Towards a Definition of Quality-of-Service Metrics for Symbolic Knowledge Injection Systems

Andrea Agiollo[a] *Andrea Rafanelli*[b,c] Andrea Omicini[a]

[a]Dipartimento di Informatica – Scienza e Ingegneria (DISI)
Alma Mater Studiorum – Università di Bologna
{*andrea.agiollo*, andrea.omicini}@unibo.it

[b]Dipartimento di Informatica – Università di Pisa

[c]Dipartimento di Informatica – Scienza e Ingegneria e Matematica (DISIM) – Università dell'Aquila
andrea.rafanelli@phd.unipi.it

23rd Workshop "From Objects to Agents",
September 01–03, 2022, Genova, Italy

Next in Line. . .

Symbolic and Sub-symbolic

Why?

- DL and ML approaches pervade our daily life
- Very complex internal processes $\Rightarrow$ interpretability issues.
- High quantity of data and training time $\Rightarrow$ not sustainable.

Benefits

- More explainable and interpretable models
- More reliable models, following a-priori unbreakable knowledge
- Less data and time to train a model
- Debugging opportunities for ML models

Symbolic Knowledge Injection

What is?

Any procedure affecting how sub-symbolic predictors draw their inferences s.t. predictions are computed as a function of, or consistent with, given symbolic knowledge.

How?

- *Logical constraint*, codification of knowledge within the loss function
- *Structural constraint*, codification of knowledge within the structure of ML model
- *Knowledge embedding*, embedding of knowledge via vector-based representation

Symbolic Knowledge Injection

Mathematically, we refer to the knowledge-aware model as:

$$\mathcal{N}^{ski}(\mathcal{K}, \mathcal{I}, \tau) \tag{1}$$

where $\mathcal{I}$ is the injection procedure, $\mathcal{K}$ is the knowledge base, $\mathcal{N}$ is the sub-symbolic model, and τ is the task to solve.

Its *uneducated* counterpart is indicated as:

$$\mathcal{N}(\tau) \tag{2}$$

Motivations

How to assess the overall quality of injection mechanism?

$$\mathcal{E}(\mathcal{I}) = \mathcal{P}(\mathcal{N}^{ski}, \tau) - \mathcal{P}(\mathcal{N}, \tau) \tag{3}$$

where $\mathcal{P}(i, j)$ is the performance measure of a model i over a task j.

Artefacts:

- **Knowledge quality and coverage**
- **Baseline mechanism quality**
- **Task at hand**

Contribution of the paper

1. The first set of reliable – artefacts-free – performance metrics for SKI
2. Focus on injection quality, and injection efficiency
3. Taxonomy proposal

Next in Line...

Overview

- **Injection quality metrics**: analyse quality achieved by the injection procedure;
- **Efficiency metrics**: analyse efficiency gains obtained with injection procedure;

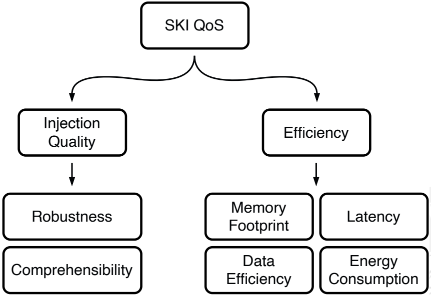

Figure: Classification of the proposed QoS metrics.

Robustness

The capability of the injection mechanism to adapt to variations of input data or knowledge base.

Intuition

An injection mechanism is robust if its prediction ability is not altered much when slight perturbations are introduced.

Formulation

$$M_r = \frac{1}{|\mathcal{P}(\mathcal{N}^{ski}(e), \tau) - \mathcal{P}(\mathcal{N}^{ski}(e_p), \tau)|} \quad \text{s.t.} \quad e_p = e + \alpha \quad \textit{with} \quad |\alpha| \leq \epsilon \tag{4}$$

Robustness

There are different possible perturbations:

1. **Injected knowledge perturbation**

$$\mathcal{K}_p = \mathcal{K} + \alpha \qquad \alpha \sim \mathcal{U}(a, b) \tag{5}$$

2. **Input perturbation**

$$x_p = x + \alpha \qquad \alpha \sim \mathcal{U}(a, b) \tag{6}$$

Comprehensibility

The capability of the injection mechanism to produce more intelligible models.

Intuition

- The introduction of symbolic knowledge within sub-symbolic models $\Rightarrow$ more comprehensible models
- *It is difficult to establish an actual metric of comprehension*

How to measure system's comprehensibility?

- Use extraction mechanisms, such as *decision trees*, and try to evaluate their comprehensibility
  1. **Assess model's complexity** (depth of the tree, or the number of the nodes) [Confalonieri et al., 2021]
  2. **Assess model's comprehensibility** through human feedbacks [Piltaver et al., 2016]

Memory Footprint

The capability of injection mechanisms to produce lightweight sub-symbolic models

Intuition

Injected knowledge lifts part of the learning burden, reducing amount of notions to be learnt data-drivenly

$$\Downarrow$$

Possibility of shrinking sub-symbolic model (FLOPs, MACs, Bytes, etc.)

