OPTILEARN

SPRYFOX

Adaptive Machine Learning via
Multi-Objective Optimisation

TECHNISCHE UNIVERSITÄT DARMSTADT

SPRYFOX

LOEWE
Exzellente Forschung für Hessens Zukunft

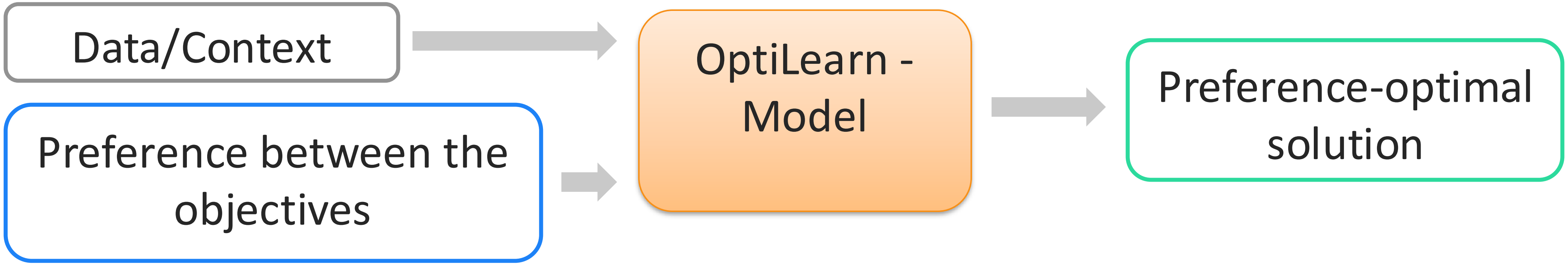
HESSEN
Hessisches Ministerium für Wissenschaft und Forschung, Kunst und Kultur

Topics

- What is Multi-Objective Optimization (MOO)?
- Why do you want it?
- How does it work?
- Case study : Simulations
- Use case : Cost optimal fabric defect classification
- OptiLearn Setting

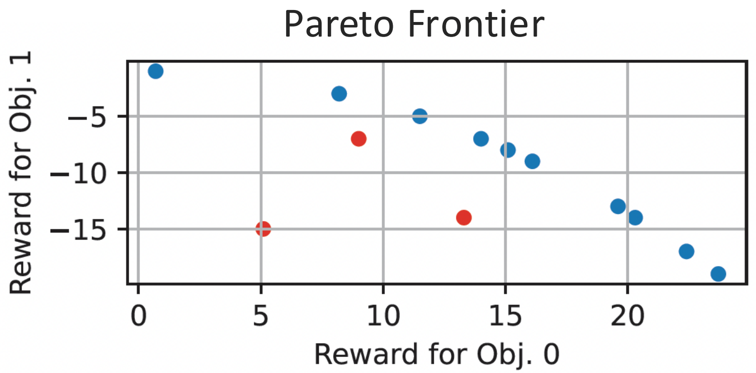
What is Multi-Objective Optimization (MOO)?

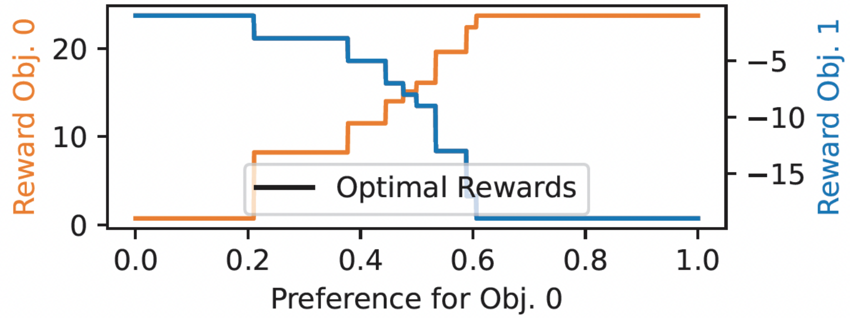
- Optimization for multiple competing objectives (e.g., precision vs. recall)
- Train parameterized model that can change their behavior during runtime



What is MOO?

- In MOO optimality is *situational*
- Depending on the current *preference* the *optimal solution* may diver
- Target: Pareto Frontier





