

# Subset Ranking using Regression

D.Cossock and T.Zhang, COLT 2006

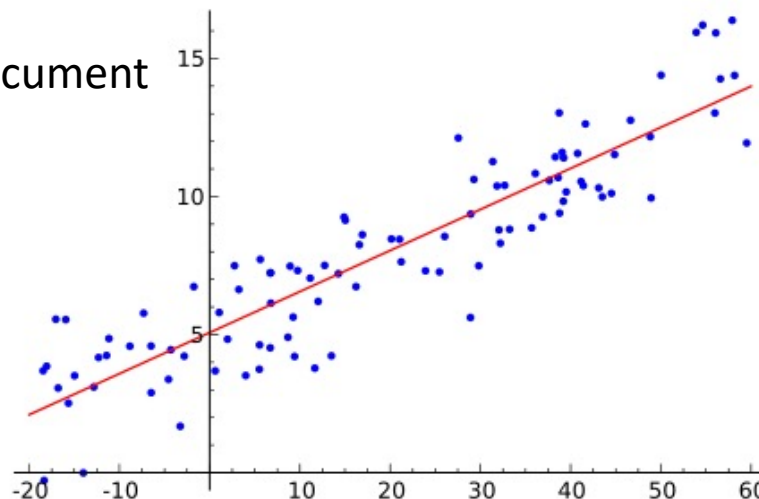
- Fit relevance labels via regression

- $\hat{f} = \arg \min_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^n \left[ \sum_{j=1}^m (f(x_{i,j}, S_i) - y_{i,j})^2 \right]$

- Emphasize more on relevant documents

- $\sum_{j=1}^m w(x_j, S) (f(x_j, S) - y_j)^2 + u \sup_j w'(x_j, S) (f(x_j, S) - \delta(x_j, S))_+^2$

Weights on each document



Most positive document