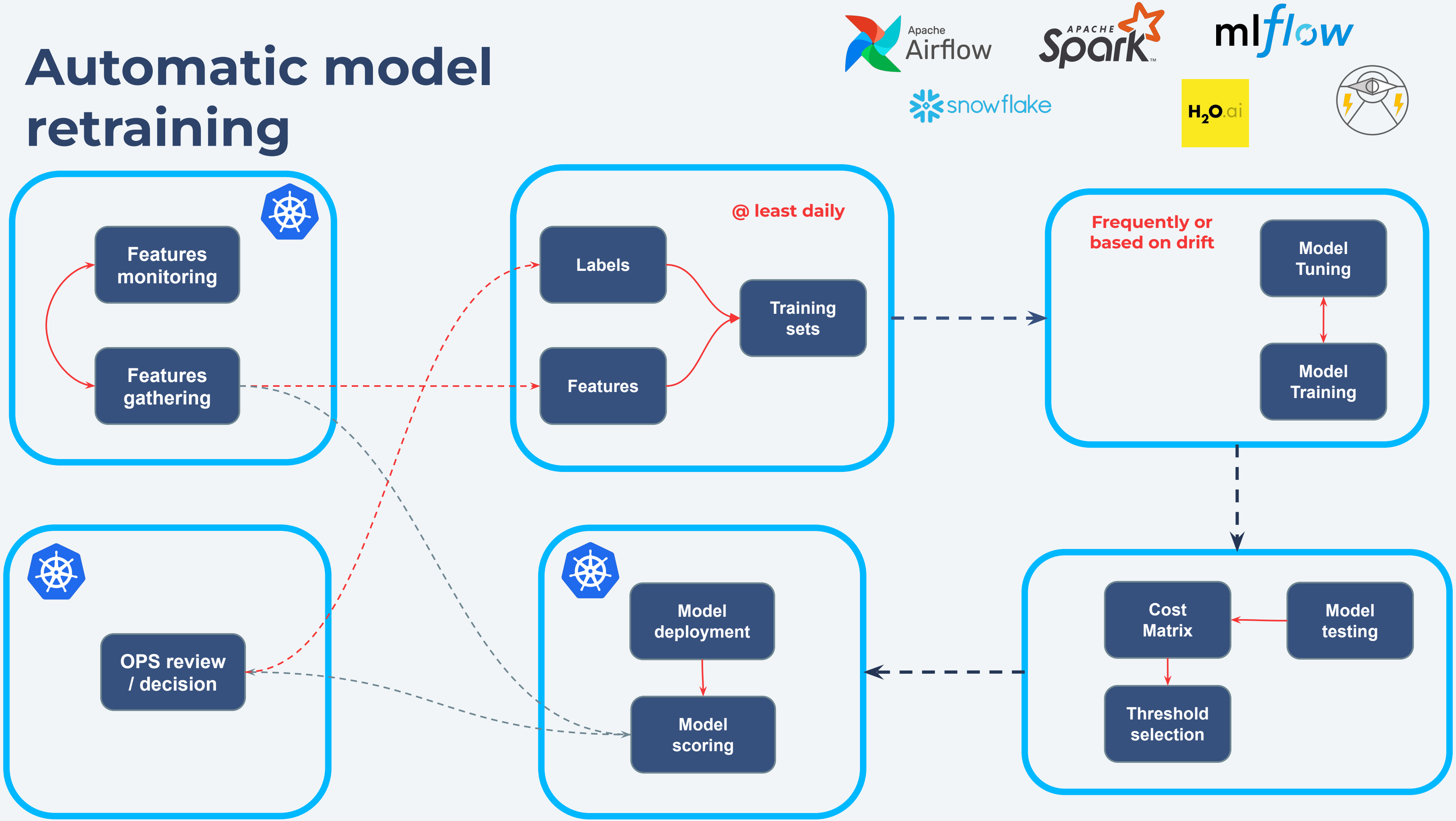
Tooling



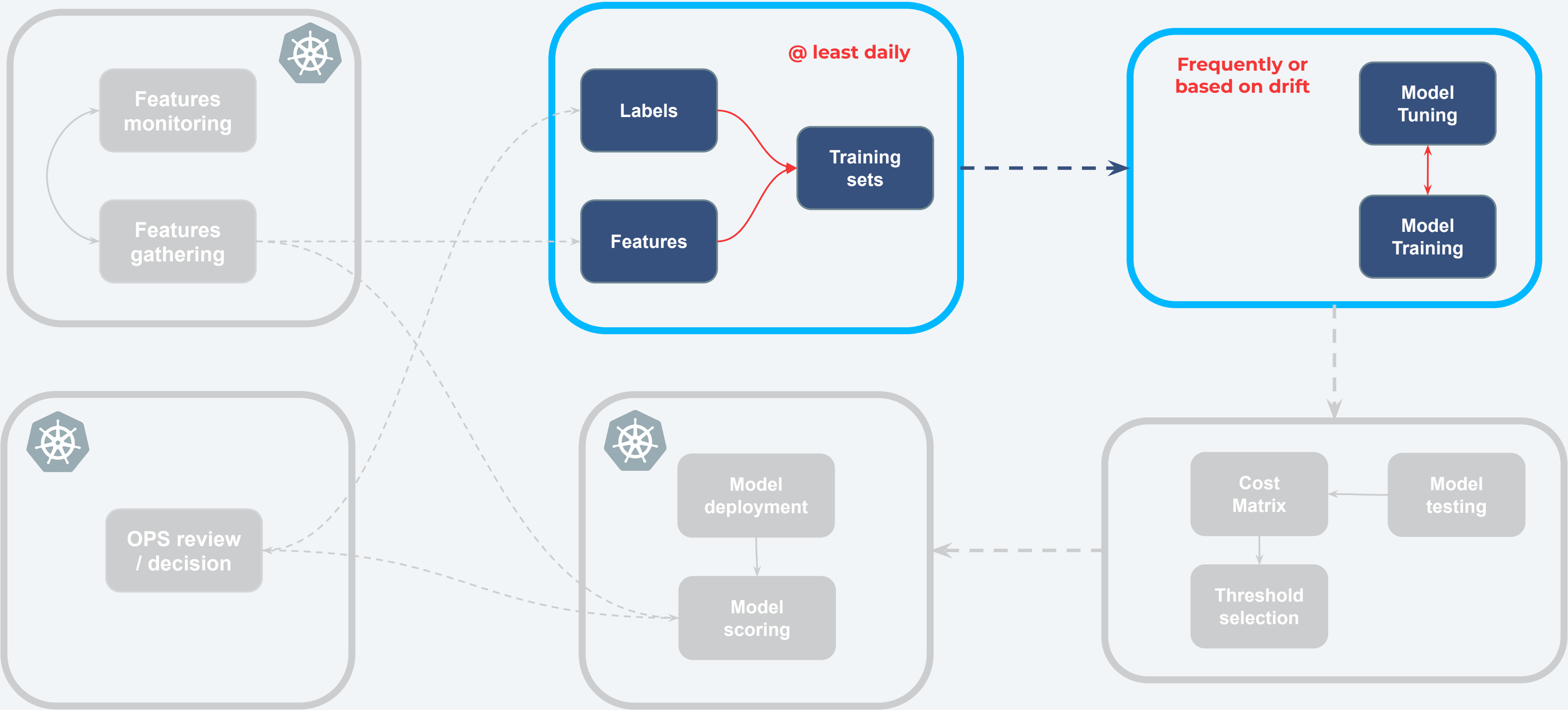
ML system components

1. Feature creation & labelling
2. **Automatic model retraining (*closed-loop learning*)**
3. Model governance
4. Model interpretability

