

How should machines learn from humans?

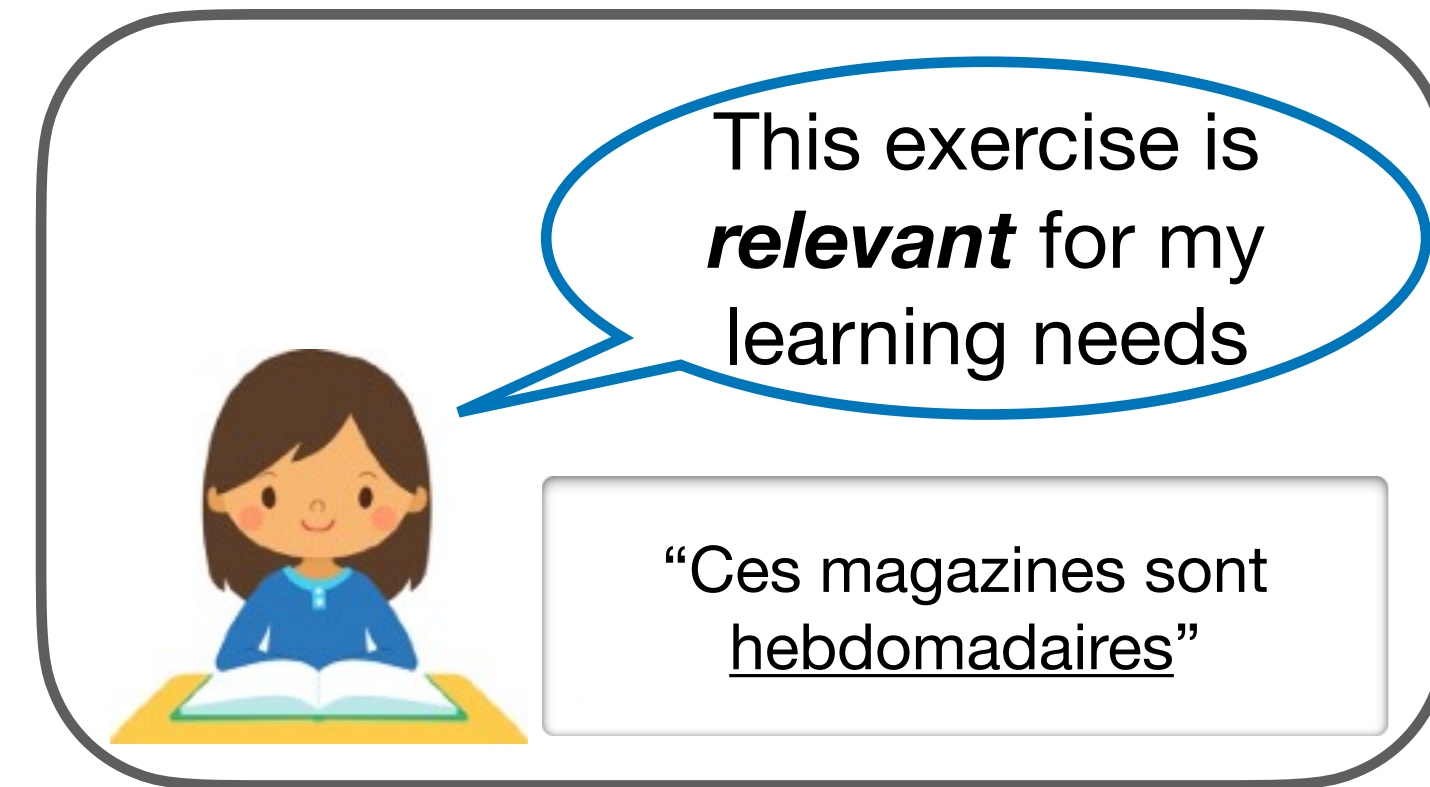
- Volume of responses needed is **massive!** → time, money, infrastructure

*How can we more effectively learn **with** human feedback and when can we use foundational tools to learn **without** additional feedback*

