Advancing Decision Science: Lessons from the Machine Learning Community

Serdar Kadıoğlu[1, 2]

[2]Department of Computer Science, Brown University, Providence, USA
[1]Artificial Intelligence Center of Excellence, Fidelity Investments, Boston, USA
{firstname.lastname@fmr.com}

Abstract

We explore the synergies between Learning and Reasoning communities as part of the AAAI-25 Constraint Programming and Machine Learning Bridge Program. We focus on software and tools available within both communities to identify similarities and new opportunities. Our ultimate goal is two-fold: i) to stimulate the development of open-source software in Decision Sciences that encompasses several paradigms in constraint and optimization technologies, and ii) to foster further collaboration for tighter integration of learning-based approaches with decision operations.

Introduction

While the Learning community has benefited tremendously from open-source software such as Hugging Face (HuggingFace 2025), which is now hosting 1M+ models, 200K+ datasets, 300K+ demos, PyTorch (Paszke et al. 2019), TensorFlow (Abadi et al. 2015), together with their well-established packages such as NumPy, SciPy, Pandas, and scikit-learn (Pedregosa et al. 2011) among others, the Reasoning community trails behind in establishing a similar software ecosystem and flywheel. This gap hinders the widespread adoption and collaboration of reasoning solvers & software, limiting its potential and impact. The AAAI 2024 Bridge Program on Constraint Programming and Machine Learning offers a unique opportunity to discuss various opportunities to bridge this gap.

To revive the success of software developments established in the learning community, we would like to foster collaboration between researchers and practitioners to promote open-source software development and establish a community-driven ecosystem for reasoning and decision science software. Improving the existing software, technology stack, and standards in decision operations is critical, especially for new participants.

Opportunities to Advance Decision Science

Potential areas to enhance the developer and user experience of Decision Science, similar to achievements in Machine Learning, can include:

Ease of Use and Accessibility

Easy Installation: Similar to the standard one-liner `pip install` in machine learning packages.

Ready-to-Run Environments: Using Docker containers hosted on public registries, e.g., GitHub Container Registry.

Standardized Interfaces: Common interfaces for solvers and models, e.g., similar to ONNX.

Low Code/No Code Environments: Drag-and-drop interfaces for constructing and modifying optimization models without writing code.

Modeling Assistants: Research and development of modeling assistants, similar to code co-pilots, e.g.,(Wasserkrug et al. 2024; Tsouros et al. 2023; Kadıoğlu et al. 2024; Singirikonda, Kadıoğlu, and Uppuluri 2024).

Integrated Environments

Constraint Model Registry: Hosting pre-trained models that can be easily downloaded and fine-tuned on problem-specific data. Base constraint models for common applications in routing, scheduling, supply chain, and resource allocation that can be tailored to specific problems at hand.

Pre-trained Embeddings: Readily available latent instance representations for Constraint Programming (CP), Mixed-Integer Programming (MIP), and Boolean Satisfiability (SAT) instances, potentially hosted as HuggingFace models[1].

Cloud Services and APIs: Online services, platforms, and API integration to deploy, host, run, test, and monitor decision models. The emergence of DecisionOps, akin to MLOps and LLMOps.

API Integration: API integration of decision models with popular programming languages and machine learning frameworks.

Advanced Features and Integration

Explainability and Interpretability: Tools that provide insights into how the optimization model makes decisions, similar to explainable AI techniques. Similarly, features that allow users to understand how changes in input parameters affect the model's output and sensitivity analysis.

[1] https://huggingface.co/models

Interactive Dashboards Tools that allow users to visualize and interact with optimization models and their results in real-time.

Automated Tuning and Model Selection: Hyperparameter optimization and automated tools for tuning the parameters of optimization models to achieve the best performance. Model Selection: Tools that can automatically select the most appropriate optimization model based on the problem characteristics.

Hybrid Software: Integrated tools and frameworks for combining learning and reasoning models.

Resources and Tools

Community-Driven Development: Collaborative development, maintenance, and events to bring participants together, including hackathons and competitions (Kadıoğlu and Kleynhans 2024). Mentorship programs to help new users and developers learn from experienced practitioners. Certification programs to validate the skills and knowledge of users in optimization and decision science.

Datasets and Leaderboards Publicly accessible datasets with data cards, models complete with model cards, benchmarks, leaderboards, and evaluations[2].

Educational Resources: Publicly available educational material for constraint and optimization solvers, e.g., interactive tutorials, blog posts, and online courses.

Showcasing Successful Applications

The Art of Possible: Demos and proof-of-concept application spaces to communicate the art of possible, as in HuggingFace Spaces[3].

Industry Applications: Showcase successful industry applications with practical impact addressing unique needs of sectors such as healthcare, finance, and manufacturing.

Call to Action

We stress the need for a community-driven ecosystem for reasoning and decision science software, similar to what has been achieved in the machine learning community. As such, let us call on researchers, practitioners, and developers to contribute to developing a robust and community-driven ecosystem that allows the creation of a data and model flywheel to achieve network effect at scale. Improving the technology stack, standards, and user experience allows us to attract new participants, foster collaboration, and advance decision science.

References

Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G. S.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Goodfellow, I.; Harp, A.; Irving, G.; Isard, M.; Jia, Y.; Jozefowicz, R.; Kaiser, L.; Kudlur, M.; Levenberg, J.; Mané, D.; Monga, R.; Moore, S.; Murray, D.; Olah, C.; Schuster, M.; Shlens, J.; Steiner, B.; Sutskever, I.; Talwar, K.; Tucker, P.; Vanhoucke, V.; Vasudevan, V.; Viégas, F.; Vinyals, O.; Warden, P.; Wattenberg, M.; Wicke, M.; Yu, Y.; and Zheng, X. 2015. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. Software available from tensorflow.org.

HuggingFace. 2025. Hugging Face, Inc. https://huggingface.co/. Accessed: 2025-02-02.

Kadıoğlu, S.; Pravin Dakle, P.; Uppuluri, K.; Politi, R.; Raghavan, P.; Rallabandi, S.; and Srinivasamurthy, R. 2024. Ner4Opt: named entity recognition for optimization modelling from natural language. *Constraints*, 1–39.

Kadıoğlu, S.; and Kleynhans, B. 2024. The Design and Organization of Educational Competitions with Anonymous and Real-Time Leaderboards in Academic and Industrial Settings. arXiv:2402.07936.

Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; et al. 2019. Pytorch: An imperative style, high-performance deep learning library. In *Advances in neural information processing systems*, 8026–8037.

Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; and Duchesnay, E. 2011. Scikit-learn: Machine Learning in Python. *JMLR*, 12: 2825–2830.

Singirikonda, A.; Kadıoğlu, S.; and Uppuluri, K. 2024. TEXT2ZINC: A Cross-Domain Dataset for Modeling Optimization and Satisfaction Problems in MiniZinc.

Tsouros, D.; Verhaeghe, H.; Kadioglu, S.; and Guns, T. 2023. Holy Grail 2.0: From Natural Language to Constraint Models. *arXiv preprint arXiv:2308.01589*.

Wasserkrug, S.; Boussioux, L.; den Hertog, D.; Mirzazadeh, F.; Birbil, I.; Kurtz, J.; and Maragno, D. 2024. From Large Language Models and Optimization to Decision Optimization CoPilot: A Research Manifesto. arXiv:2402.16269.

[2] https://huggingface.co/docs/leaderboards

[3] https://huggingface.co/spaces