



Contribution ID: 18

Type: Poster

## Toward a Structured Reinforcement Learning Pipeline for Real-World Systems

Recent advances in reinforcement learning (RL) have shown great potential for managing complex systems in robotics, manufacturing, and beyond. However, translating RL successes from controlled experiments to real-world scenarios remains a significant challenge due to the absence of a standardized engineering pipeline that prioritizes thorough problem formulation. While data science and control engineering benefit from robust workflows, such as CRISP-DM and standard control-design processes, current RL practices often focus narrowly on hyperparameter tuning, overlooking foundational tasks like environmental design and reward specification.

In this work, we introduce an *RL engineering pipeline* that bridges existing gaps by integrating several key components. First, our approach begins with systematic problem identification and the formalization of the Markov Decision Process - defining states, actions, and rewards in alignment with physical constraints. Next, we ensure a careful setup of the environment, optimization objectives, and training procedures that respect the inherent nature of the optimization problem. These steps are followed by iterative agent training, hyperparameter optimization, evaluation, and eventual deployment. By adapting best practices from both CRISP-DM and classical control, our methodology enhances reproducibility and efficiency in RL development.

We validate our framework through a case study on a 1-degree-of-freedom (1-DoF) helicopter testbed. Our experiments indicate that targeted modifications, such as normalizing observations, randomizing initial conditions, extending episode horizons, and incorporating action penalties, yield measurable improvements in sample efficiency and training stability in both simulation and hardware. Moreover, our results suggest that this pipeline can be generalized to more complex systems, paving the way for more robust real-world RL applications.

Overall, our framework not only clarifies design decisions in RL but also offers a promising pathway to overcome long-standing challenges in deploying RL solutions outside of controlled experimental settings.

**Authors:** SCHÄFER, Georg; Dr REHRL, Jakob (Salzburg University of Applied Sciences); HIRLAENDER, Simon (PLUS University Salzburg); Dr HUBER, Stefan (Salzburg University of Applied Sciences)

**Presenter:** SCHÄFER, Georg

**Session Classification:** Poster session