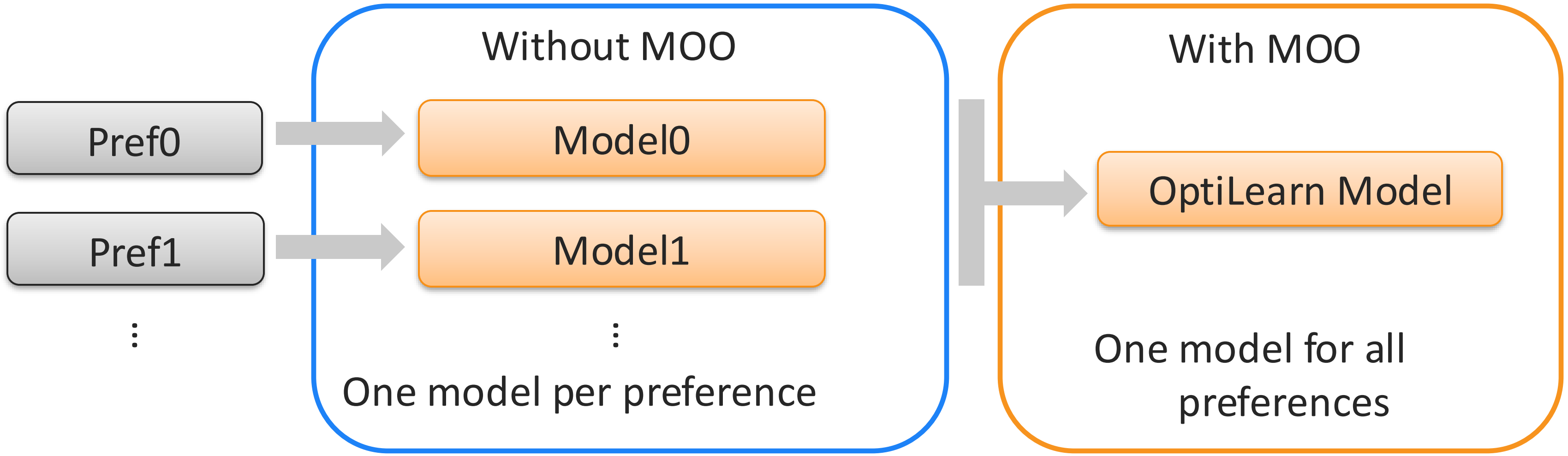
Preference Optimal Solution, S*(Pref):

Utility(Pref, Solution) = Pref***Obj0**(Solution) + (1-Pref)***Obj1**(Solution)

S*(Pref) = argmax[**Utility**(Pref, Solution)]
Solution

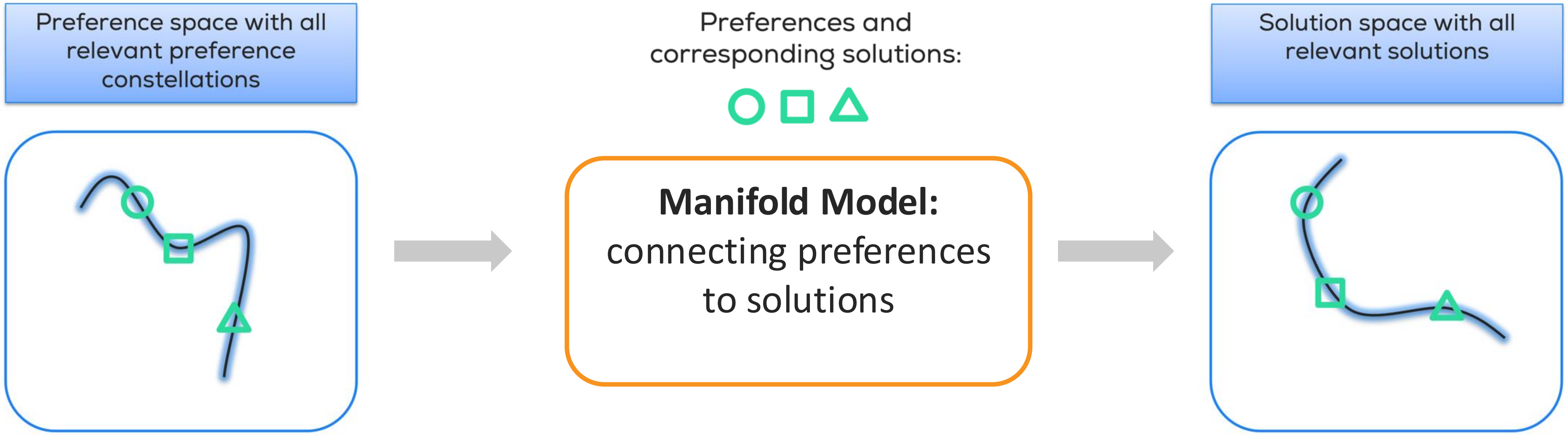
Why do you want it?

- Adjust model behavior to changing circumstances
- Shift focus by changing preferences
- **All that without the need for retraining**



