

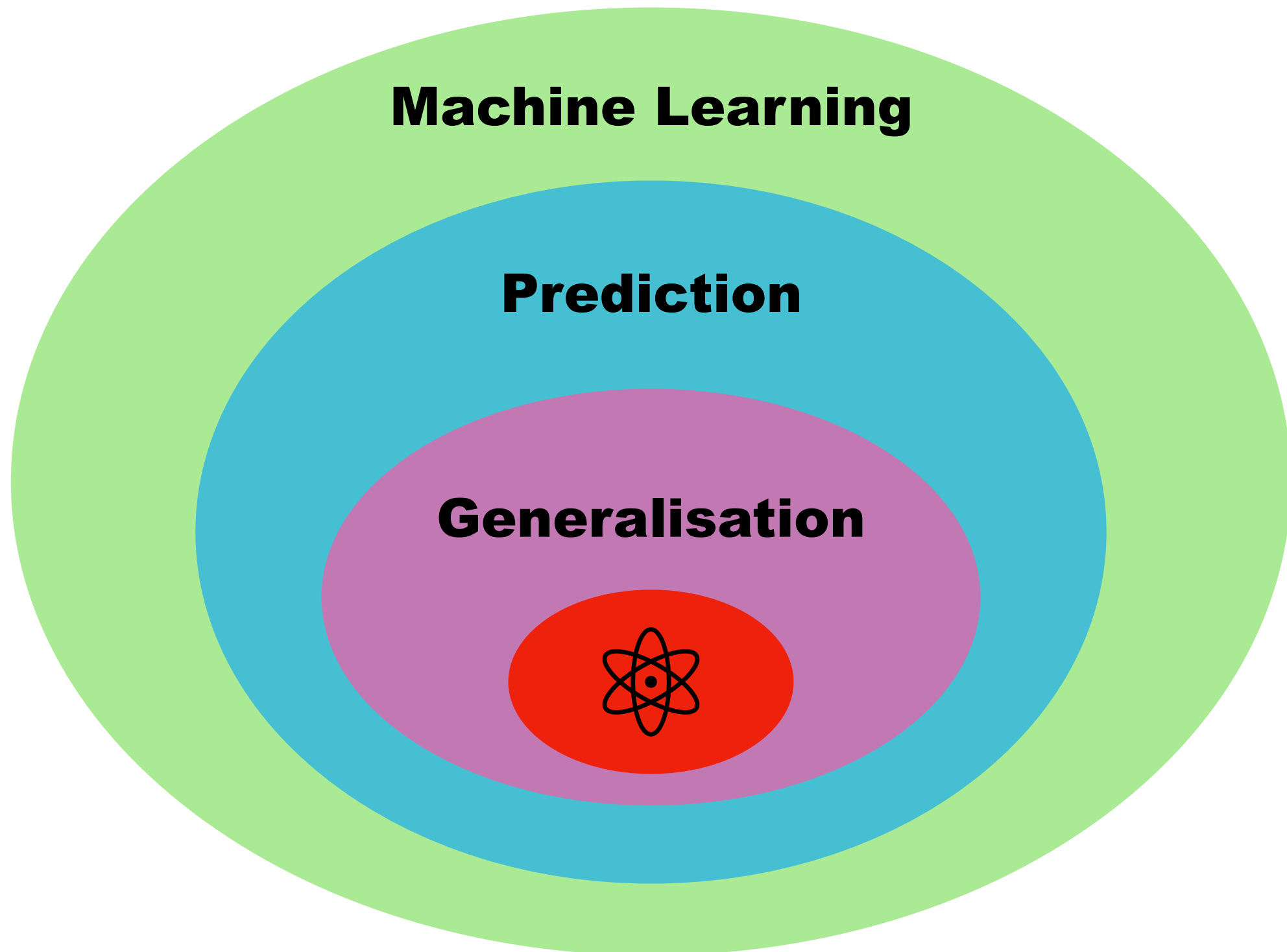
The Agnostic Hypothesis

Machine Learning's Achilles Heel or Midas Touch?

