

Can Machines Learn Weak Signals?

Dacheng Xiu

**Booth School of Business, University of Chicago
(with Zhouyu Shen)**

In high-dimensional regression scenarios with low signal-to-noise ratios, we assess the predictive performance of several machine learning algorithms. Theoretical insights show Ridge regression's superiority in exploiting weak signals, surpassing a zero benchmark. In contrast, Lasso fails to exceed this baseline, indicating its learning limitations. Simulations reveal that Random Forest generally outperforms Gradient Boosted Regression Trees when signals are weak. Moreover, Neural Networks with ℓ_2 -regularization excel in capturing nonlinear functions of weak signals. Our empirical analysis across six economic datasets suggests that the weakness of signals, not necessarily the absence of sparsity, may be Lasso's major limitation in economic predictions.