PAC-Bayesian Generalization Bound on Constrained f-Entropic Risk Measures

Hind Atbir[1,◇] Farah Cherfaoui[1,◇] Guillaume Metzler[3] Emilie Morvant[1] **Paul Viallard**[4]

[1] Univ Jean Monnet St-Étienne, CNRS, IOGS,
Laboratoire Hubert Curien UMR 5516, [◇] Inria, St-Etienne,
[3] Univ de Lyon, Lyon 2, ERIC UR3083, Bron,
[4] Univ Rennes, Inria, CNRS IRISA - UMR 6074, Rennes

AISTATS 2026

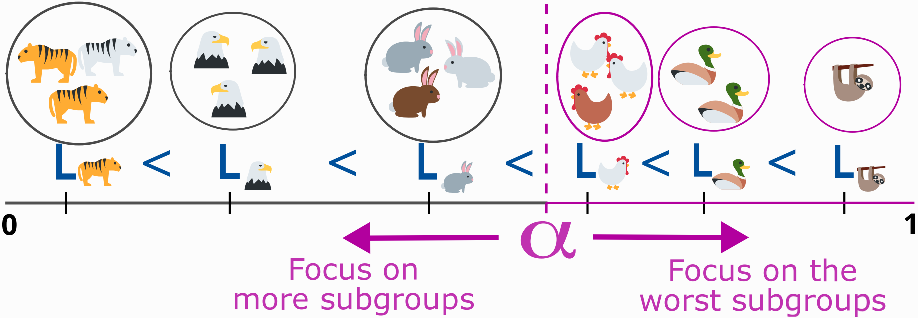
Tight and Minimizable PAC-Bayesian Generalization Bound on Constrained f-Entropic Risk Measures for Subgroup Learning

Hind Atbir[1,◇] Farah Cherfaoui[1,◇] Guillaume Metzler[3] Emilie Morvant[1] **Paul Viallard**[4]

[1] Univ Jean Monnet St-Étienne, CNRS, IOGS,
Laboratoire Hubert Curien UMR 5516, [◇] Inria, St-Etienne,
[3] Univ de Lyon, Lyon 2, ERIC UR3083, Bron,
[4] Univ Rennes, Inria, CNRS IRISA - UMR 6074, Rennes

AISTATS 2026

From the CVaR to Constrained f-Entropic Risk Measures



Conditional Value at Risk

With **reference distribution** π on the subgroups and $\alpha \in (0,1]$,

$$\mathbf{CVaR}_{\alpha}(h) = \sup_{\rho \in \mathrm{E}_{\alpha}} \mathop{\mathbb{E}}_{\mathrm{A}\sim\rho} L_{\mathrm{A}}(h) \quad \text{with } \mathrm{E}_{\alpha} = \left\{ \rho \;\middle|\; \rho \ll \pi \text{ and } \frac{d\rho}{d\pi} \leq \frac{1}{\alpha} \right\}$$

Risk of the subgroup A

From the CVaR to Constrained f-Entropic Risk Measures

Constrained f-Entropic Risk Measures

For **any convex function** f such that $D_f(\rho \| \pi)$ is a f-divergence,

$$\mathcal{R}(h) = \sup_{\rho \in E} \mathop{\mathbb{E}}_{\mathrm{A} \sim \rho} L_{\mathrm{A}}(h) \qquad \text{with } E = \left\{ \rho \;\middle|\; \rho \ll \pi \text{ and } \mathop{\mathbb{E}}_{\mathrm{A} \sim \pi} f\left(\frac{d\rho}{d\pi}(\mathrm{A})\right) \leq \beta \text{ and } \frac{d\rho}{d\pi} \leq \frac{1}{\alpha} \right\}$$

$$\Uparrow$$

Conditional Value at Risk

With **reference distribution** π on the subgroups and $\alpha \in (0, 1]$,

$$\mathbf{CVaR}_{\boldsymbol{\alpha}}(h) = \sup_{\rho \in \mathrm{E}_\alpha} \mathop{\mathbb{E}}_{\mathrm{A} \sim \rho} \underbrace{L_{\mathrm{A}}(h)}_{\text{Risk of the subgroup } \mathrm{A}} \quad \text{with } \mathrm{E}_\alpha = \left\{ \rho \;\middle|\; \rho \ll \pi \text{ and } \frac{d\rho}{d\pi} \leq \frac{1}{\alpha} \right\}$$

What can we guarantee?

Our (Disintegrated) PAC-Bayesian Generalization Bound

We have with probabilty $1-\delta$ over $\mathcal{S} \sim D^m$ and $h \sim Q_{\mathcal{S}}$,

$$\underbrace{\sup_{\rho \in E} \mathop{\mathbb{E}}_{\mathrm{A} \sim \rho} L_{\mathrm{A}}(h)}_{\substack{\text{True risks}\\ \text{of the worst subgroups}}} \leq \underbrace{\sup_{\rho \in E} \mathop{\mathbb{E}}_{\mathrm{A} \sim \rho} \hat{L}_{\mathrm{A}}(h)}_{\substack{\text{Empirical risks}\\ \text{of the worst subgroups}}} + \sqrt{\mathop{\mathbb{E}}_{\mathrm{A} \sim \pi} \frac{1}{2\,\alpha\, m_{\mathrm{A}}} \left[\underbrace{\ln^{+}\!\left(\frac{dQ_{\mathcal{S}}}{dP}(h)\right)}_{\text{Divergence btw } Q_{\mathcal{S}} \text{ and } P} + \ln \frac{2n\sqrt{m_{\mathrm{A}}}}{\delta} \right]}$$

m_{A} size of group A, n number of subgroups

What do we do in practice?

Self-bounding algorithm to minimize the bound

$\Downarrow$

Models with built-in **subgroup generalization guarantees**

PAC-Bayesian Generalization Bound on Constrained f-Entropic Risk Measures

Goal: Ensure good **generalization** for **every subgroup** A in the data

Our contribution

Tight and **Minimizable**
PAC-Bayesian Generalization Bound
on **Constrained f-Entropic Risk Measures**
for **Subgroup Learning**

Scan for paper

Poster 176

Hind Atbir Farah Cherfaoui Guillaume Metzler Emilie Morvant **Paul Viallard**