ICLR
International Conference On Learning Representations

UNIVERSITY OF OXFORD

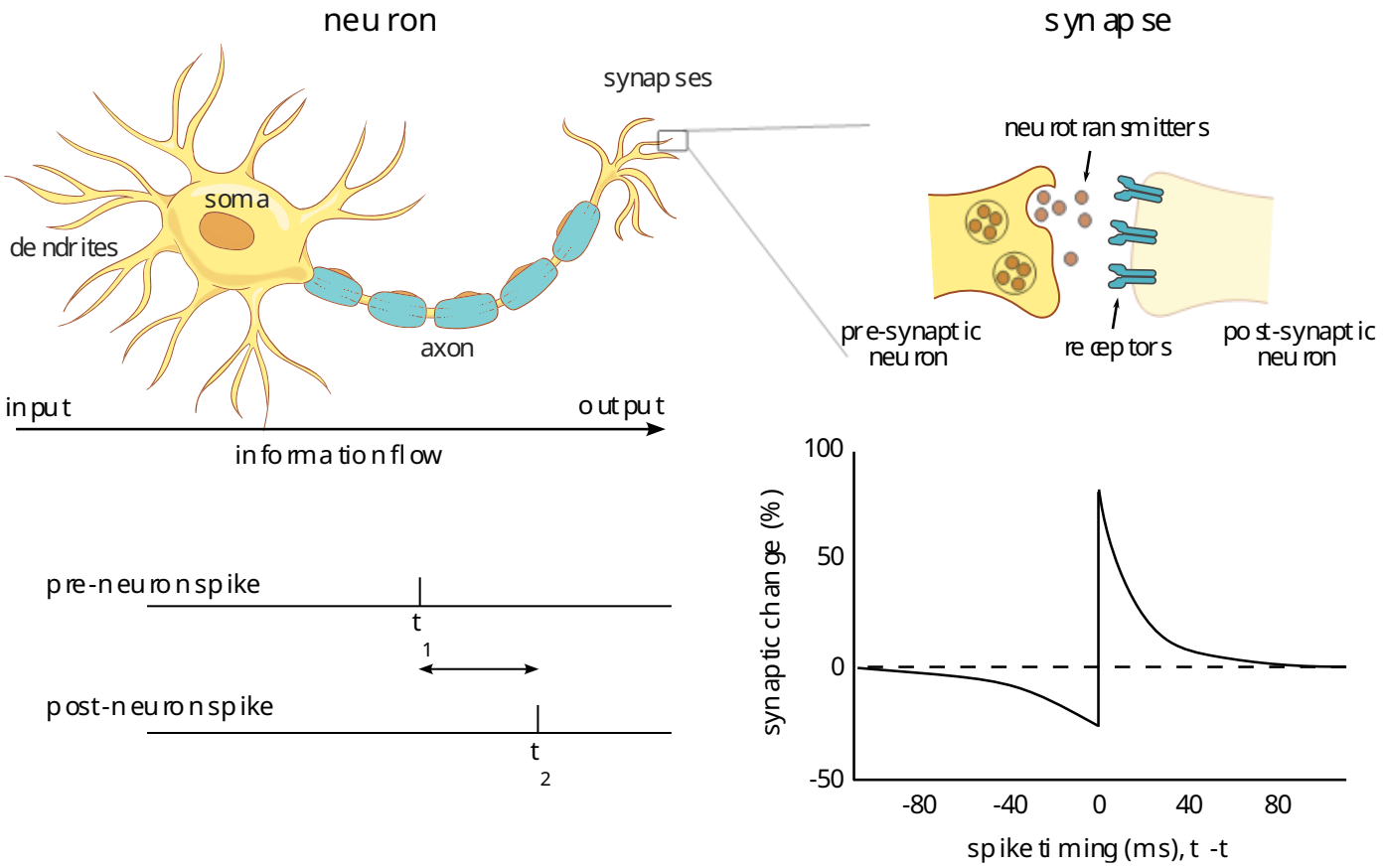
UK Research and Innovation

Bidirectional predictive coding

Gaspard Oliviers, Mufeng Tang, & Rafal Bogacz
Nuffield Department of Clinical Neurosciences, University of Oxford, UK
MRC Centre of Research Excellence in Restorative Neural Dynamics, UK
Correspondence to: rafal.bogacz@ndcn.ox.ac.uk

