

Learning with Fenchel-Young Losses

- **Mathieu Blondel** (Google Brain Paris)
- André Martins (Unbabel & Instituto de Telecomunicações)
- Vlad Niculae (Instituto de Telecomunicações Lisbon, Portugal)

Keywords: loss functions, output regularization, convex duality, structured prediction.

Abstract: Over the past decades, numerous loss functions have been proposed for a variety of supervised learning tasks, including regression, classification, ranking, and more generally structured prediction. Understanding the core principles and theoretical properties underpinning these losses is key to choose the right loss for the right problem, as well as to create new losses which combine their strengths. In this paper, we introduce Fenchel-Young losses, a generic way to construct a convex loss function for a regularized prediction function. We provide an in-depth study of their properties in a very broad setting, covering all the aforementioned supervised learning tasks, and revealing new connections between sparsity, generalized entropies, and separation margins. We show that Fenchel-Young losses unify many well-known loss functions and allow to create useful new ones easily. Finally, we derive efficient predictive and training algorithms, making Fenchel-Young losses appealing both in theory and practice.

References:

- [1] Blondel, M., Martins, A. F., and Niculae, V. (2020). Learning with Fenchel-Young losses. *J. Mach. Learn. Res.*, 21(35), 1-69.