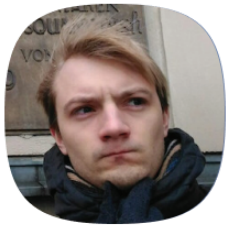
Maxim Borisyak



Learning to pivot

with Adversarial Networks

2021

MLHEP 2021
HIGHER SCHOOL OF ECONOMICS
NATIONAL RESEARCH UNIVERSITY
LAMBDA • HSE
Yandex
SCHOOL OF DATA ANALYSIS
EPFL
Schaffhausen Institute of Technology

Learning to pivot

Original paper

Learning to Pivot with Adversarial Networks

Gilles Louppe
New York University
g.louppe@nyu.edu

Michael Kagan
SLAC National Accelerator Laboratory
makagan@slac.stanford.edu

Kyle Cranmer
New York University
kyle.cranmer@nyu.edu

Problem statement

Make output of a classifier/regressor $f : \mathcal{X} \mapsto \mathcal{Y}$ independent from a nuisance variable $Z \in \mathcal{Z}$.

- particle identification;
- independent from:
  - momentum;
  - total number of tracks;
  - pseudo-rapidity.

- trigger;
- independent from:
  - MC vs real data.

Independence (statistical)

Random variables $F = f(X)$ and Z are independent iff

$$P(Z \mid F) = P(Z).$$

Independence (informational)

$$(F \perp Z) \iff P(Z \mid F) = P(Z).$$

- conditional entropy:

$$\mathbb{H}(Z \mid F) \leq \mathbb{H}(Z);$$

- equality occurs only when $P(F \mid Y) = P(Z)$;
- i.e., when F and Z are independent:

$$\text{maximize } \mathbb{H}(Z \mid F) \iff \text{achive } F \perp Z.$$

Conditional Entropy

$$\mathbb{H}(Z \mid F) = - \mathop{\mathbb{E}}_{x,z \sim P(X,Z)} \log P\big(z \mid f(x)\big)$$

Recovering $P\big(Z \mid f(X)\big)$ is a regression/classification problem:

$$f(X) \mapsto Z$$

Adversary

$$\mathbb{H}(Z \mid F) = -\mathop{\mathbb{E}}_{x,z \sim P(X,Z)} \log P\big(z \mid f(x)\big) = \min_r \left[-\mathop{\mathbb{E}}_{x,z \sim P(X,Z)} \log r\big(z \mid f(x)\big) \right]$$

- adversary $r : \mathcal{Y} \mapsto \Pi(\mathcal{Z})$:
  - $\Pi(\mathcal{Z})$ — set of probability distributions;

Adversary

Discrete $\mathcal{Z}$:

- r — classifier, e.g.:

$$r(\gamma) = \mathrm{softmax}(r_1(\gamma), r_2(\gamma), \dots, r_n(\gamma))$$

Continuous $\mathcal{Z}$:

- r — general regressor:
- MSE = approximation via a conditional Gaussian;
- approximation via conditional Gaussian Mixtures e.g.:

$$r(\gamma) = \alpha_1(\gamma)\Phi\big(\mu_1(\gamma), \sigma_2^2(\gamma)\big) + \ldots + \alpha_n(\gamma)\Phi\big(\mu_n(\gamma), \sigma_n^2(\gamma)\big)$$