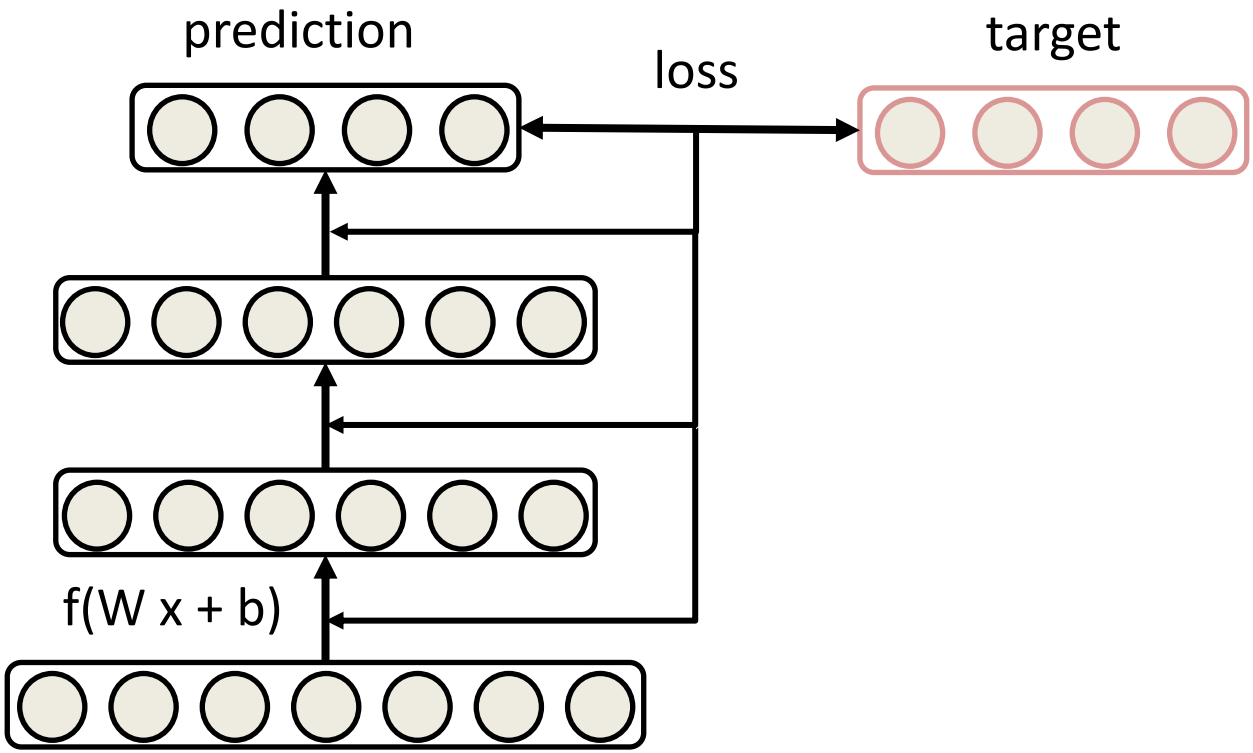
Learning in the brain

Synaptic weight depends only on their pre- and post-synaptic neurons

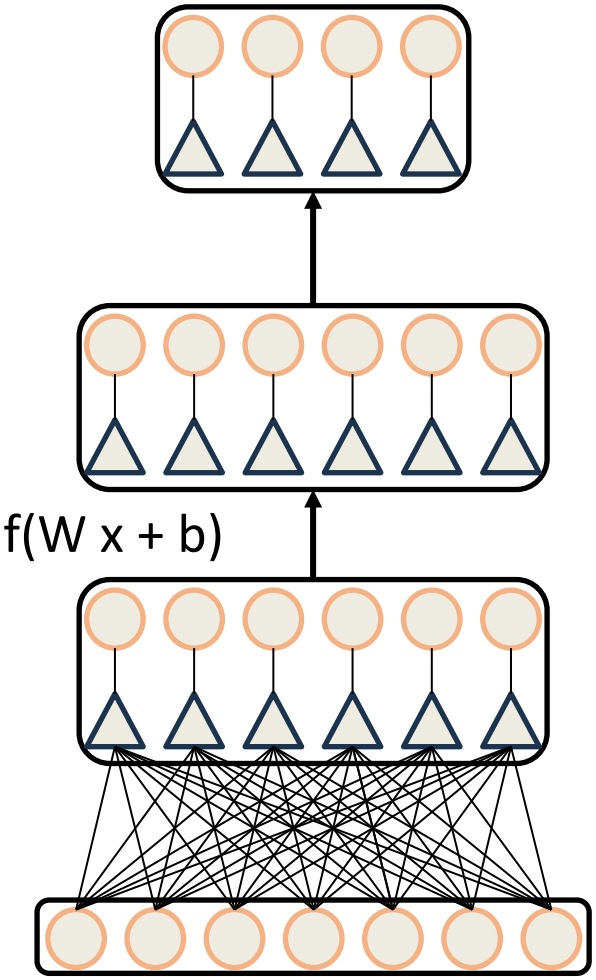


Neuron illustration from NIH BioArt

Backpropagation is not local



Predictive coding for local plasticity



state of neuron, x_l

synaptic input of neuron, $f(Wx_{l-1} + b)$

local loss = $(x_l - f(Wx_{l-1} + b))^2$

Predictive coding use cases

Explain the activity patterns of biological neurons
(Rao, Ballard. Nature. 1999)

Approximating backprop in the brain
(Whittington, Bogacz. Neural Comp. 2017)

Unsupervised learning
(Salvatori et al. NeurIPS 2022)

