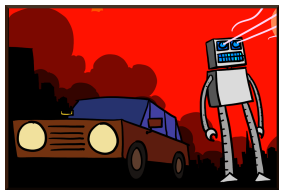
Multilabel Prediction via Compressed Sensing

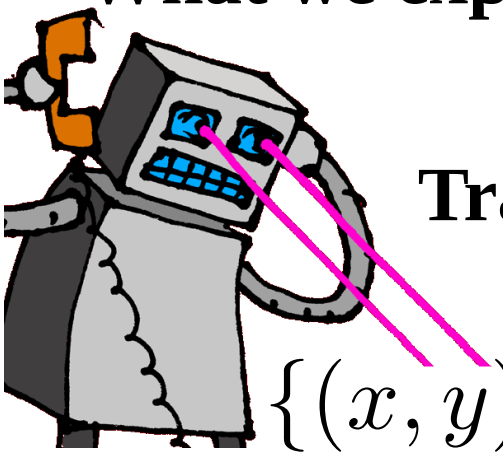
Daniel Hsu (UCSD), Sham M. Kakade (TTI-C), John Langford (Yahoo!), Tong Zhang (Rutgers)

Goal: multilabel prediction; # of possible labels d is very large.

$$f\left(\underbrace{\text{  }}_{x}\right) = \{\text{car, robot, lazer eyes}\} \in \mathbb{R}^d$$

$$(y_{\text{car}} = 1,\ y_{\text{robot}} = 1,\ y_{\text{lazer eyes}} = 1)$$

What we exploit: output sparsity, *i.e.* $\mathbb{E}[y|x]$ is k-sparse, $k \ll d$.



Training and Prediction:

$A \in \mathbb{R}^{m \times d}, \quad m = O(k \log d)$

$$\{(x, y)\} \xrightarrow{\text{Compressed sensing}} \{(x, Ay)\} \mapsto \hat{g} : \mathcal{X} \to \mathbb{R}^m$$

Training data — Compressed training data — Predictor of compressed labels

$$x \mapsto \hat{g}(x) \in \mathbb{R}^m \xrightarrow{\text{Sparse reconstruction}} \hat{y} \in \mathbb{R}^d$$

Test point — Predicted compressed label — Reconstructed sparse label