Chaochao Lu

CBL Tea Talk | 4th May 2020

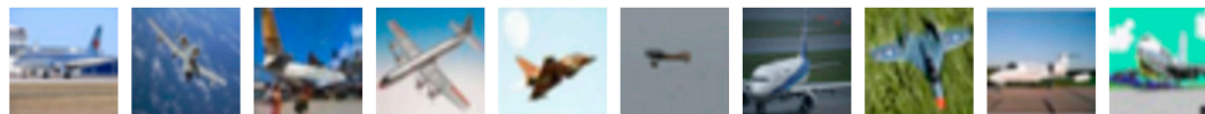
Core Question



Motivating Example

Is image classification a causal problem?

airplane



automobile



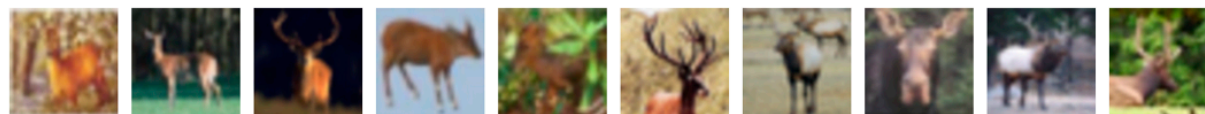
bird



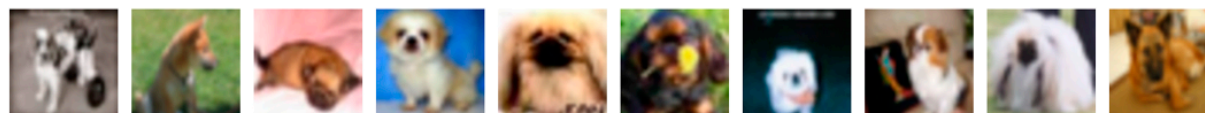
cat



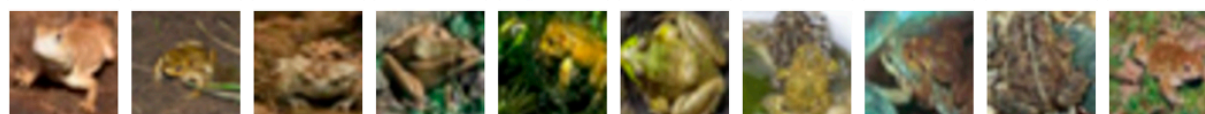
deer



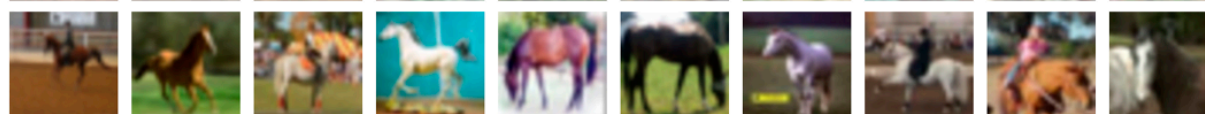
dog



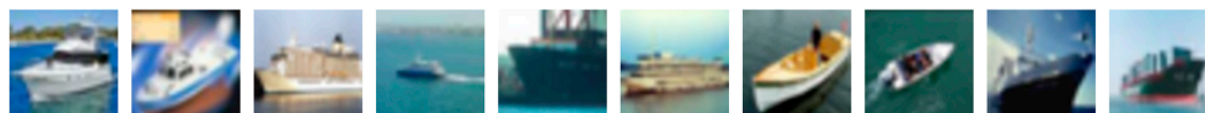
frog