Loss function

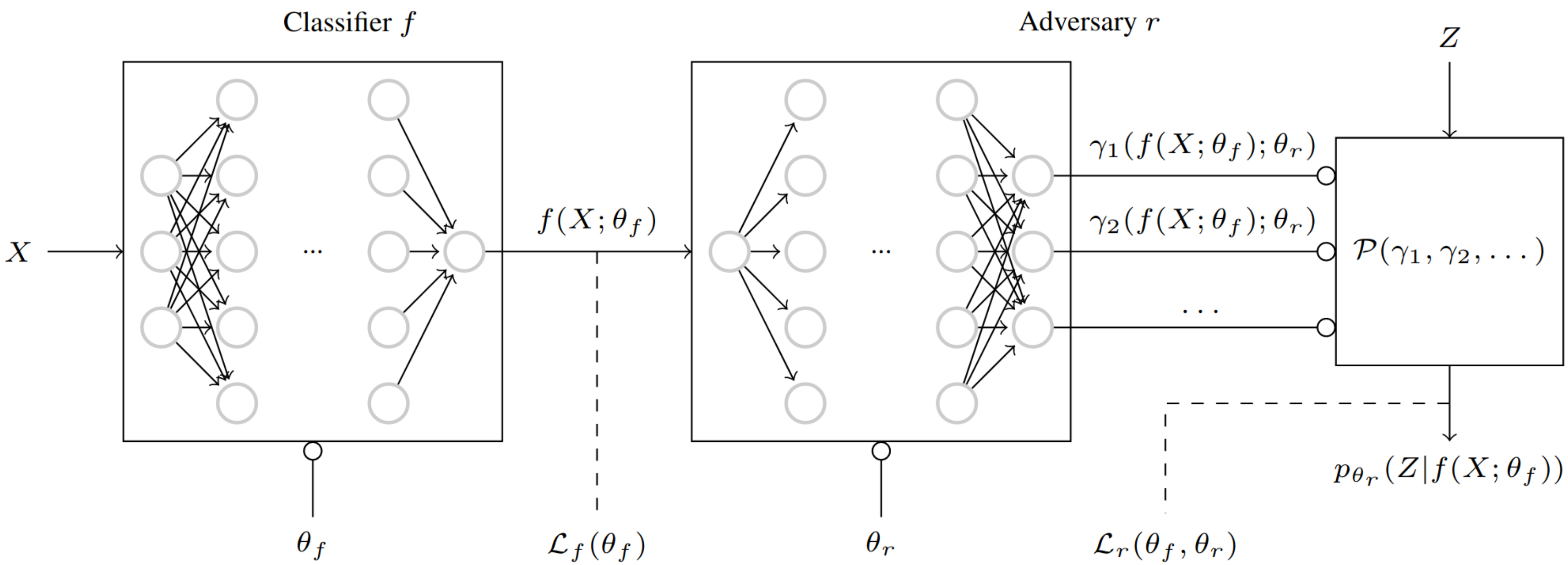
$$\mathcal{L}(f) = \mathcal{L}_f(f) - \mathcal{L}_r(f) \underset{f}{\to} \min;$$

$$\mathcal{L}_f(f) = - \mathop{\mathbb{E}}_{x,y \sim P(X,Y)} \log P\big(y \mid f(x)\big);$$

$$\mathcal{L}_r(f) = \mathbb{H}(Z \mid Y) = \min_r \left[- \mathop{\mathbb{E}}_{x,z \sim P(X,Z)} \log r\big(z \mid f(x)\big)\right];$$

- $\mathcal{L}_f$ — original loss;
- $\mathcal{L}_r$ — pivoting loss.

Pivoting


Classifier f
Adversary r
Z
X
...
$f(X;\theta_f)$
...
$\gamma_1(f(X;\theta_f);\theta_r)$
$\gamma_2(f(X;\theta_f);\theta_r)$
$\mathcal{P}(\gamma_1, \gamma_2, \dots)$
$\dots$
$p_{\theta_r}(Z|f(X;\theta_f))$
θ_f
$\mathcal{L}_f(\theta_f)$
θ_r
$\mathcal{L}_r(\theta_f, \theta_r)$