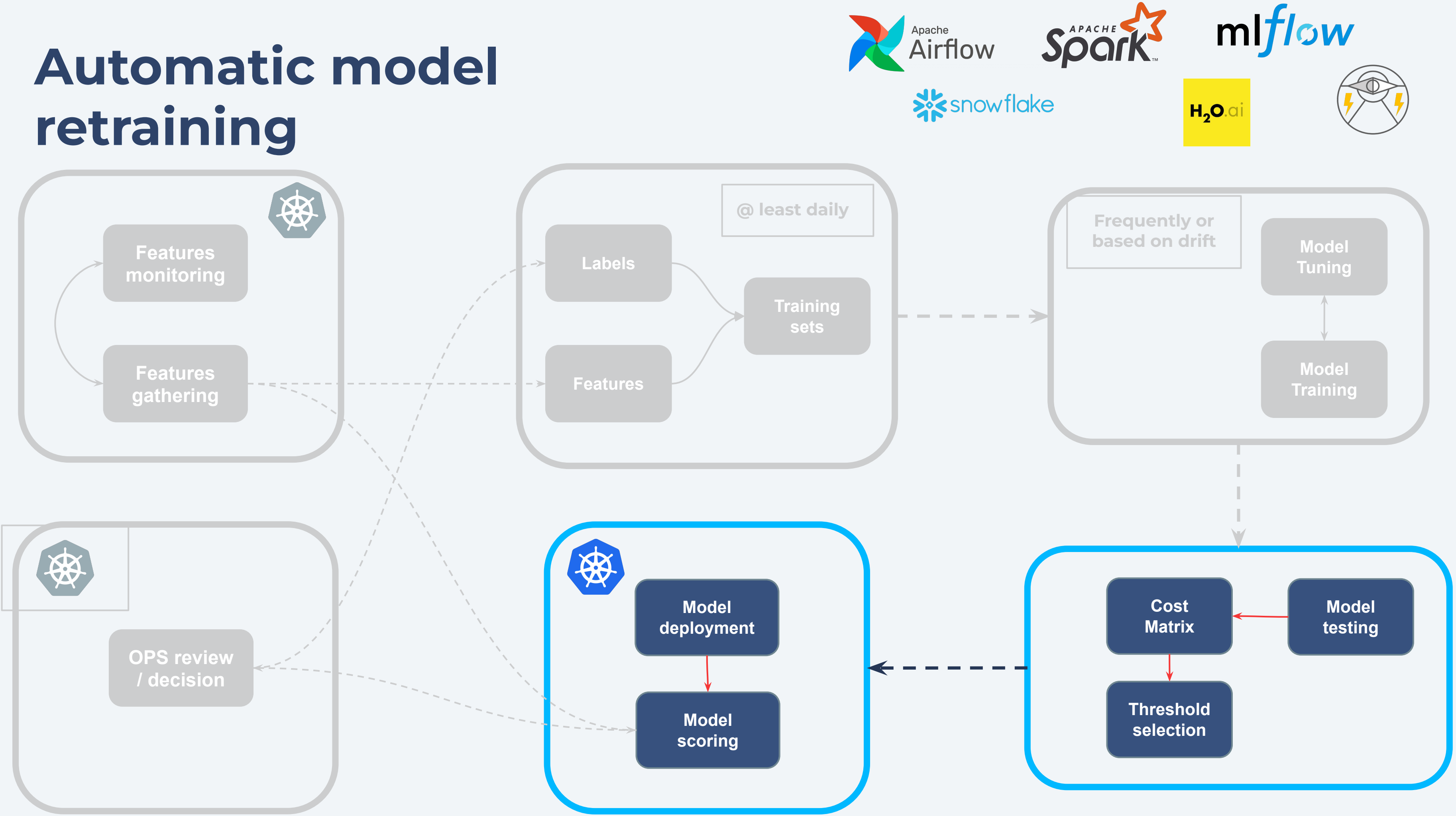
Automatic model retraining



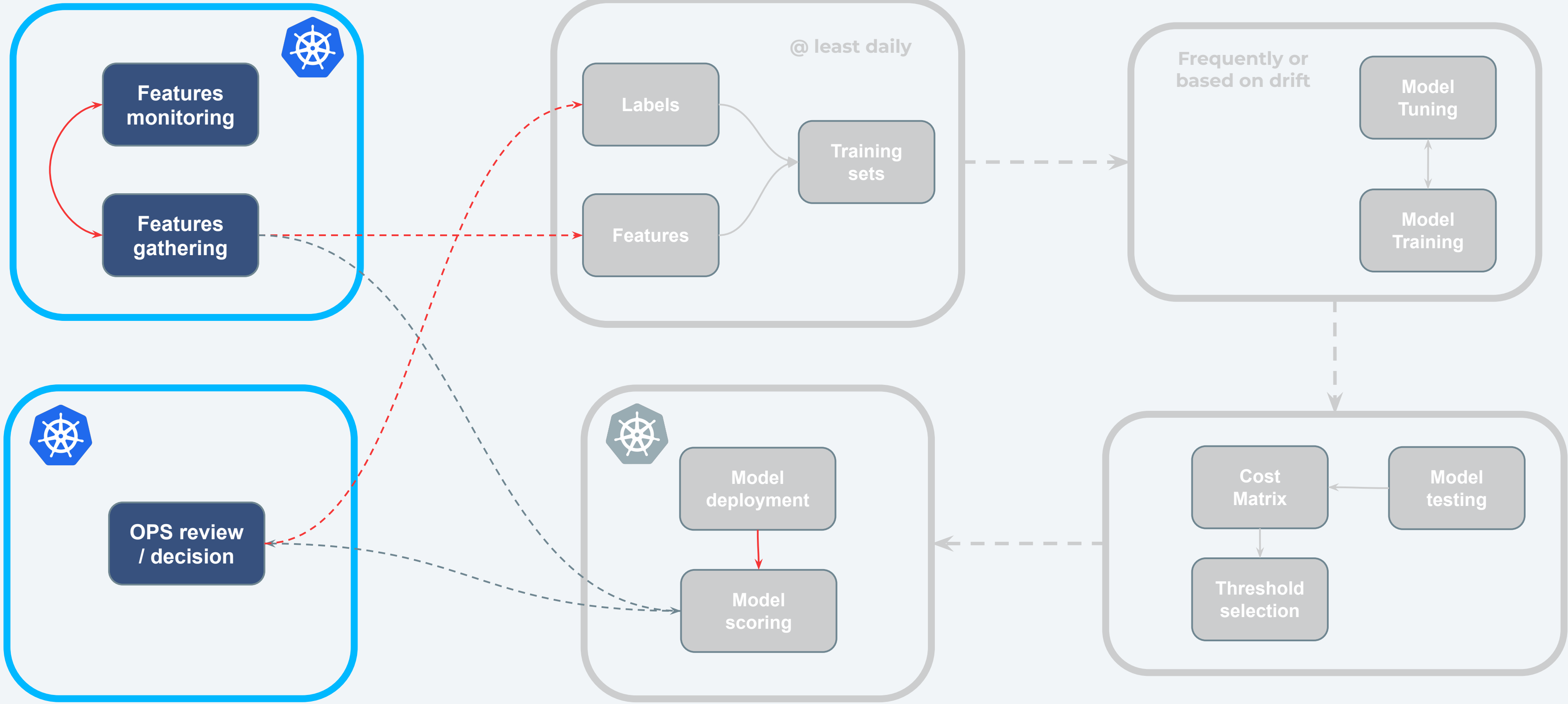
Automatic model retraining



Automatic model retraining



Closed-loop learning

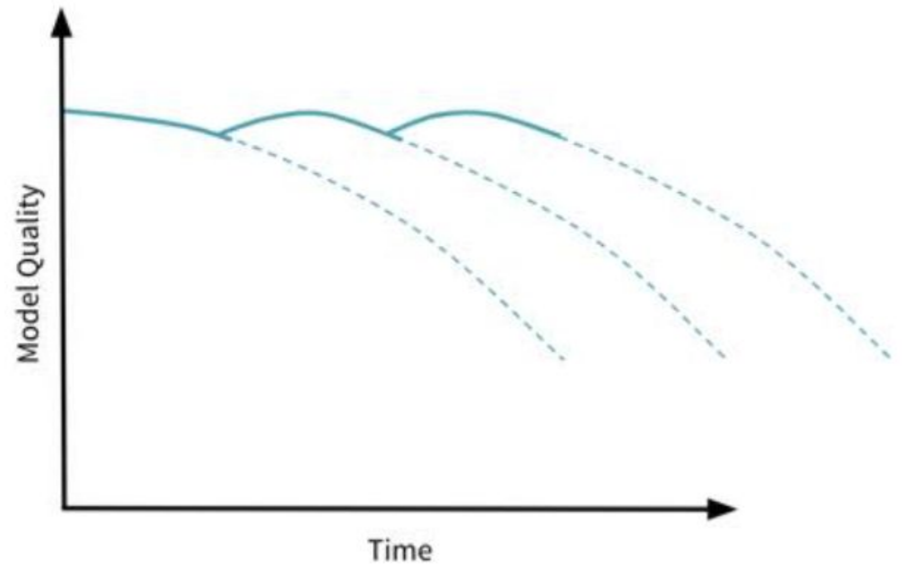


ML system components

1. Feature creation & labelling
2. Automatic model retraining (*closed-loop learning*)
3. **Model governance**
4. Model interpretability

Model governance

- Product coverage
- Risk typology coverage
- Data always up-to-date
- Update models frequently (no do-and-forget attitude)
- Tracking all important model metrics (in prod)
- Features & labels quality monitoring



https://databricks.com/blog/2019/09/18/productionizing-machine-learning-from-deployment-to-drift-detection.html

ML system components

