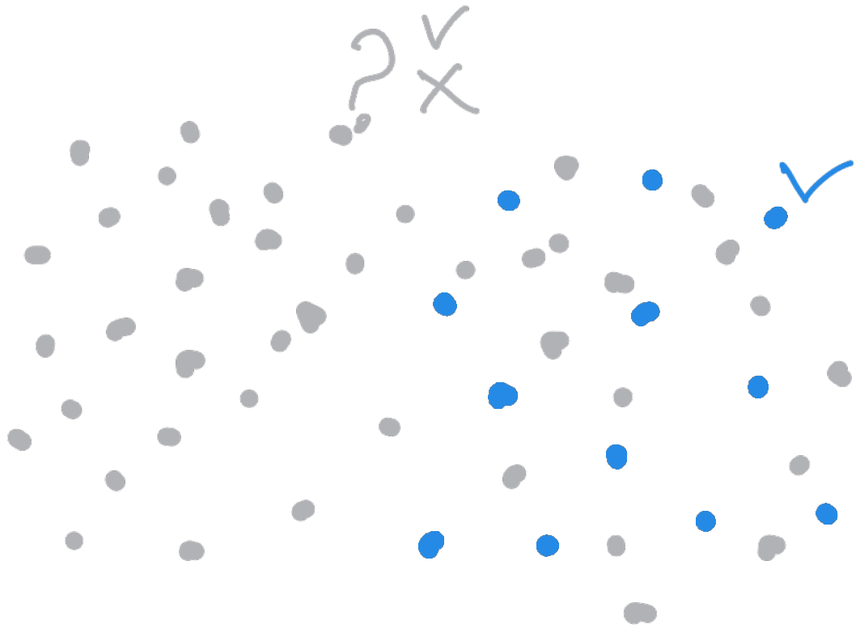
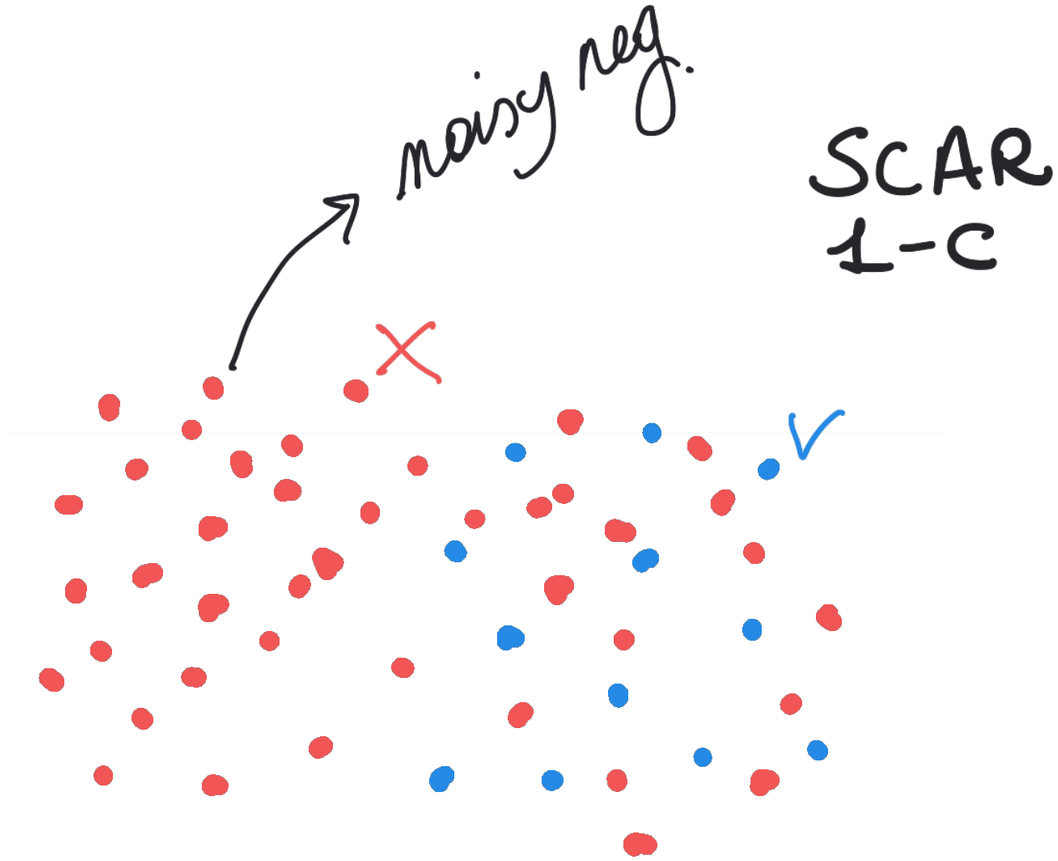
Learning from positive and unlabeled data