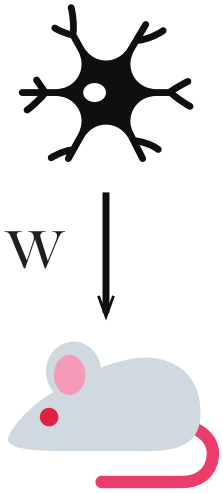
Classification accuracy similar to backprop
(Pinchetti et al. ICLR. 2025)

Better learning than backprop on biologically relevant tasks
(Song et al. Nature. 2024)

Problem statement - PC can be

Generative

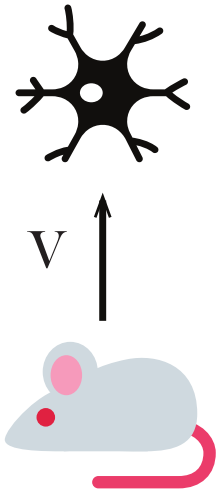
- Bayesian inference
- accurate unsupervised learning



$$E_g = \sum \frac{1}{2} ||x^l - W^{l+1} f(x^{l+1})||_2^2$$

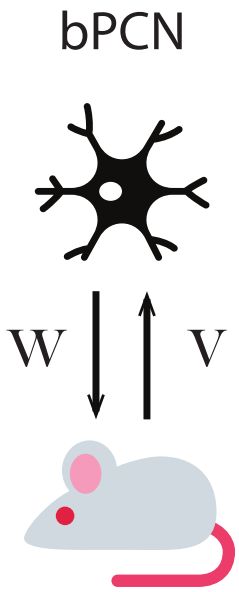
Discriminative

- fast inference
- accurate supervised learning



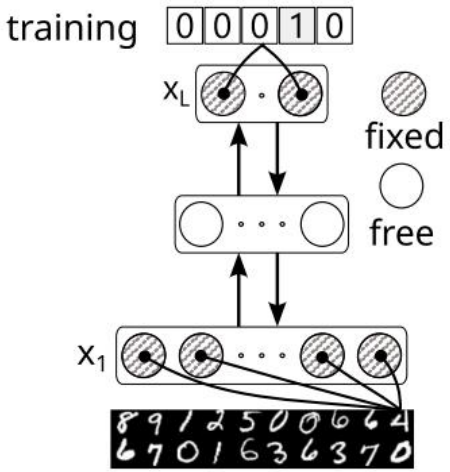
$$E_d = \sum \frac{1}{2} ||x^{l+1} - V^l f(x^l)||_2^2$$