Formulation

$$\begin{aligned} M_m = & \Psi\left(\mathcal{N}^{ski}(\mathcal{K}, \mathcal{I}, \tau)\right) - \Psi(\mathcal{N}(\tau)) \\ \text{s.t.} \quad & \mathcal{P}(\mathcal{N}^{ski}, \tau) \geq \mathcal{P}(\mathcal{N}, \tau) \end{aligned} \tag{7}$$

Energy Consumption

The capability of injection mechanisms to produce energy-friendly sub-symbolic models

Intuition

- Resource hungry popular AI solutions $\Rightarrow$ not sustainable
- Knowledge injection reduces learning complexity $\Rightarrow$ reduces amount of computations for training and running
- Model *training* and *deployment* are the most resource hungry

Formulation

$$\begin{aligned} M_e = & \Upsilon_t \left(\mathcal{N}^{ski}\left(\mathcal{K},\mathcal{I},\tau\right)\right) + \alpha\Upsilon_d \left(\mathcal{N}^{ski}\left(\mathcal{K},\mathcal{I},\tau\right)\right) - \\ & - \left[\Upsilon_t(\mathcal{N}(\tau)) + \alpha\Upsilon_d(\mathcal{N}(\tau))\right] \\ \text{s.t.} \quad & \mathcal{P}(\mathcal{N}^{ski},\tau) \geq \mathcal{P}(\mathcal{N},\tau) \end{aligned} \tag{8}$$

Latency

The capability of injection mechanisms to produce faster sub-symbolic models

Intuition

- Latency crucial in AI models for time-costly scenarios and for collaboration between multiple intelligent entities (MAS)
- Injection remove unnecessary computations $\Rightarrow$ reduce latency
- SKI may introduce delays linked with analysis of KB (e.g., grounding issues [Tsamoura et al., 2020]) $\Rightarrow$ increase latency

Formulation

$$\begin{aligned} M_l = & \mathcal{T}(\mathcal{N}^{ski}(\mathcal{K}, \mathcal{I}, \tau)) - \mathcal{T}(\mathcal{N}(\tau)) \\ \text{s.t.} \quad & \mathcal{P}(\mathcal{N}^{ski}, \tau) \geq \mathcal{P}(\mathcal{N}, \tau) \end{aligned} \tag{9}$$

Data Efficiency

The capability of injection mechanisms to produce data-frugal sub-symbolic models

Intuition

- Popular DL approaches are data inefficient
- Injection mechanisms play a significant role in data-frugal proposals [Xu et al., 2018]
- Several concepts are injected automatically $\Rightarrow$ Portions of training data are not required

Formulation

$$\begin{aligned} M_d = & \mathcal{D}(\mathcal{N}(\tau)) - \mathcal{D}(\mathcal{N}^{ski}(\mathcal{K}, \mathcal{I}, \tau)) \\ \text{s.t.} \quad & \mathcal{P}(\mathcal{N}^{ski}, \tau) \geq \mathcal{P}(\mathcal{N}, \tau) \end{aligned} \tag{10}$$

Taxonomy

- Efficiency QoS metrics obviously related with AI sustainability field
- Increased robustness and comprehensibility relate SKI to XAI easing the explanation/understanding process
- Simpler models (*memory footprint* and *data efficiency*) obtained using SKI should be easier to analyse, understand and explain *rightarrow* XAI

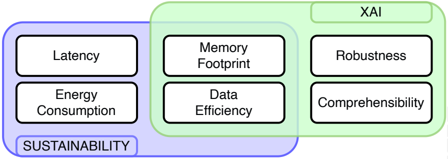

Figure: Taxonomy of the proposed QoS metrics.

Next in Line...

Conclusions & future works

Summing up

- We propose a set of QoS metrics for SKI, aiming to overcome shortcomings of measuring only accuracy improvements in SKI
- Distinguish between injection quality and efficiency metrics
- We identify a taxonomy for the proposed QoS metrics, analysing relations with XAI and AI sustainability

Future works

- Test efficacy of proposed QoS metrics, benchmarking state-of-the-art SKI mechanisms
- Develop a SKI-QoS library and make it available as a tool to automatically benchmark SKI mechanisms

Towards a Definition of Quality-of-Service Metrics for Symbolic Knowledge Injection Systems

Andrea Agiollo[a] *Andrea Rafanelli*[b,c] Andrea Omicini[a]

[a]Dipartimento di Informatica – Scienza e Ingegneria (DISI)
Alma Mater Studiorum – Università di Bologna
{*andrea.agiollo*, andrea.omicini}@unibo.it

[b]Dipartimento di Informatica – Università di Pisa

[c]Dipartimento di Informatica – Scienza e Ingegneria
e Matematica (DISIM) – Università dell'Aquila
andrea.rafanelli@phd.unipi.it

23rd Workshop "From Objects to Agents",
September 01–03, 2022, Genova, Italy

References I

[Confalonieri et al., 2021] Confalonieri, R., Weyde, T., Besold, T. R., and Martín, F. M. d. P. (2021).

Using ontologies to enhance human understandability of global post-hoc explanations of black-box models.

Artif. Intell., 296:103471

DOI:10.1016/j.artint.2021.103471.

[Piltaver et al., 2016] Piltaver, R., Luštrek, M., Gams, M., and Martinčić-Ipšić, S. (2016).

What makes classification trees comprehensible?

Expert Systems with Applications, 62:333–346

DOI:https://doi.org/10.1016/j.eswa.2016.06.009.

References II

[Tsamoura et al., 2020] Tsamoura, E., Gutiérrez-Basulto, V., and Kimmig, A. (2020).
Beyond the grounding bottleneck: Datalog techniques for inference in probabilistic logic programs.
In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pages 10284–10291. AAAI Press
https://ojs.aaai.org/index.php/AAAI/article/view/6591.

[Xu et al., 2018] Xu, J., Zhang, Z., Friedman, T., Liang, Y., and Van den Broeck, G. (2018).
A semantic loss function for deep learning with symbolic knowledge.
In Dy, J. G. and Krause, A., editors, *Proceedings of the 35th International Conference on Machine Learning (ICML), Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018*, volume 80 of *Proceedings of Machine Learning Research*, pages 5498–5507. PMLR
http://proceedings.mlr.press/v80/xu18h.html.