horse



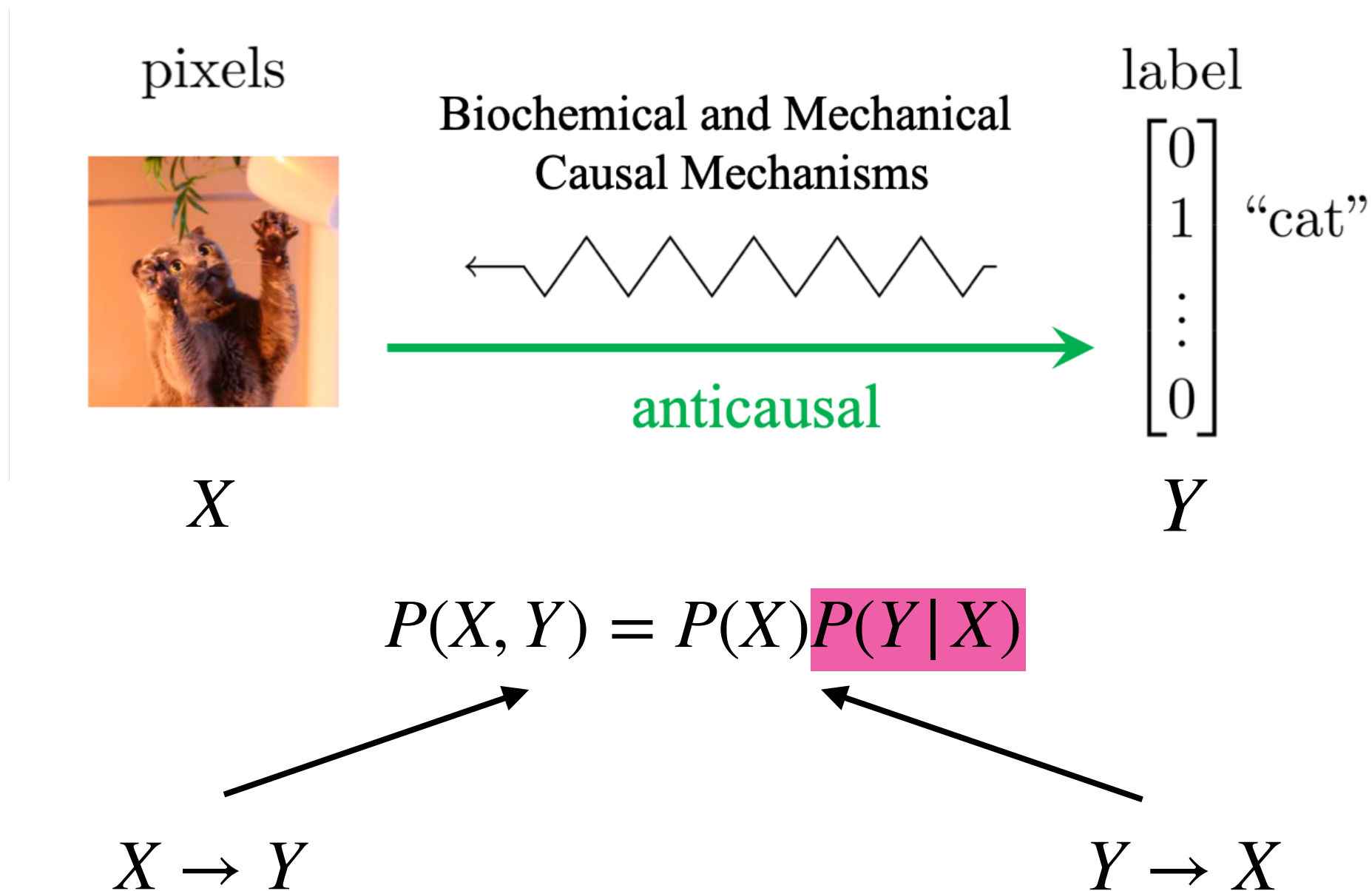
ship



truck

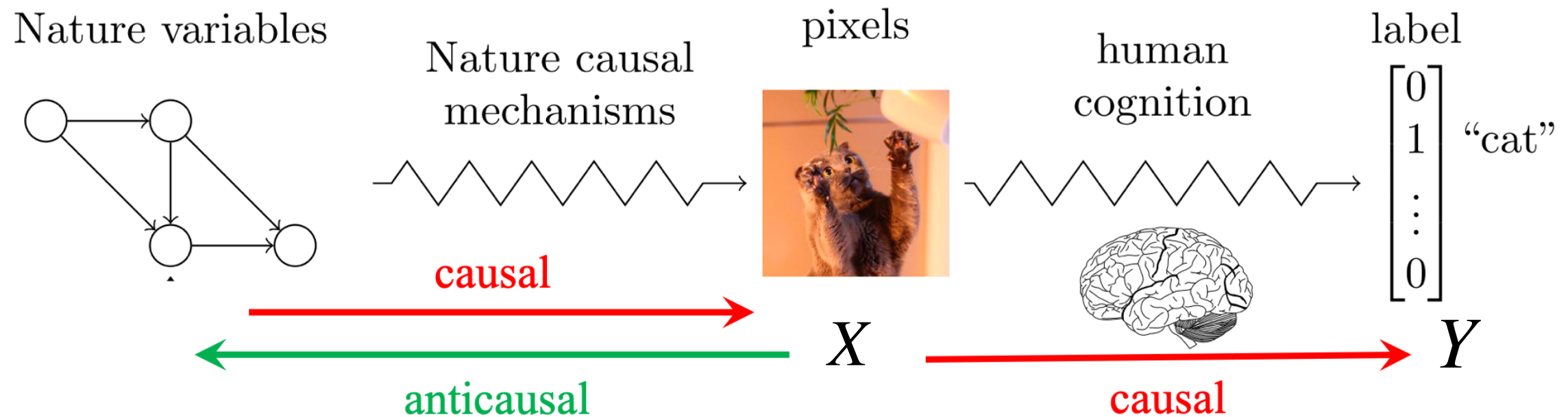


Opinion 1: Anticausal



Schölkopf et al. On Causal and Anticausal Learning, 2012

Opinion 2: Causal



- In the causal direction, Nature variables (e.g., colour, light, angle, animal, etc.) produce images through nature causal mechanisms.
- In the anticausal direction, we attempt to disentangle the underlying causal factors of variation behind images (i.e., Nature variables).
- Disentanglement vs. Inference
- Hierarchy of Nature Variables vs. Occam's Razor

- predict human annotations from images in order to imitate the cognitive process (i.e., humans produce labels by following a causal and cognitive process after observing images.)
- $P(Y|X)$ should be stable across environments or domains. Hence, empirical risk minimisation (ERM) should work quite well.