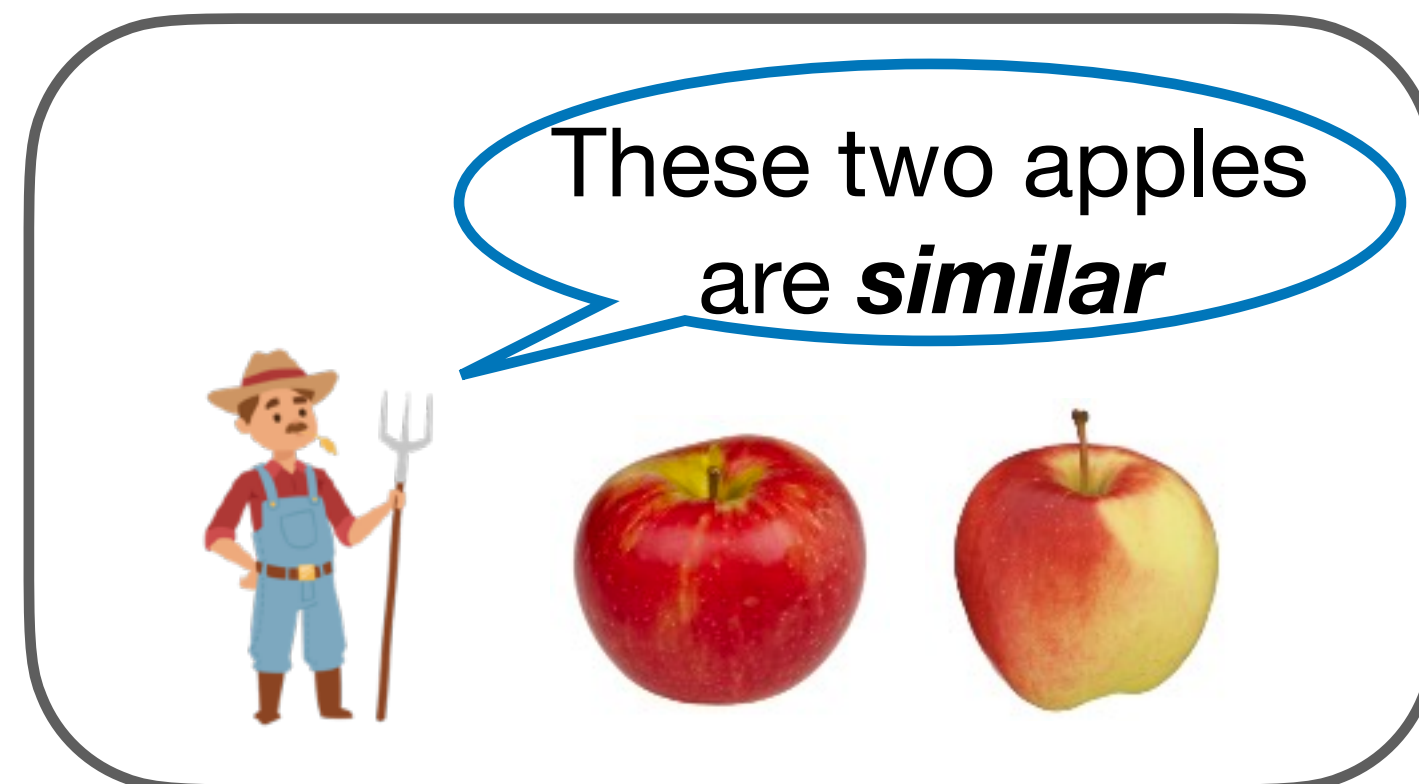
Learning from human feedback is crucial for modern ML systems!



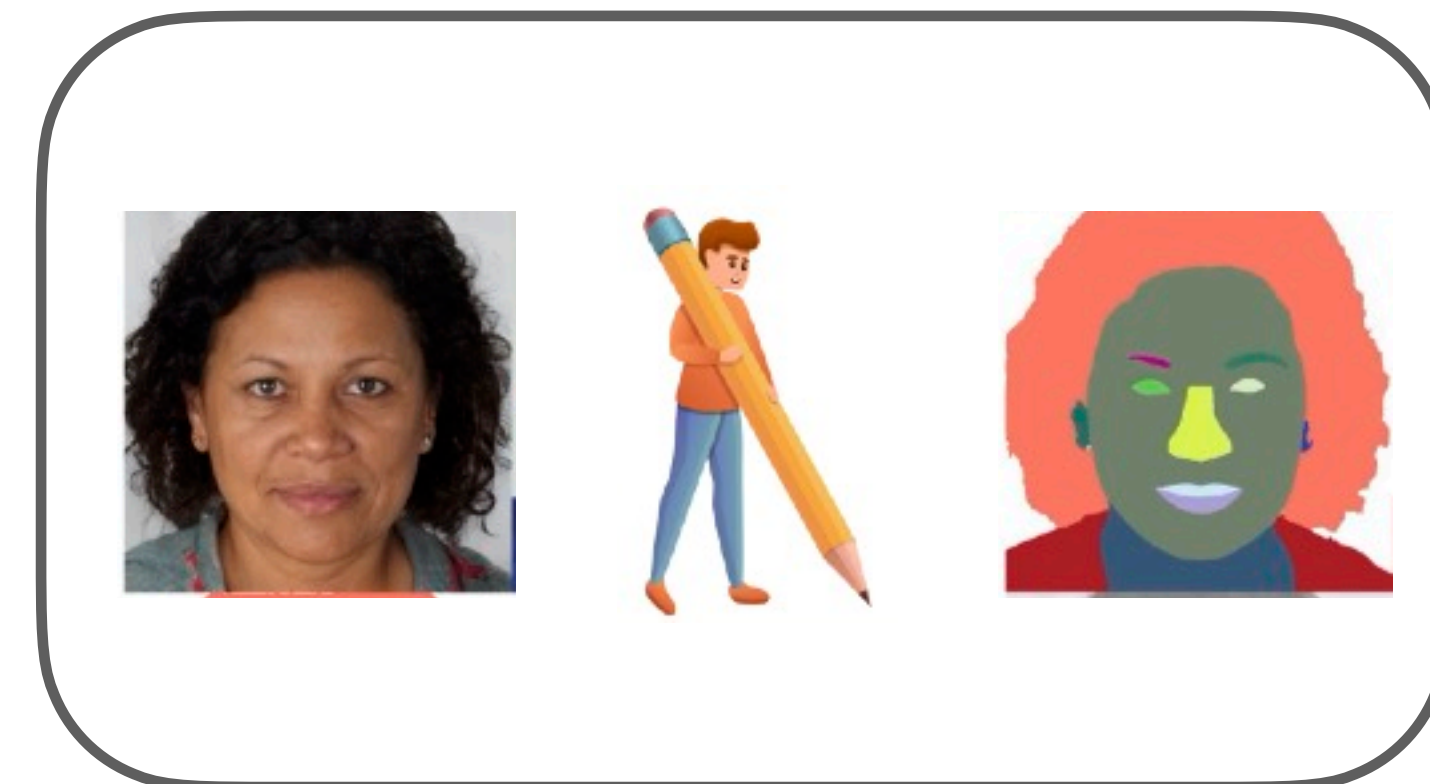
Preference



Relevance



Perception



Labels