Towards Production-Ready Data Management with Deep Reinforcement Learning

Gabriel Campero Durand, Pavlo Shevchenko, Milena Malysheva, Ivan Prymak, and Gunter Saake

Email: *{firstname.lastname@ovgu.de}*

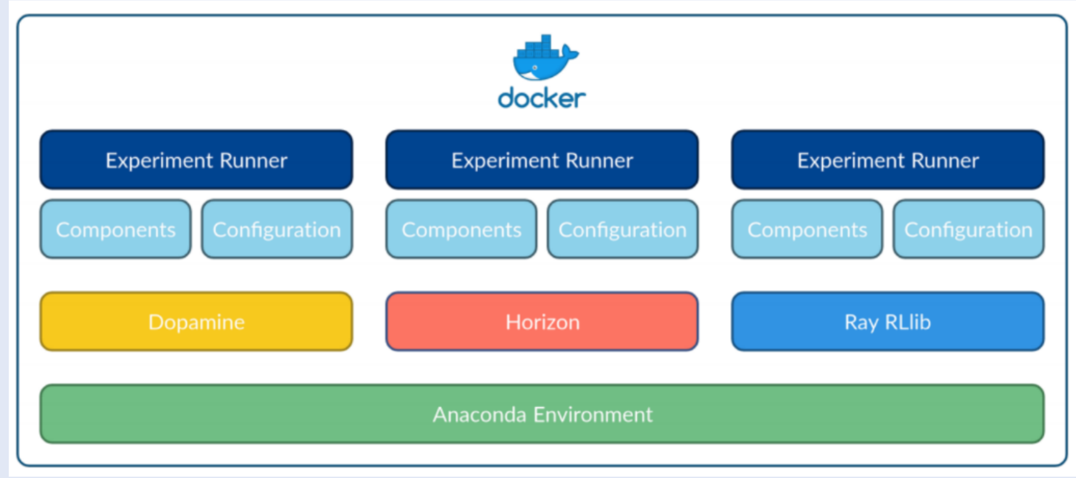
Development process

1. Application-side contribution
2. Problem framing
3. Advanced model understanding
4. Training configuration

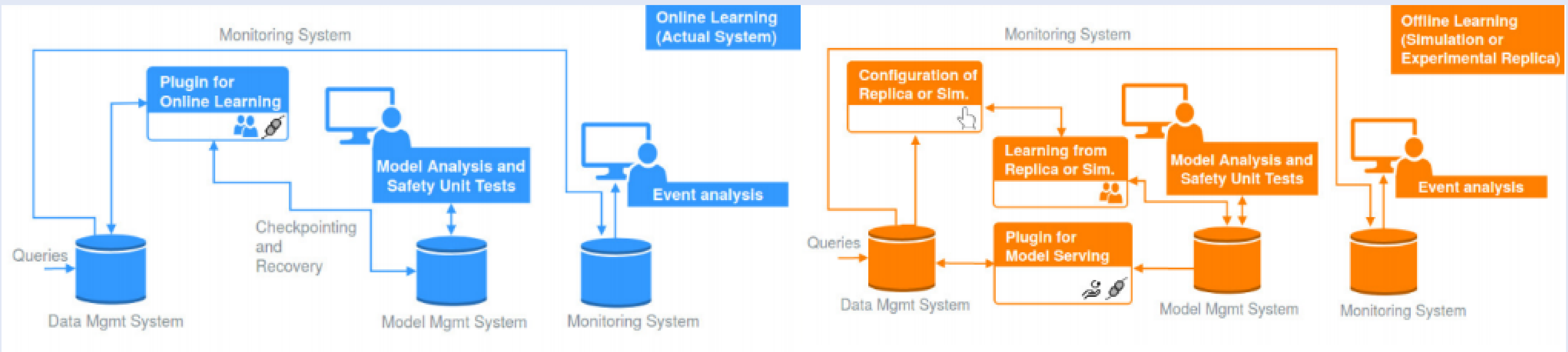
Project Goals

1. **Comparison of frameworks**.
2. **Partitioning with DRL.**
3. Studying how to provide **safety guarantees.**