The Agnostic Hypothesis

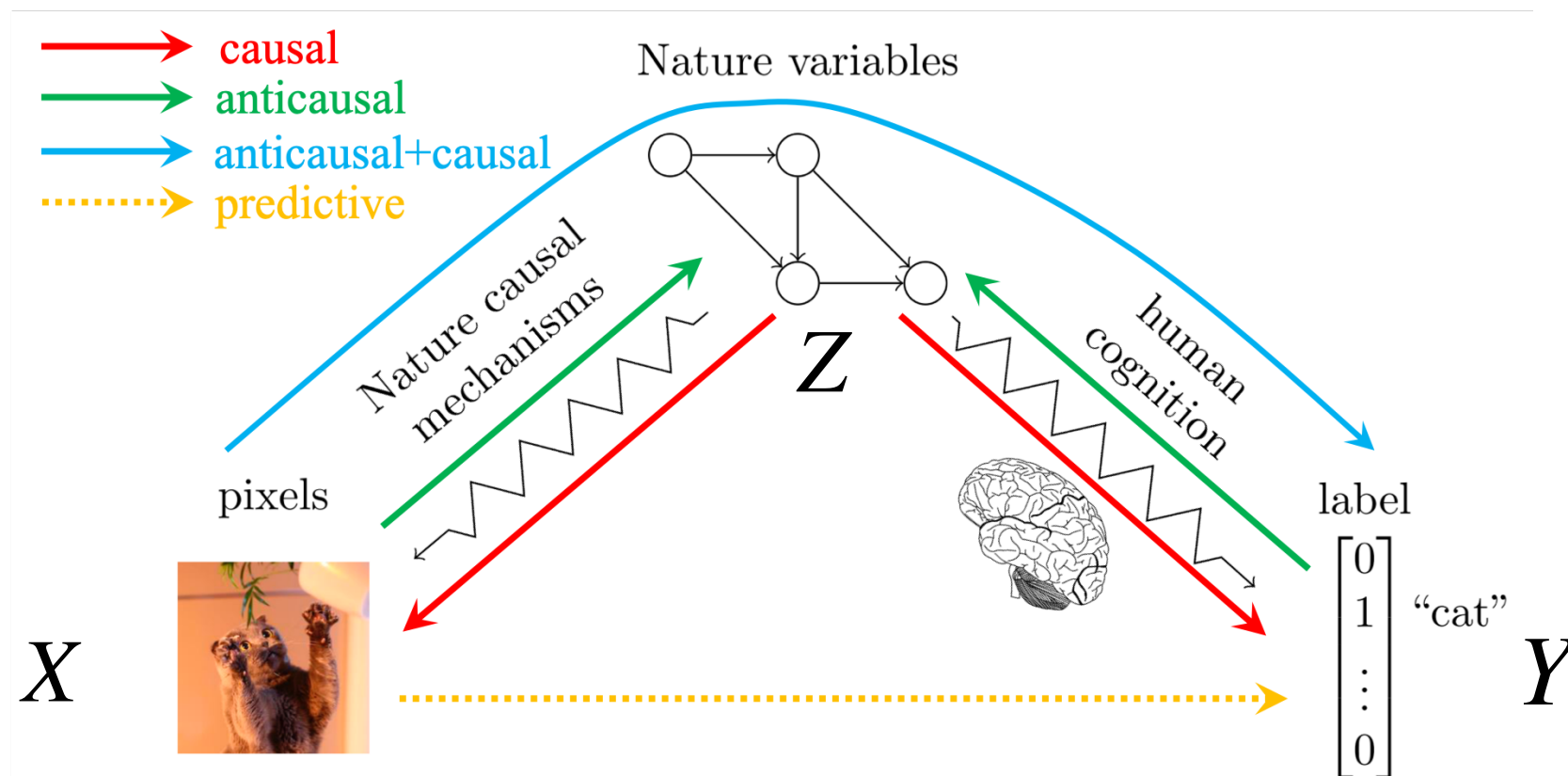


image and label as two different representation spaces

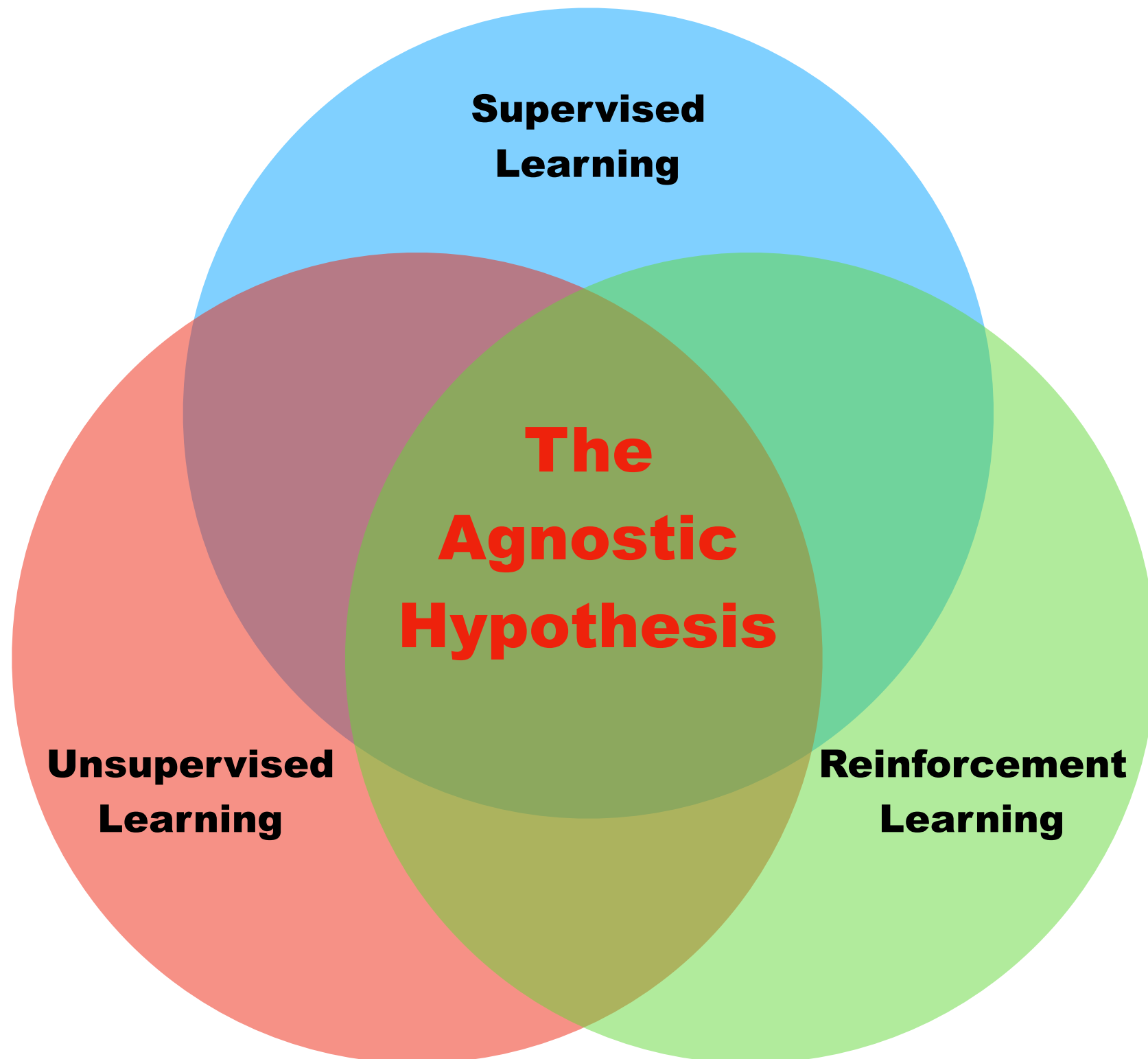
$$X \leftarrow Z \rightarrow Y$$

a feature extractor and a classifier