5. Biased Learning

Section 5.2 in the survey paper

PU learning = learning with noisy negative labels





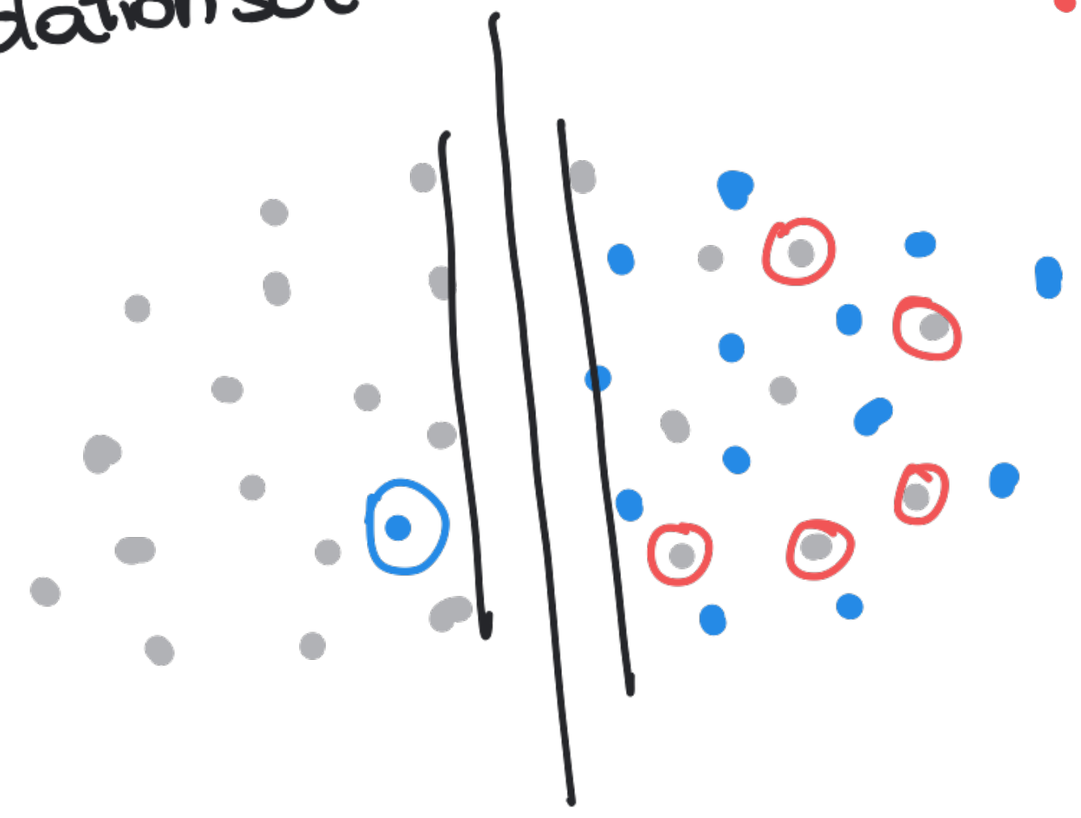
- Deal with the noise:
  - Higher penalties on misclassified positive examples
  - Higher importance on positive examples