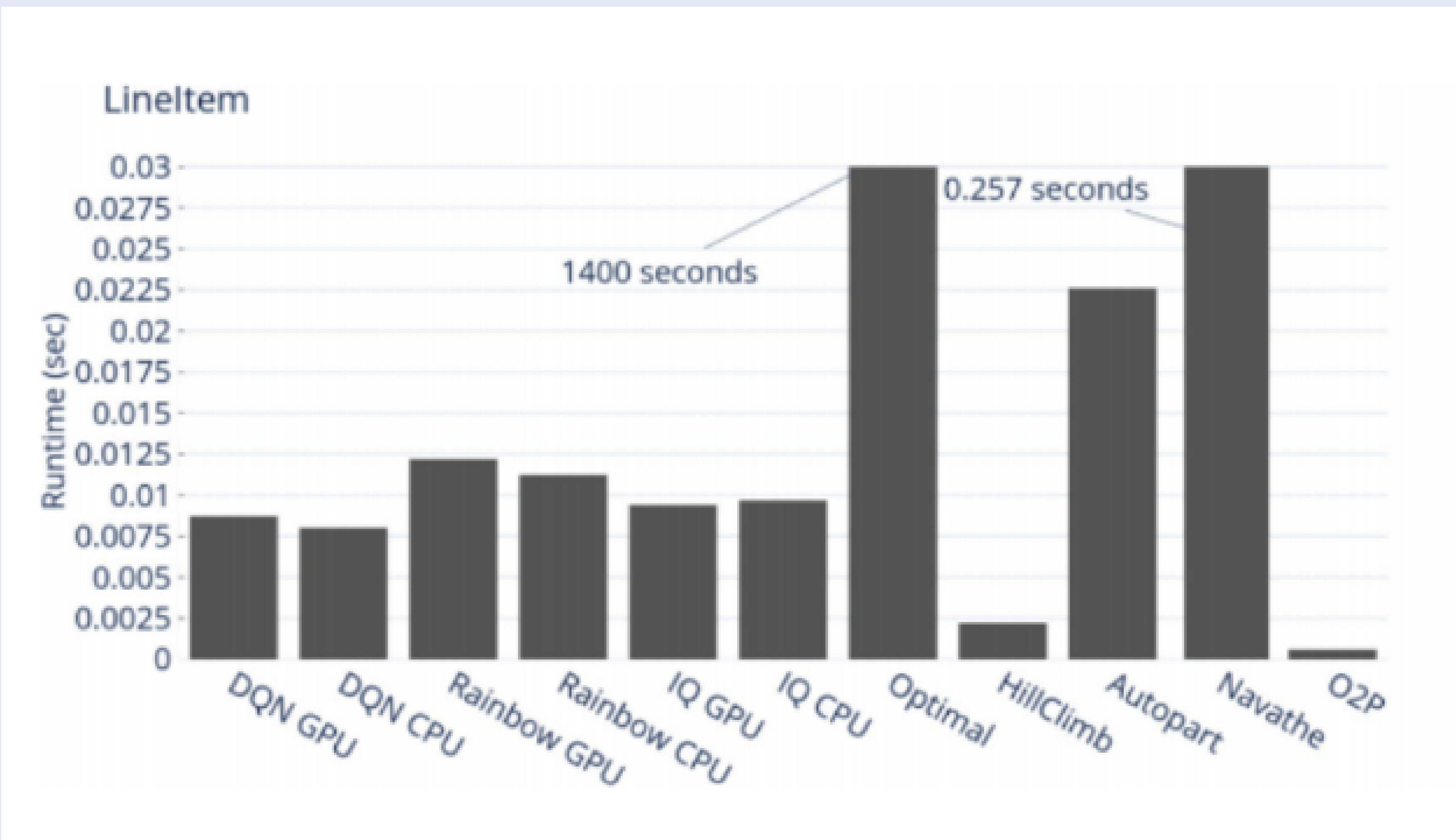
Extensible DRL



Deep Reinforcement Learning in Data Management

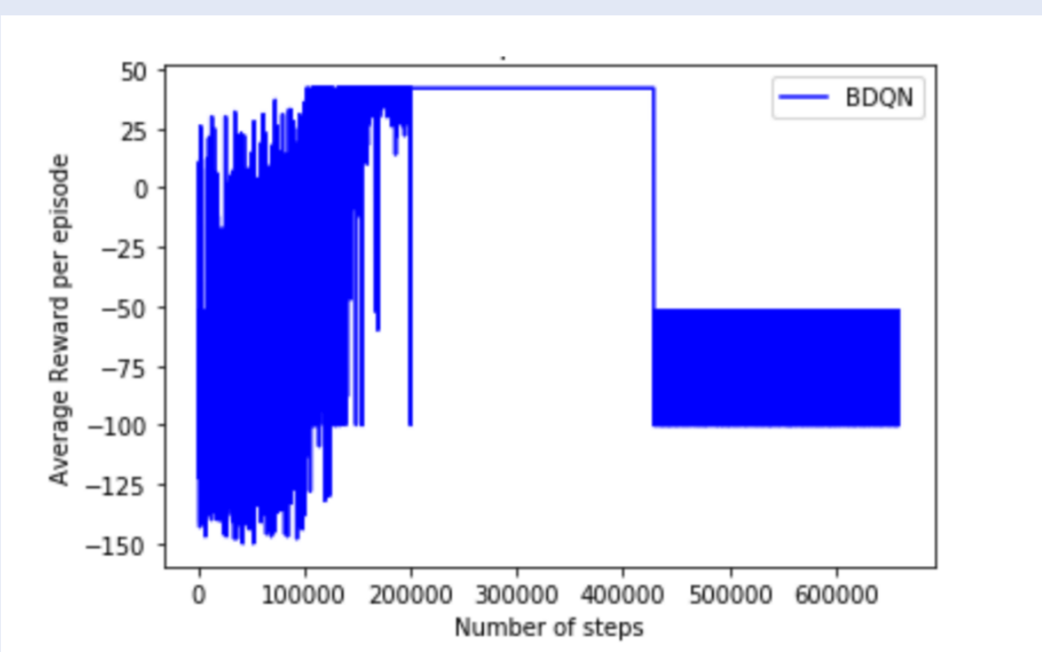


Findings: Partitioning and Comparison



Property	Dopamine	Horizon	Ray RLlib
Framework's focus	Research Education	Production	Distributed learning
Environments	OpenAI Gym User-defined	OpenAI Gym User-defined	OpenAI Gym ELF User-defined
Agents	Value-based	Value-based Policy-based	Value-based Policy-based
Setting	Online	Online Offline	Online Offline
Supported DL libraries	TensorFlow	PyTorch	PyTorch TensorFlow
GPU support	✓	✓	✓
Distributed learning	✗	✓	✓
Code base (LoC)	35 files (4798)	126 files (23,434)	2k files (590,924)
DQN implementation[1]	Convergence: ✗ Runtime rank: 3	Convergence: ✗ Runtime rank: 2	Convergence: ✓ Runtime rank: 1
GPU speedup	2,54%	69,4%	0,3%
Runtime increase (0% to 100% saves)	1,9%	4%	3,9%
Most performant implementation	IQN	DQN	PPO
Number of actions per second (DQN)	CartPole: 1257 QOpt: 97	CartPole: 1385 QOpt: N/A	CartPole: 726 QOpt: 125

Findings: Safety



Safety concerns include **robustness** (adversarial robustness, distribution shift, safe exploration) and **specification** (reward gaming, safe interruptibility, side-effect avoidance) aspects.

Distributional shift requires proper models and is challenging to study.

Future Directions

1. Extensions to our cross-DRL-framework tool. Model understanding and learning from demos. Structured training and hierarchical designs.
2. Production-ready solutions for partitioning and join ordering.
3. Robustness with Bayesian uncertainty estimates.

Acknowledgements

This work was partially funded by the DFG (grant: SA 465/50-1).