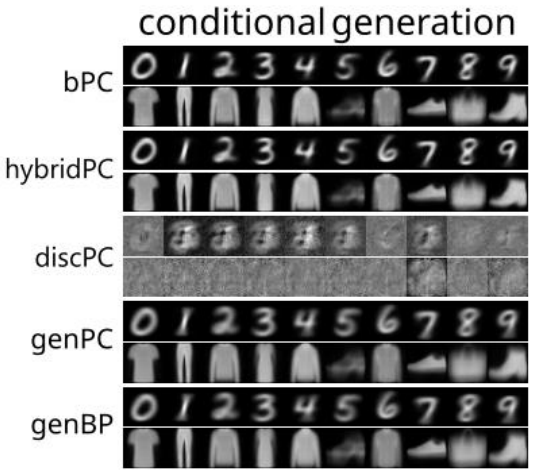
Bidirectional predictive coding

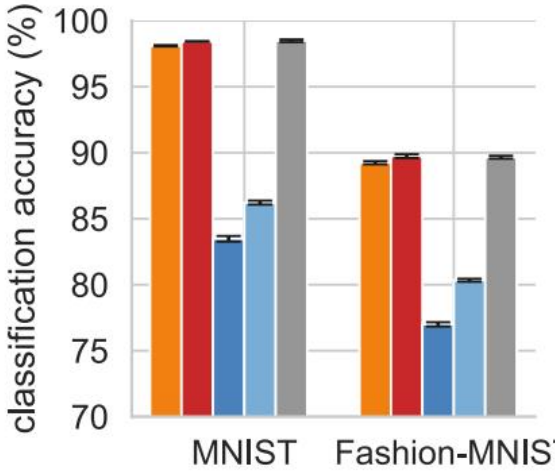


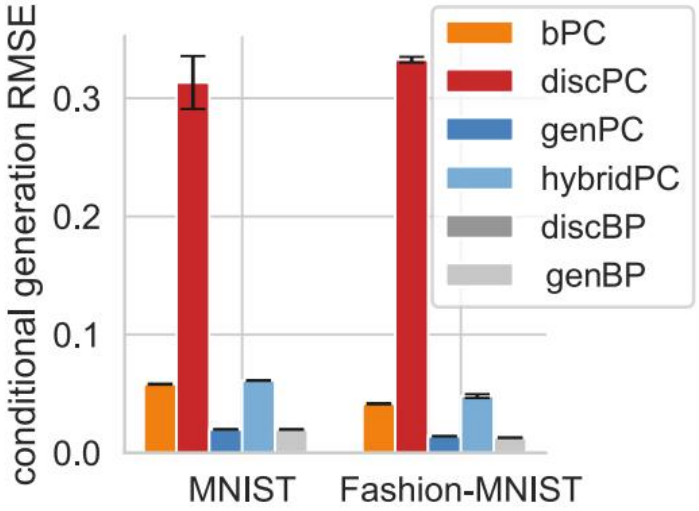
$$E = \frac{1}{\sigma_g}\sum\frac{1}{2}||x^l - W^{l+1}f(x^{l+1})||_2^2 + \frac{1}{\sigma_d}\sum\frac{1}{2}||x^{l+1} - V^l f(x^l)||_2^2$$

