

Twitter Thread by Jerome Pesenti



Jerome Pesenti

@an_open_mind



Since [@ylecun](#) introduced his cake analogy at NeurIPS 5 years ago, unsupervised learning and the self-supervised approach in particular, has been a major focus of [@facebookai](#) and other AI labs. Progress has been steady and remarkable almost everywhere - NLU, MT, Speech, CV, Video

■ "Pure" Reinforcement Learning (cherry)

- ▶ The machine predicts a scalar reward given once in a while.

- ▶ **A few bits for some samples**

■ Supervised Learning (icing)

- ▶ The machine predicts a category or a few numbers for each input
- ▶ Predicting human-supplied data

- ▶ **10→10,000 bits per sample**

■ Unsupervised/Predictive Learning (cake)

- ▶ The machine predicts any part of its input for any observed part.
- ▶ Predicts future frames in videos

- ▶ **Millions of bits per sample**



■ (Yes, I know, this picture is slightly offensive to RL folks. But I'll make it up)

Not only does this approach often lead to better performance when combined with small amount of supervision, but it offers the promise of a more inclusive and private AI, one that works for all languages and cultures and doesn't require extensive human labeling. #ResponsibleAI

Here are pointers to some of our latest work relying on self-supervised learning:

Speech <https://t.co/4K03h47klI>

CV <https://t.co/FcS4QmYQMJ>

NLU <https://t.co/wCl2q93ybc>

MT <https://t.co/GllyZ8qmZ5>

Videos <https://t.co/o4VWF0r0vK>

Overall approach <https://t.co/Yc7JoJs95j>