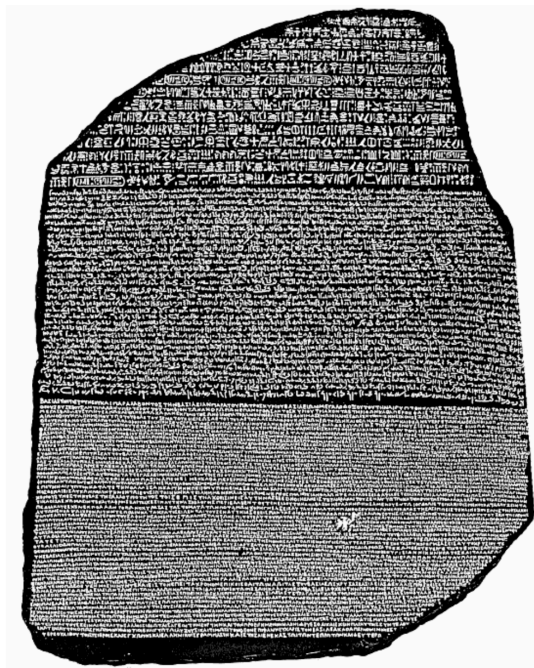


- **Plato's Theory of Forms** (Credit to Hannes)
- **Manipulability Theory**
- **Principle of Common Cause**
- **Theory of Linguistics** (Credit to Rebecca)