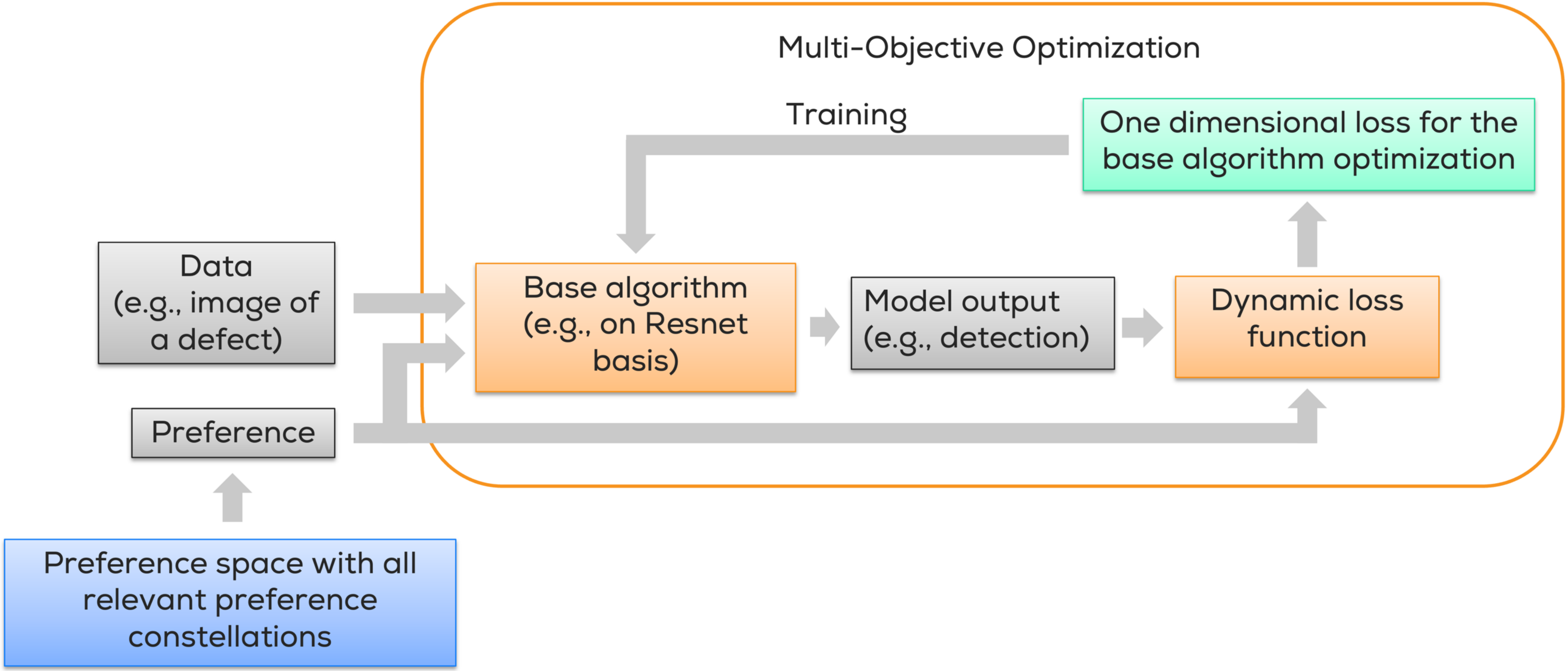
How does it work?

- We train the model on a preference space rather than one or set of preferences
- Generating a parameterized model to form a manifold connecting preference and solution space



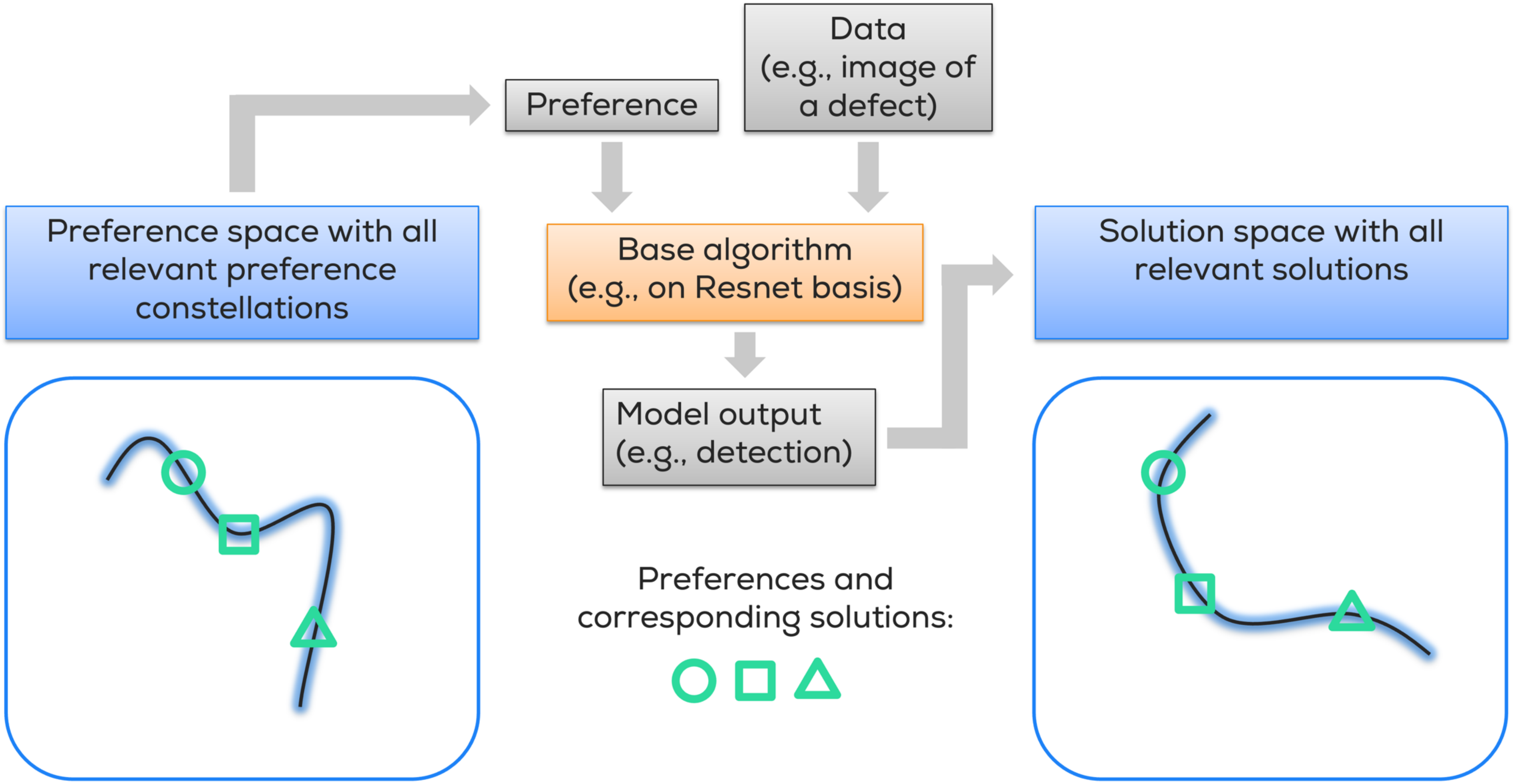
How does it work?