1. Feature creation & labelling
2. Automatic model retraining (*closed-loop learning*)
3. Model governance
4. **Model interpretability**

Model interpretability

Shapley or Lime for ‘local’ interpretability

0.79

Feature influence

Name	Value	Importance
feature #38	10	0.277
feature #2	0	0.240
feature #22	30	0.238
feature #6	0	0.187
feature #45	xyz	0.167
feature #13	0.9	0.163
feature #21	rrr	0.140
feature #23	rrrr	0.070
feature #17	eeee	0.069
feature #31	0.99	0.053
feature #54	0.45	-0.136
feature #9	0.7	-0.150
feature #18	www	-0.153
feature #7	-1	-0.159
feature #28	11	-0.167
feature #5	iiii	-0.235
feature #44	uuu	-0.281
feature #32	0.1	-0.332
feature #78	10	-0.334
feature #1	0	-0.421

Model interpretability

Shapley or Lime for ‘local’ interpretability

Feature influence

Name	Value	Importance
feature #38	10	0.277
feature #2	0	0.240
feature #22	30	0.238
feature #6	0	0.187
feature #45	xyz	0.167
feature #13	0.9	0.163
feature #21	rrr	0.140
feature #23	rrrr	0.070
feature #17	eeee	0.069
feature #31	0.99	0.053
feature #54	0.45	-0.136
feature #9	0.7	-0.150
feature #18	www	-0.153
feature #7	-1	-0.159
feature #28	11	-0.167
feature #5	iiii	-0.235
feature #44	uuu	-0.281
feature #32	0.1	-0.332
feature #78	10	-0.334
feature #1	0	-0.421

Features linked to a specific pattern or behaviour - easier to interpret / explain the model decision

Current challenges .

Challenges

1. Scaling of ML infrastructure (with growth)
2. Features group interpretability (generating narratives)
3. Enhancing features and labels quality monitoring

🙏 Thank you!