Connections to ML



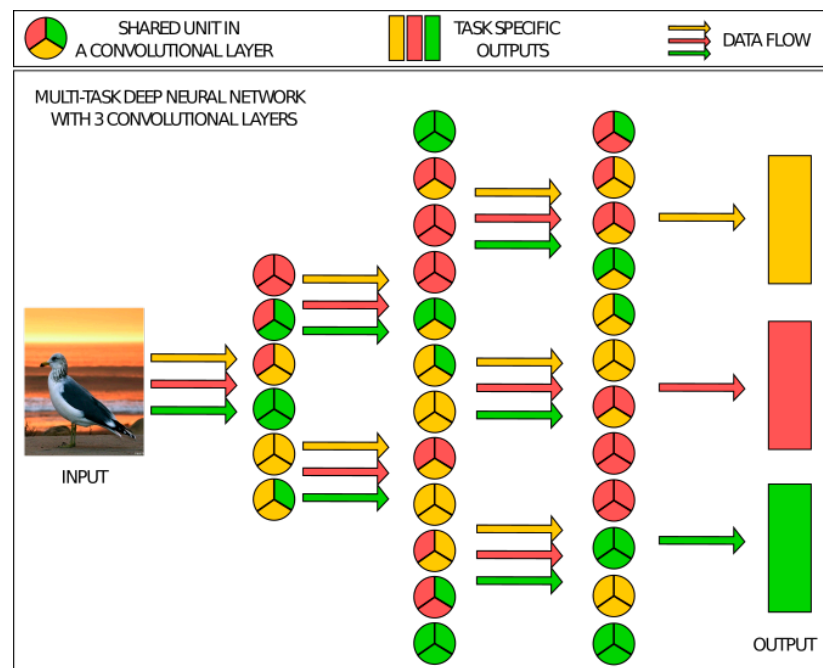
Concrete Examples



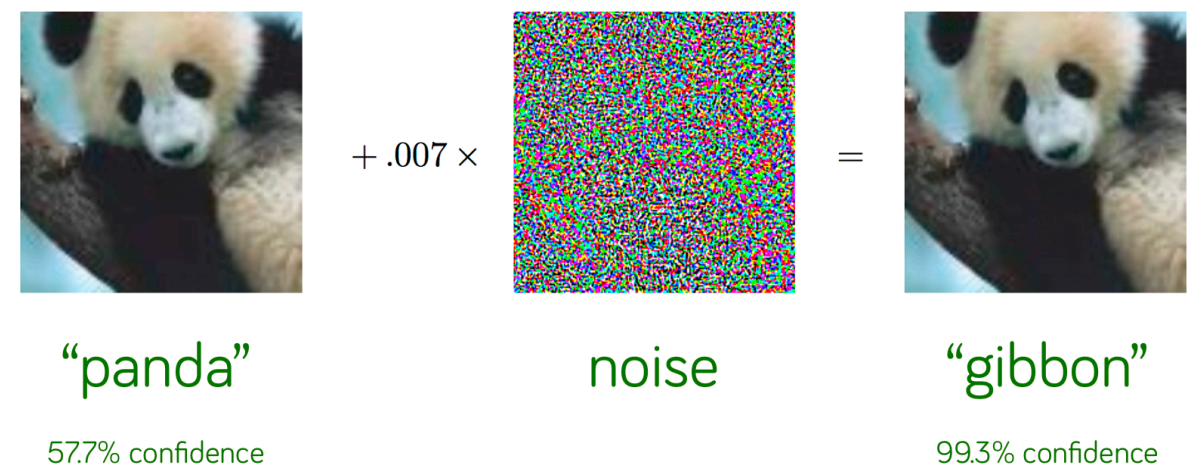
Multiview Nonlinear ICA



Time Series Prediction



Multi-Task Learning



Adversarial Attack

Thank you.

Any discussion is welcome.

<https://causallu.com/2020/04/05/is-image-classification-a-causal-problem/>