Training



How does it work?

Inference



Case Study

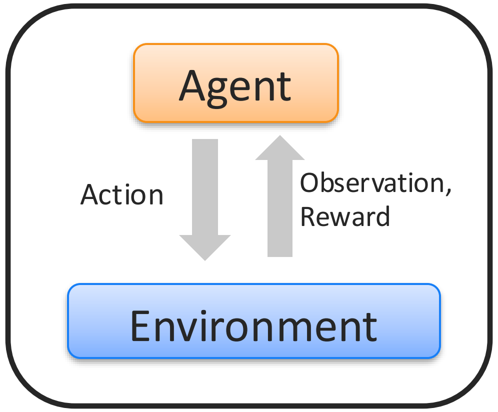
Simulations

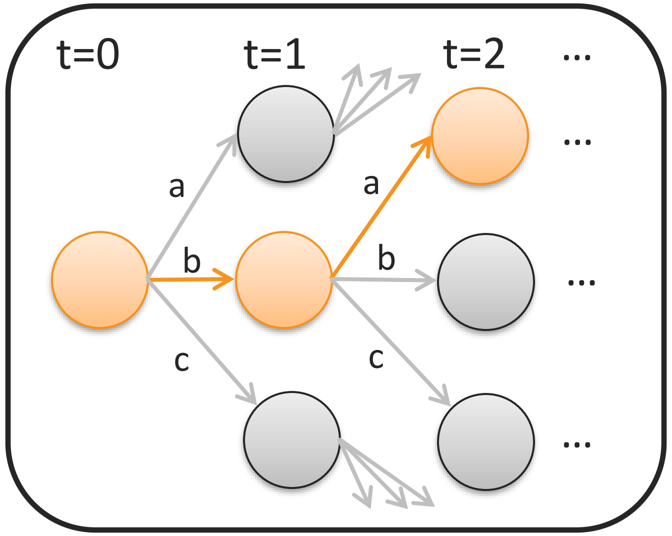
Setting:

- Sequential decision making
- Reinforcement learning setup

Evaluation:

- Simple grid world example
- Complex robot control environments





Objective:
Provide optimal policy for the given preferences.
The model is be used in the field under varying conditions without retraining.

Case Study

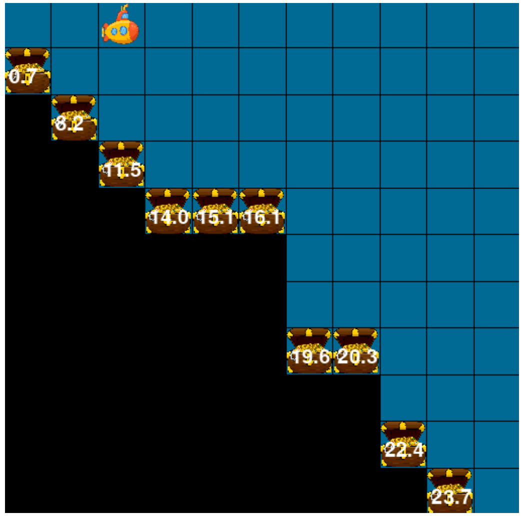
Simulations – Simple Grid World

Deep-Sea-Treasure (DST) is a simple well known multi-objective environment
(Vamplew et al., 2011).

- A grid world with sparse rewards
- The objectives are to maximize the reward value and to minimize the steps to find it
- Experiments are done with different utility functions

$$\text{linear: } \bar{u}(\boldsymbol{r}, \boldsymbol{w}) = \boldsymbol{w}^\top \boldsymbol{r}$$

$$\text{log: } \tilde{u}(\boldsymbol{r}, \boldsymbol{w}) = \sum_{i=0}^{n} \operatorname{sign}(r_i) \log(|r_i| + \epsilon) w_i$$