Combining classification and generation

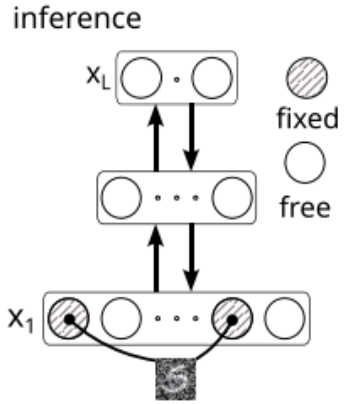


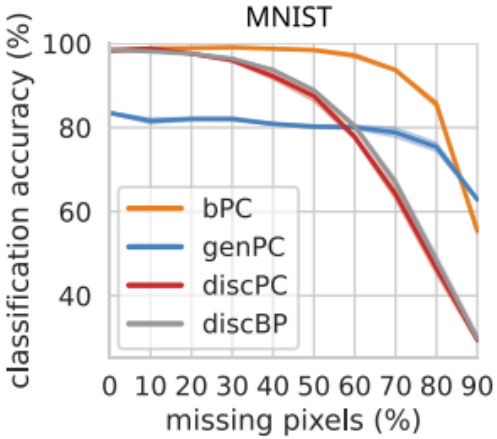


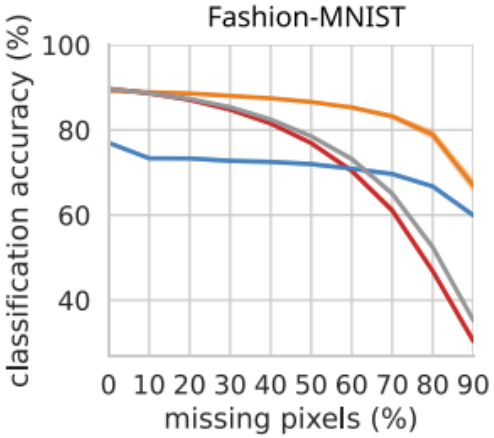


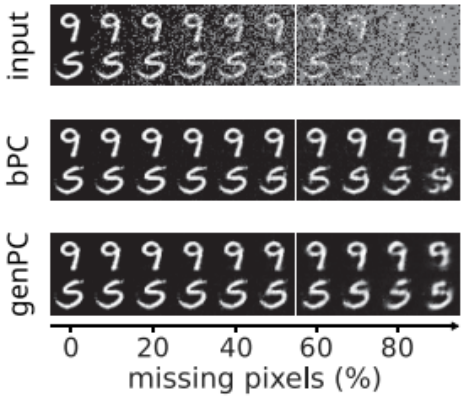


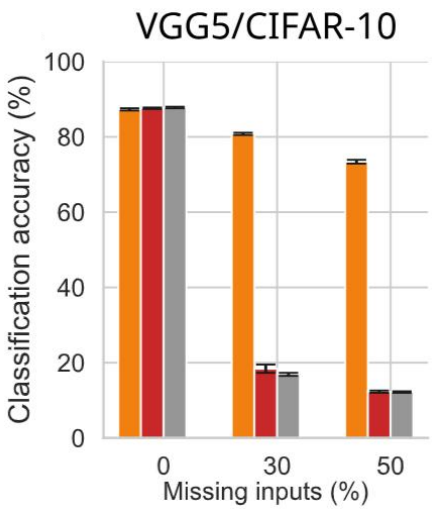
Benefits of combined discriminative and generative processing