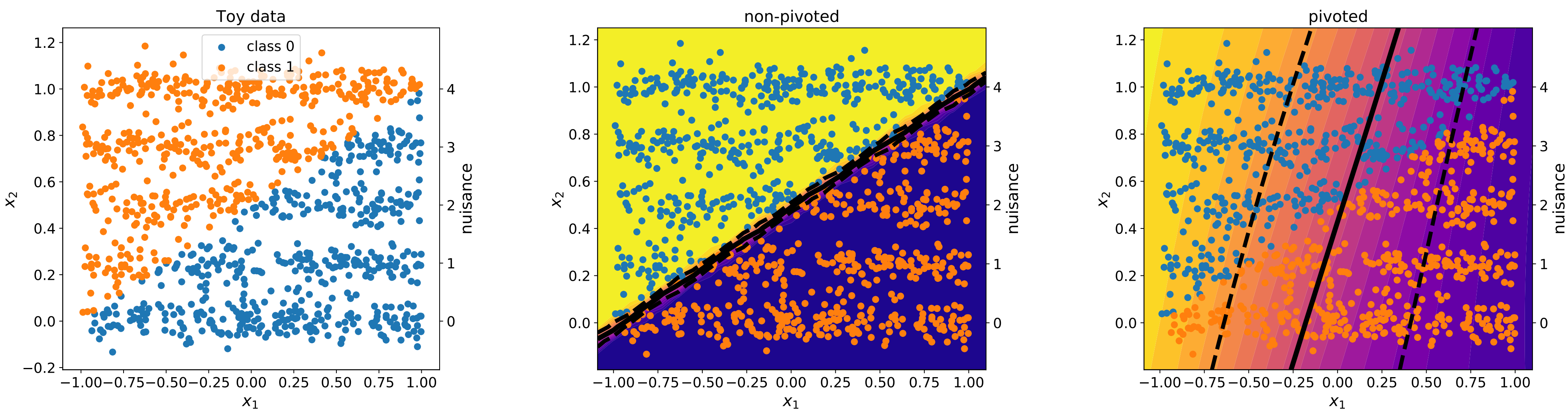
Adversarial Learning

while not converged do
    for $i = 1 \ldots N$ do
        sample x, z
        $G_\psi \leftarrow -\nabla_\psi \log r_\psi\big(z \mid f_\theta(x)\big)$
        $\psi \leftarrow \mathrm{adam}(G_\psi)$
    end for
    sample x, y
    $G^f_\theta \leftarrow -\nabla_\theta \log P\big(y \mid f_\theta(x)\big)$
    $G^r_\theta \leftarrow -\nabla_\theta \log r_\psi\big(z \mid f_\theta(x)\big)$
    $\theta \leftarrow \mathrm{adam}(G^f_\theta - G^r_\theta)$
end **while**

Example

$$x_1 \sim U[0, 1]; \qquad Z \sim U\{0, 1, 2, 3, 4\};$$

$$x_2 \sim N(0, \sigma^2) + Z; \qquad y = \mathbb{I}\left[x_2 > x_1\right].$$



Conditional pivoting

Target variable Y might depend on nuisance Z.

Trade-off:

$$\mathcal{L}(f) = \mathcal{L}_f(f) - \lambda\, \mathcal{L}_r(f);$$

Condition on Y:

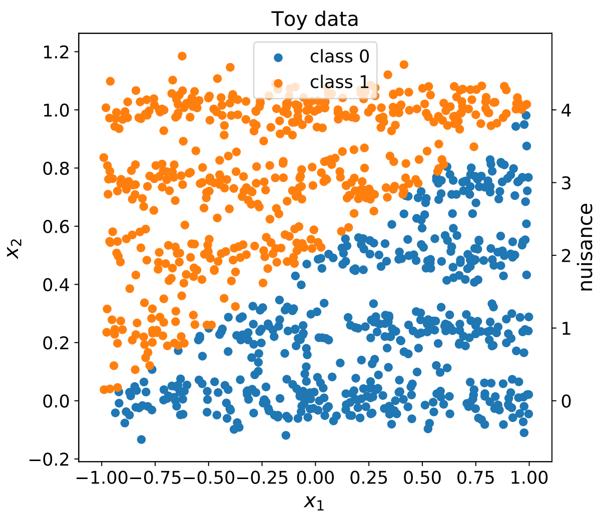
$$r\big(f(x)\big) \to r\big(f(x), y\big);$$

