Metadata Representations for Queryable Repositories of Machine Learning Models

Ziyu Li[1]
Supervisors: Alessandro Bozzon [1], Rihan Hai [1], and Asterios Katsifodimos [1]

[1] Department of Software Technology, Delft University of Technology, Netherlands
z.li-14,{initial.surname}@ugent.be

Machine learning (ML) practitioners and organizations are building model repositories of pre-trained models, referred to as model zoos. Examples include HuggingFace, TensorflowHub, and PyTorch Hub[1]. These model zoos contain metadata describing the properties of the ML models and datasets. The metadata serves crucial roles for reporting, auditing, ensuring reproducibility, and enhancing interpretability. Despite the growing adoption of descriptive formats like datasheets and model cards, the metadata available in existing model zoos remains notably limited. Moreover, existing formats have limited expressiveness, thus constraining the potential use of model repositories, extending their purpose beyond mere storage for pre-trained models.

In this work, we advocate for expressive metadata representation for model zoos. Beyond the current state-of-the-art (and practices) [1, 2], we propose a metadata format that can capture information about relevant artifacts (e.g., models, datasets, data instances, training configurations, evaluation) and their relationships. We also describe the design and current implementation of an advanced ML models management platform called *Macaroni* [3][2] that can be used to query and make use of such metadata. Our proposed metadata representations cover various categories of meatadata and can support different ML applications, e.g., optimization of ML inference queries [4]

References

[1] Sebastian Schelter, Joos-Hendrik Boese, Johannes Kirschnick, Thoralf Klein, and Stephan Seufert. Automatically tracking metadata and provenance of machine learning experiments. In *Machine Learning Systems Workshop at NIPS*, pages 27–29, 2017.

[2] Pulkit Agrawal, Rajat Arya, Aanchal Bindal, Sandeep Bhatia, Anupriya Gagneja, Joseph Godlewski, Yucheng Low, Timothy Muss, Mudit Manu Paliwal, Sethu Raman, et al. form for machine learning. In *Proceedings of the 2019 SIGMOD*, pages 1803–1816, 2019.

[3] Ziyu Li, Henk Kant, Rihan Hai, Asterios Katsifodimos, and Alessandro Bozzon. Macaroni: Crawling and enriching metadata from public model zoos. In *ICWE*, pages 376–380. Springer, 2023.

[4] Ziyu Li, Mariette Schönfeld, Wenbo Sun, Marios Fragkoulis, Rihan Hai, Alessandro Bozzon, and Asterios Katsifodimos. Optimizing ml inference queries under constraints. In *ICWE*, pages 51–66. Springer, 2023.

[1] `https://huggingface.co/`, `https://www.tensorflow.org/`, `https://pytorch.org/hub/`
[2] Prototype available at `https://sites.google.com/view/macaroni-model-zoo/home`