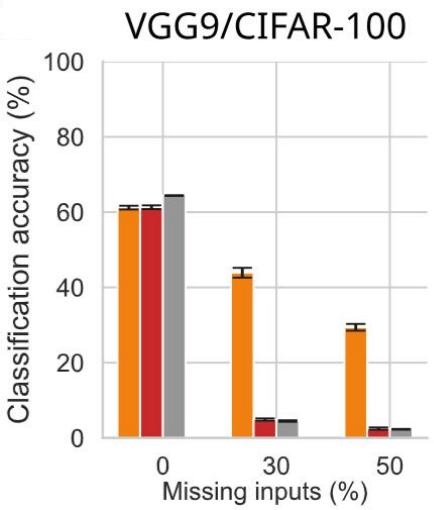


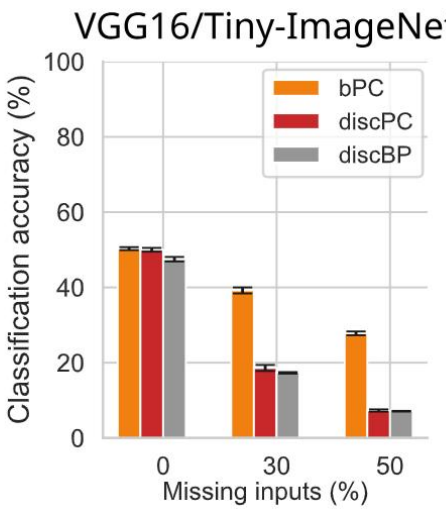


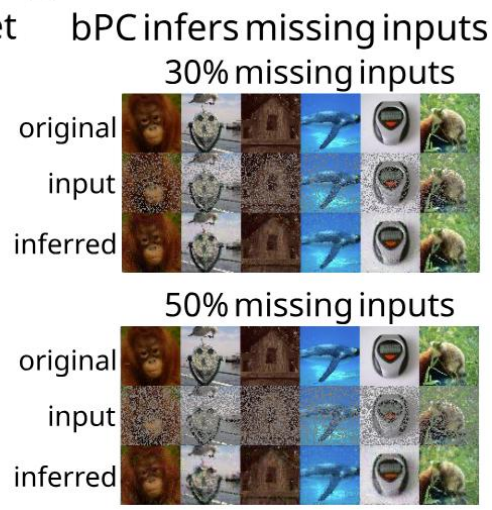






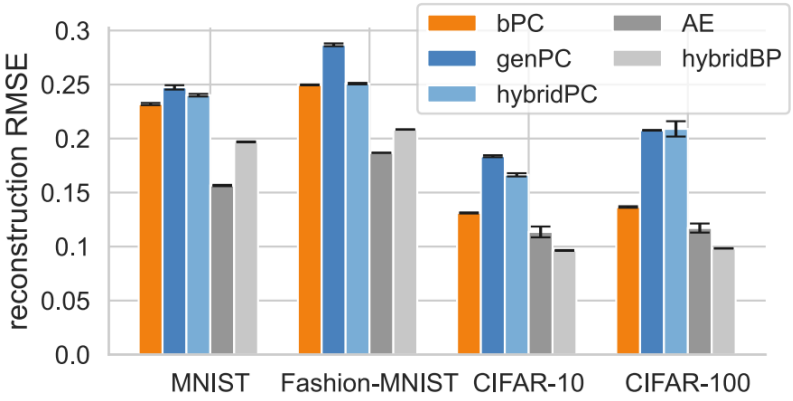




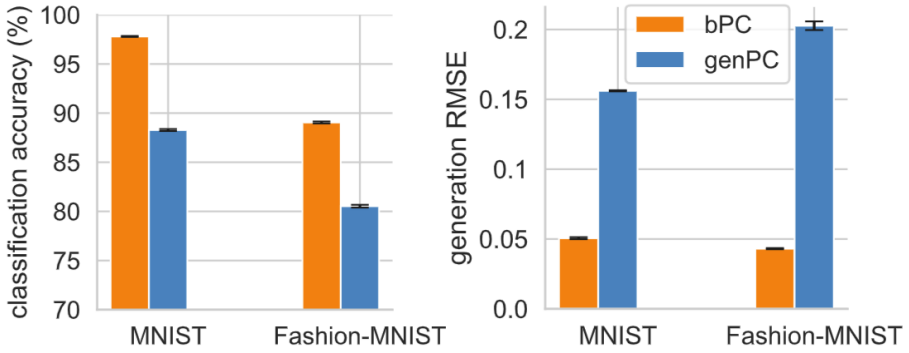


Additional benefits

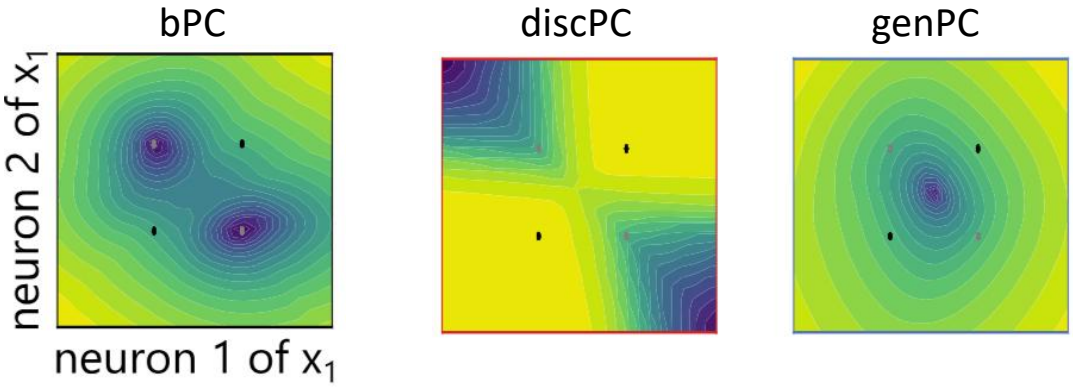
Unsupervised learning



Bimodal information transfer



Robust to out-of-distribution samples



For more results and details see:

Paper

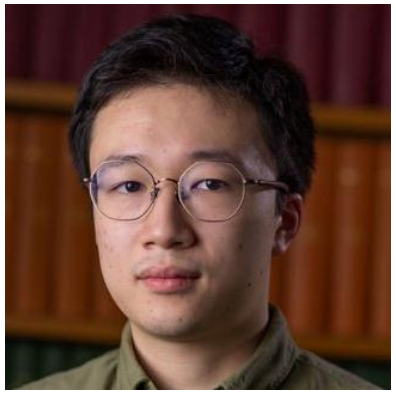
Codebase

Predictive coding tutorials

Research team

Gaspard Oliviers

Mufeng Tang



Rafal Bogacz