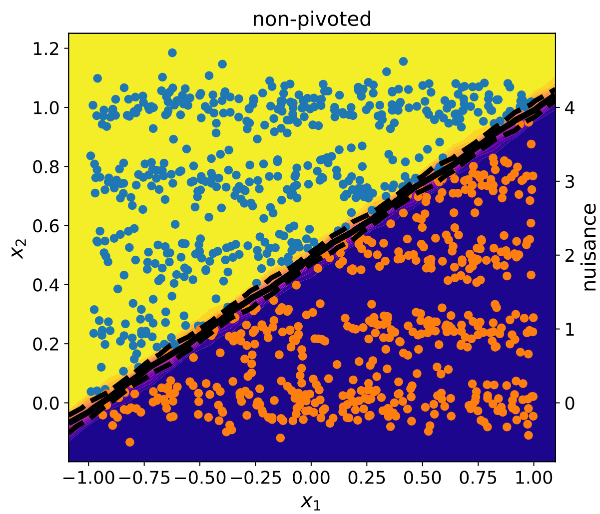
- pivots each class independently;
- combined predictions are **not** independent from Z.

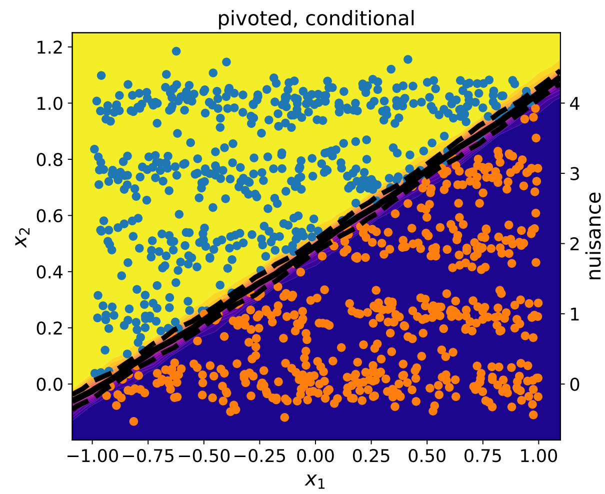
Example: conditional pivoting

$$x_1 \sim U[0, 1]; \qquad Z \sim U\{0, 1, 2, 3, 4\};$$

$$x_2 \sim N(0, \sigma^2) + Z; \qquad y = \mathbb{I}\left[x_2 > x_1\right].$$

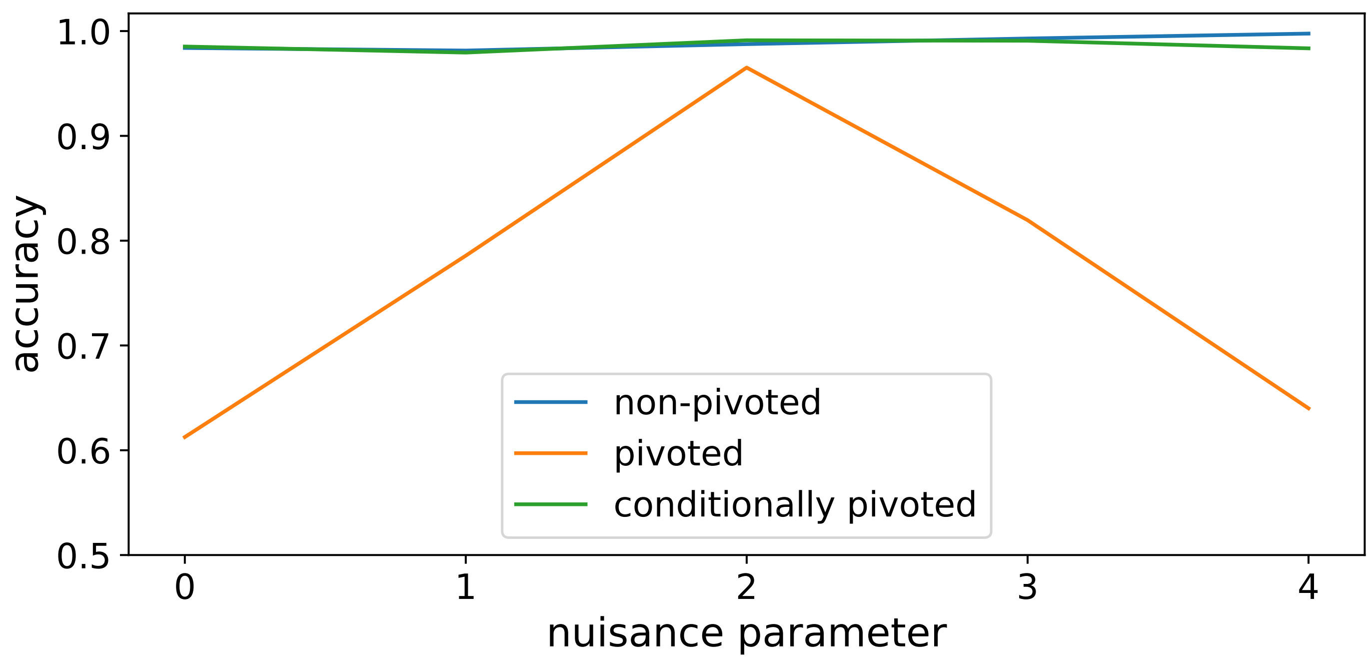




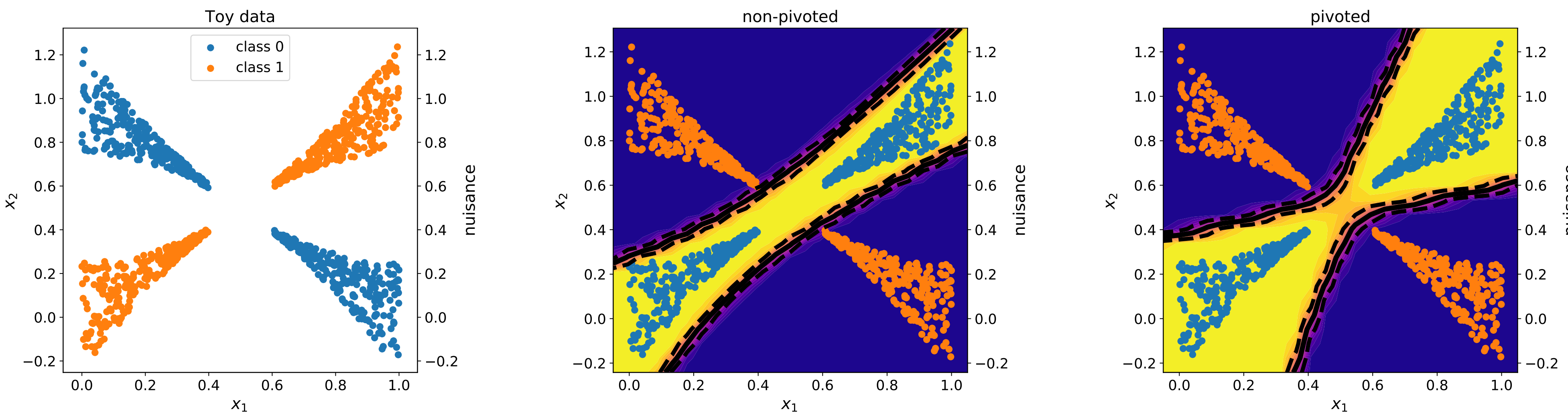


Performance

- unconditional pivoting might degrade performance;
- conditional does not.



Independence $\neq$ invariance



Summary

Summary

Adversarial learning:

- allows to measure and reduce dependencies;

Pivoting:

- makes predictions independent from nuisance;
- conditional pivoting acts within each class independently;
- conditional pivoting does not degrade performance.

Independence $\neq$ invariance.

References

- Louppe G, Kagan M, Cranmer K. Learning to pivot with adversarial networks. In Advances in neural information processing systems 2017 (pp. 981-990).
- Ganin Y, Lempitsky V. Unsupervised domain adaptation by backpropagation. In International conference on machine learning 2015 Jun 1 (pp. 1180-1189).
- Derkach D, Hushchyn M, Kazeev N. Machine Learning based Global Particle Identification Algorithms at the LHCb Experiment. In EPJ Web of Conferences 2019 (Vol. 214, p. 06011). EDP Sciences.