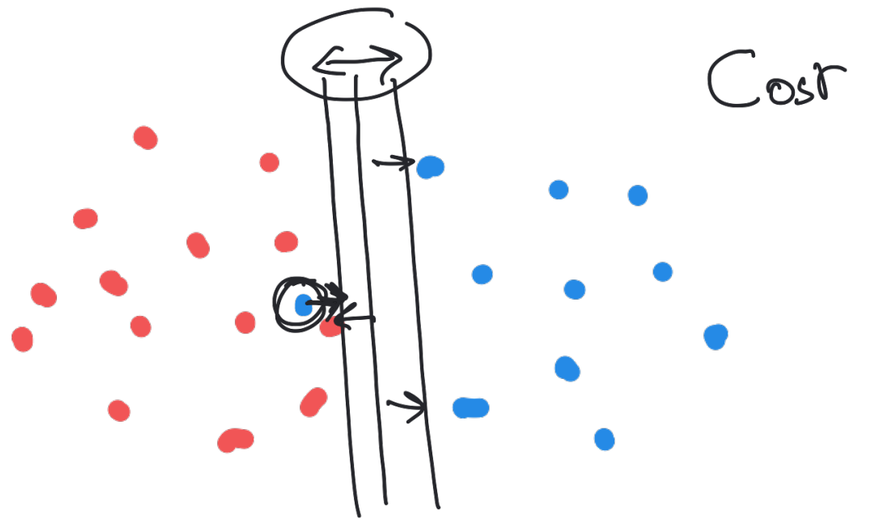
Biased SVM

hyper parameters
tune on PU validation set

$$F_1' = \frac{pr}{\Pr(y=1)}$$

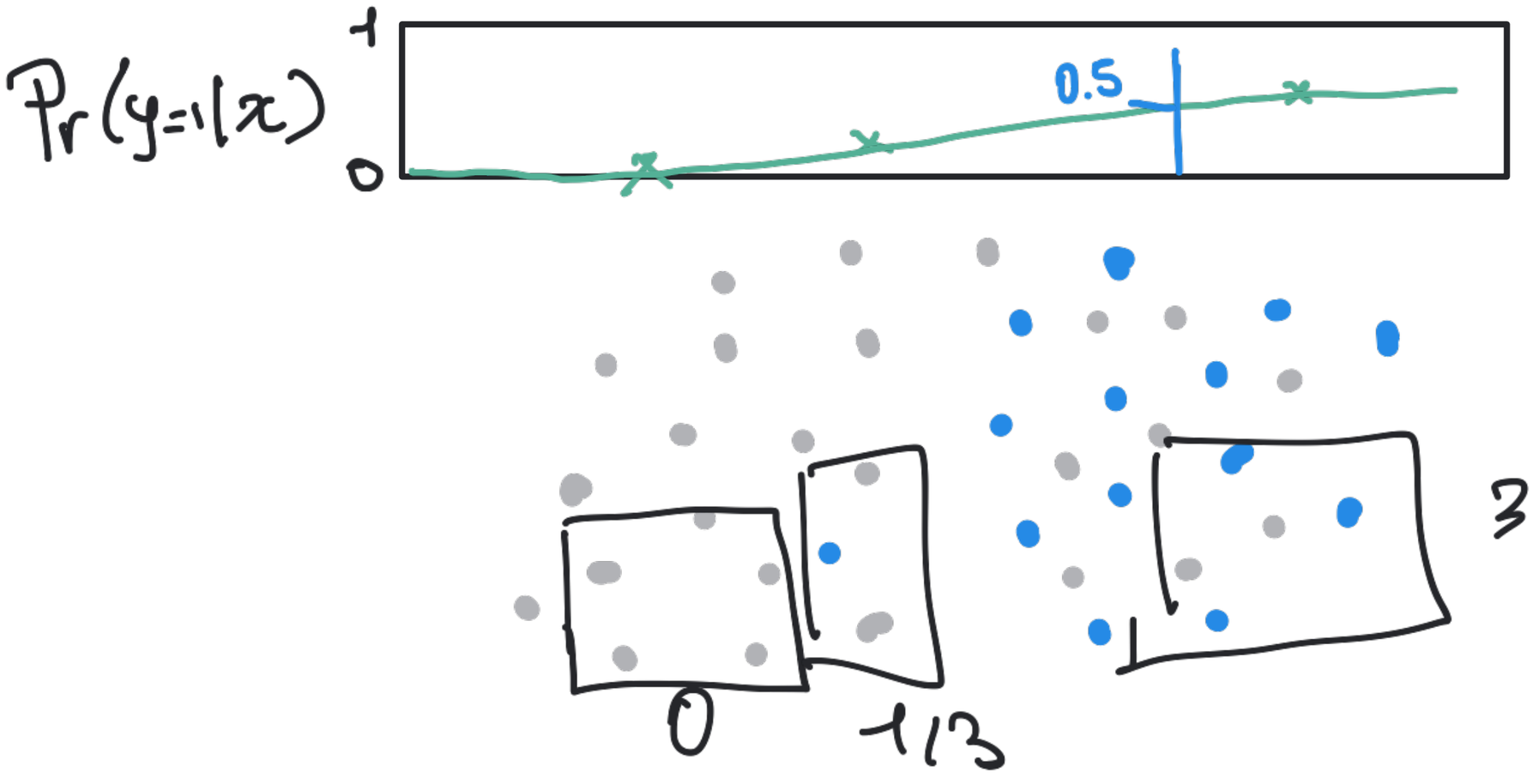
$$= \frac{r^2}{\Pr(\hat{y}=1)}$$





[1] Liu et al. Building Text Classifiers Using Positive and Unlabeled Examples. ICDM. 2003