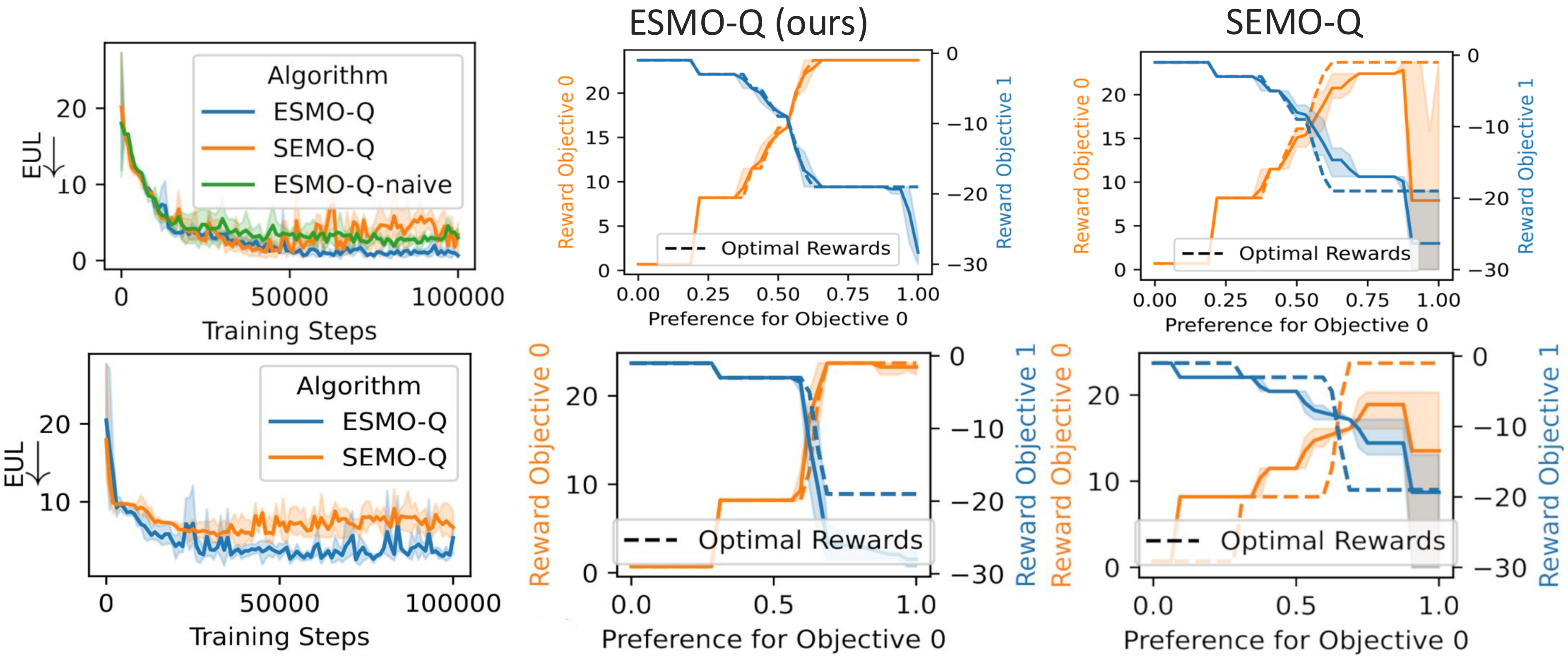
Case Study

Simulations – Simple Grid World

Our algorithm improved on stability and performance for linear and non-linear utility functions

Linear:

Log:

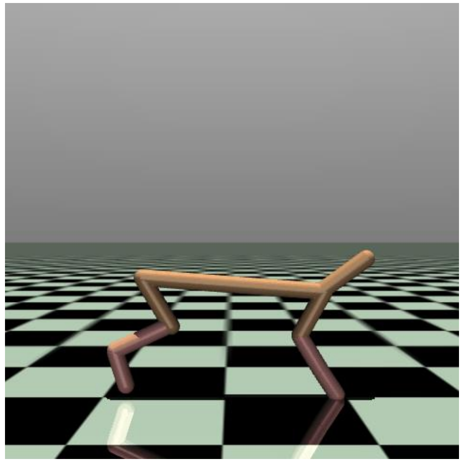


SPRYFOX

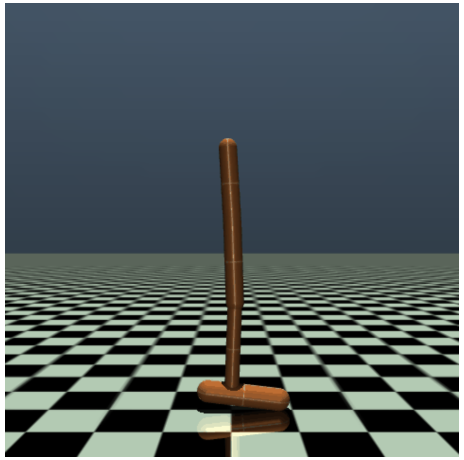
Case Study

Simulations – Robot Control

- Half Cheetah
  - Energy consumption vs. distance travelled
  - State space dimensions: 17
  - Action space dimensions: 6



- Hopper
  - Energy consumption vs. distance travelled
  - State space dimensions: 11
  - Action space dimensions: 3



Both environments are complex control problems and part of the mo-gymnasium library (Alegre et al., 2022).

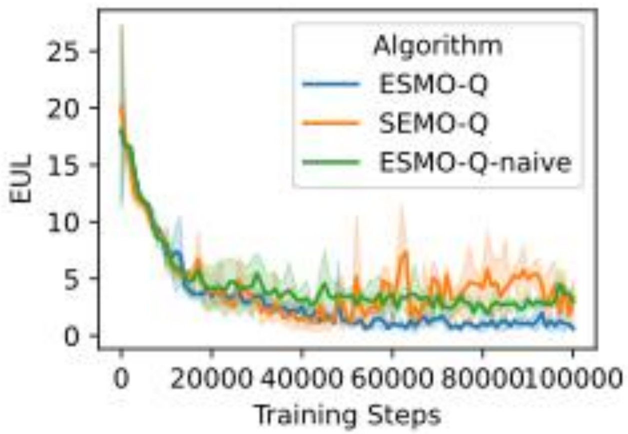
Case Study

Simulations – Robot Control

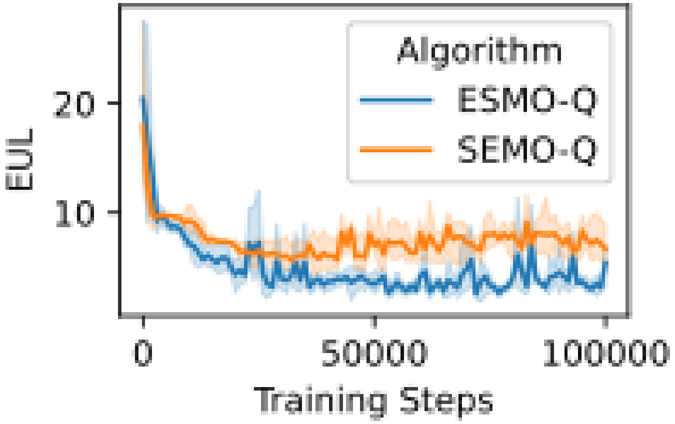
For the complex environment, our algorithm improved on stability and performance

- Expected Utility Loss (EUL) describes the distance to the preference optimal outcome

Linear:



Log:



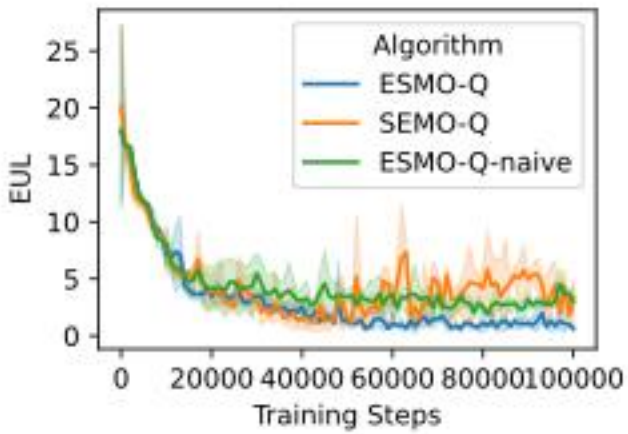
Case Study

Simulations – Robot Control

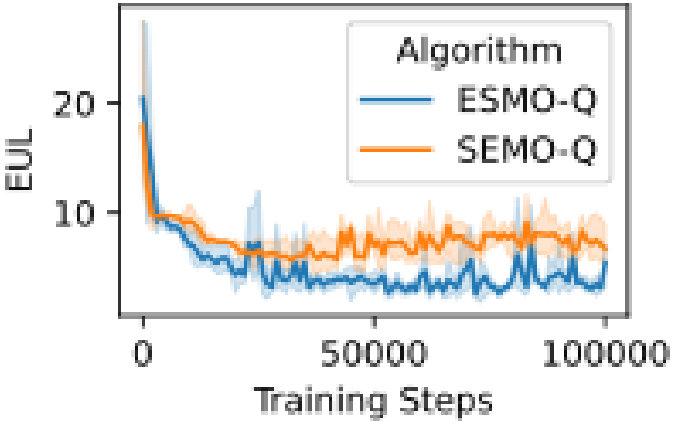
For the complex environment, our algorithm improved on stability and performance

- Expected Utility Loss (EUL) describes the distance to the preference optimal outcome
- Demo Time!

Linear:



Log:



Use case

Textile fabric defect classification

Setting:

- Input: Images of fabric
- Output: defect classifications (holes, stains, etc.)

Problem statement:

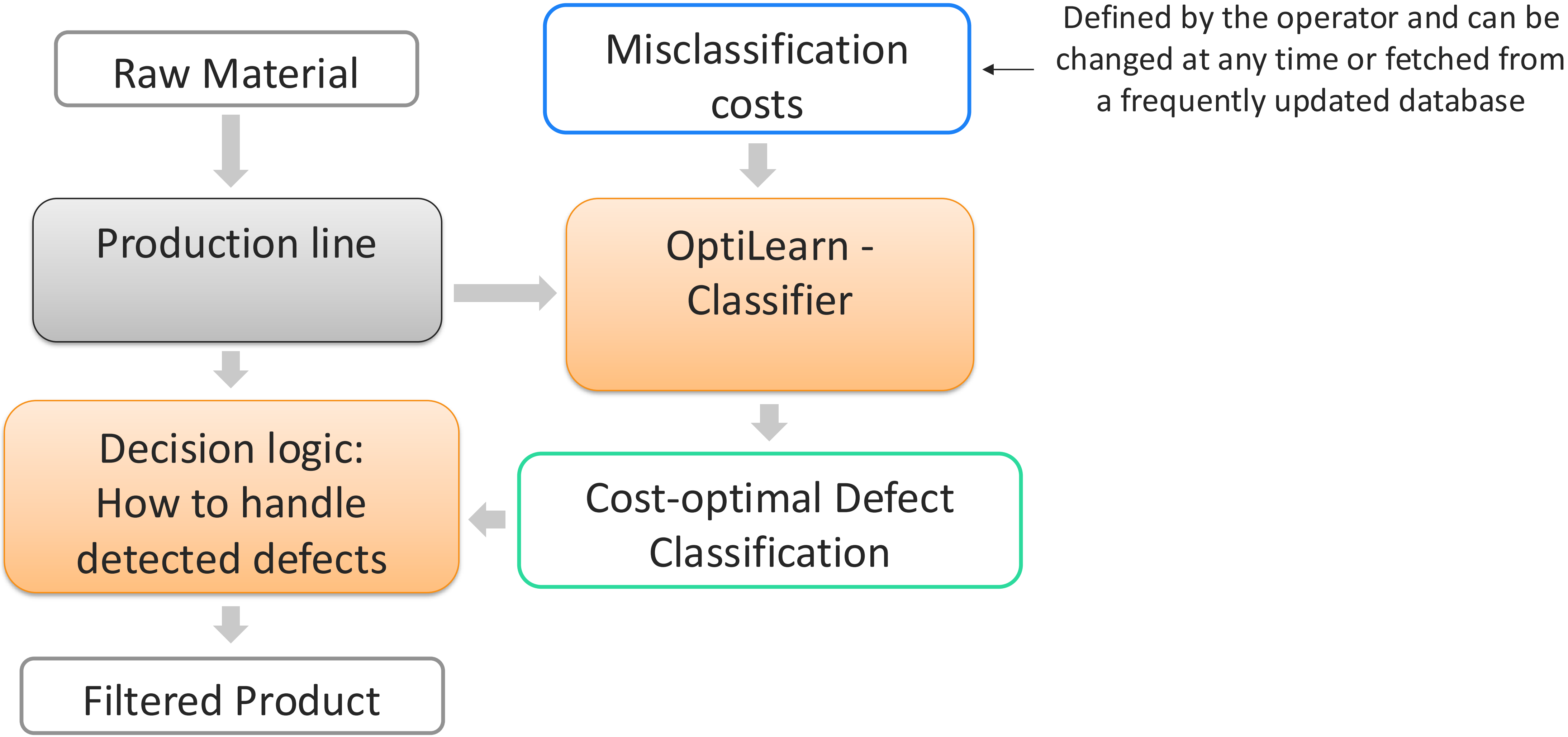
- False alarm costs: interrupt the production for a **false positive**
- Missed defect cost: miss a critical defect (**false negative**)

Objective:

Provide optimal policy for the given costs.
The model is be used in the field under varying conditions without retraining.

Use case

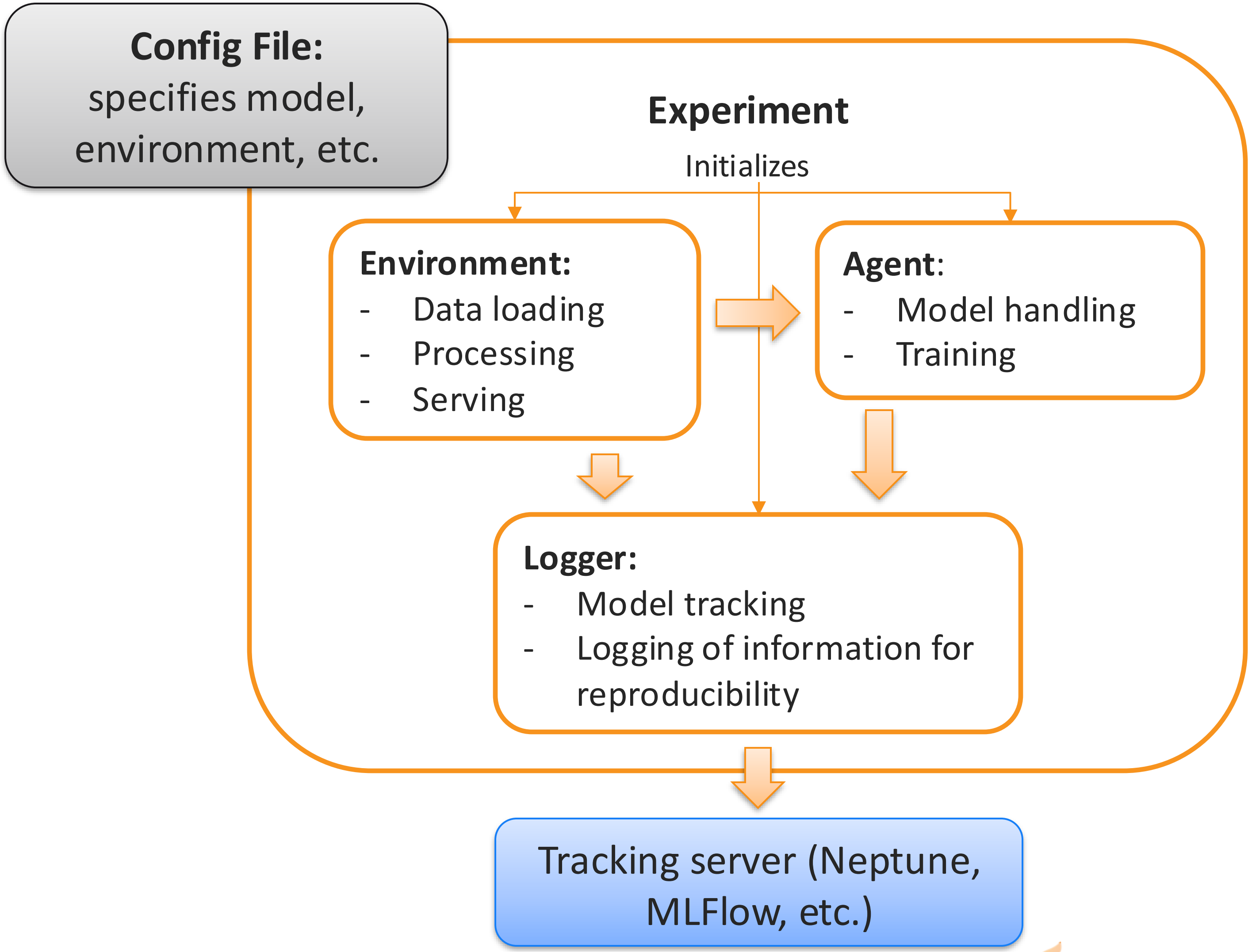
Textile fabric defect classification



OptiLearn Setting

Seamless workflow:

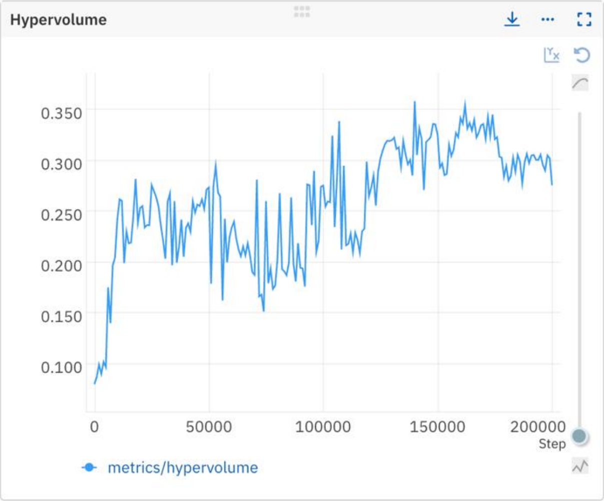
- Training
- Tracking & Evaluation
- Publishing
- Life cycle management

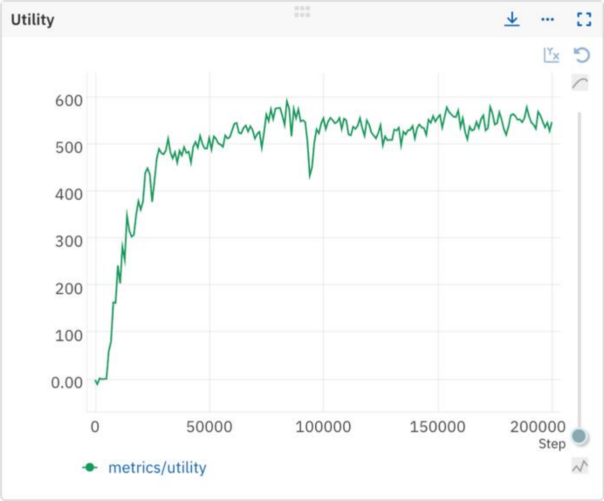


OptiLearn Setting

Seamless workflow:

- Training
- Tracking & Evaluation
- Publishing
- Life cycle management





Id	Stage
OP-CLANNLIN-14	STAGING
OP-CLANNLIN-13	NONE
OP-CLANNLIN-12	NONE
OP-CLANNLIN-11	PRODUCTION

Interest has been sparked?

- You want to build you own solution?
- You want to apply MOO to your use cases?
- You are curious about the math behind all this?

➔ Contact us!

Eike.Mentzendorff@Spryfox.de
Christian.Debes@Spryfox.de

TECHNISCHE UNIVERSITÄT DARMSTADT

SPRYFOX

LOEWE
Exzellente Forschung für Hessens Zukunft

HESSEN
Hessisches Ministerium für Wissenschaft und Kunst

References

- Vamplew, P., Dazeley, R., Berry, A. et al. Empirical evaluation methods for multiobjective reinforcement learning algorithms. Mach Learn 84, 51–80 (2011). https://doi.org/10.1007/s10994-010-5232-5

- Nunes Alegre, L., Felten, F., Talbi, E-G., Danoy, G., Nowé, A., Bazzan, A., & C. da Silva, B. (2022). MO-Gym: A Library of Multi-Objective Reinforcement Learning Environments. In Proceedings of the 34th Benelux Conference on Artificial Intelligence BNAIC/Benelearn 2022 https://bnaic2022.uantwerpen.be/wp-content/uploads/BNAICBeNeLearn_2022_submission_6485.pdf

Acknowledgements

This project (HA project no. 1397/22-102) is financed with funds of LOEWE – Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz, Förderlinie 3: KMU-Verbundvorhaben (State Offensive for the Development of Scientificand Economic Excellence).

Dieses Projekt (HA-Projekt-Nr.: 1397/22-102) wird im Rahmen der Innovationsförderung Hessen aus Mitteln der LOEWE – Landes- Offensive zur Entwicklung Wissenschaftlichökonomischer Exzellenz, Förderlinie 3: KMUVerbundvorhaben gefördert.

SPRYFOX