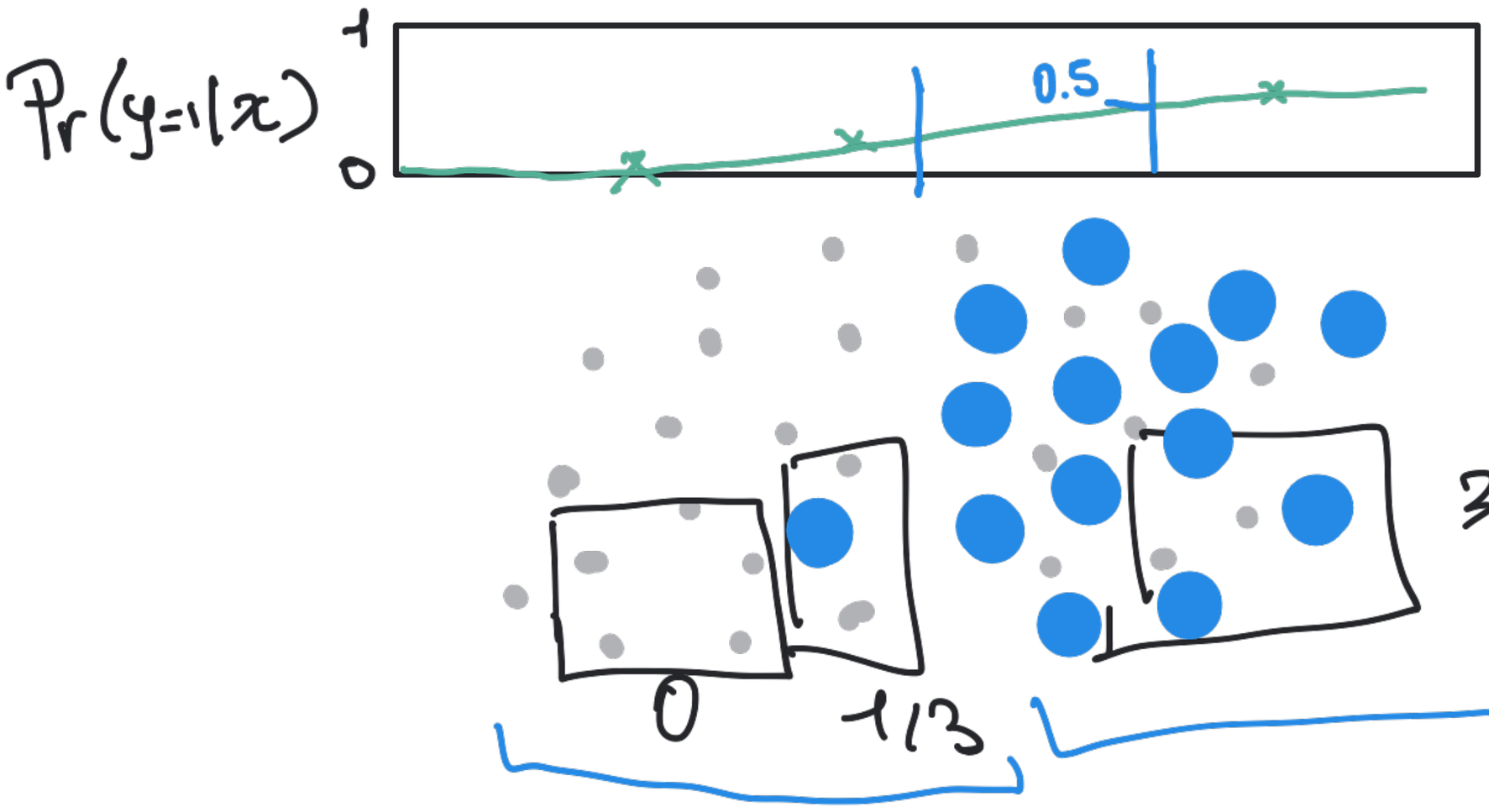
Weighted Logistic Regression



[1] Lee & Liu. Learning with positive and unlabeled examples using weighted logistic regression. ICML. 2003

Weighted Logistic Regression

$\alpha = 0.5$ OR separable

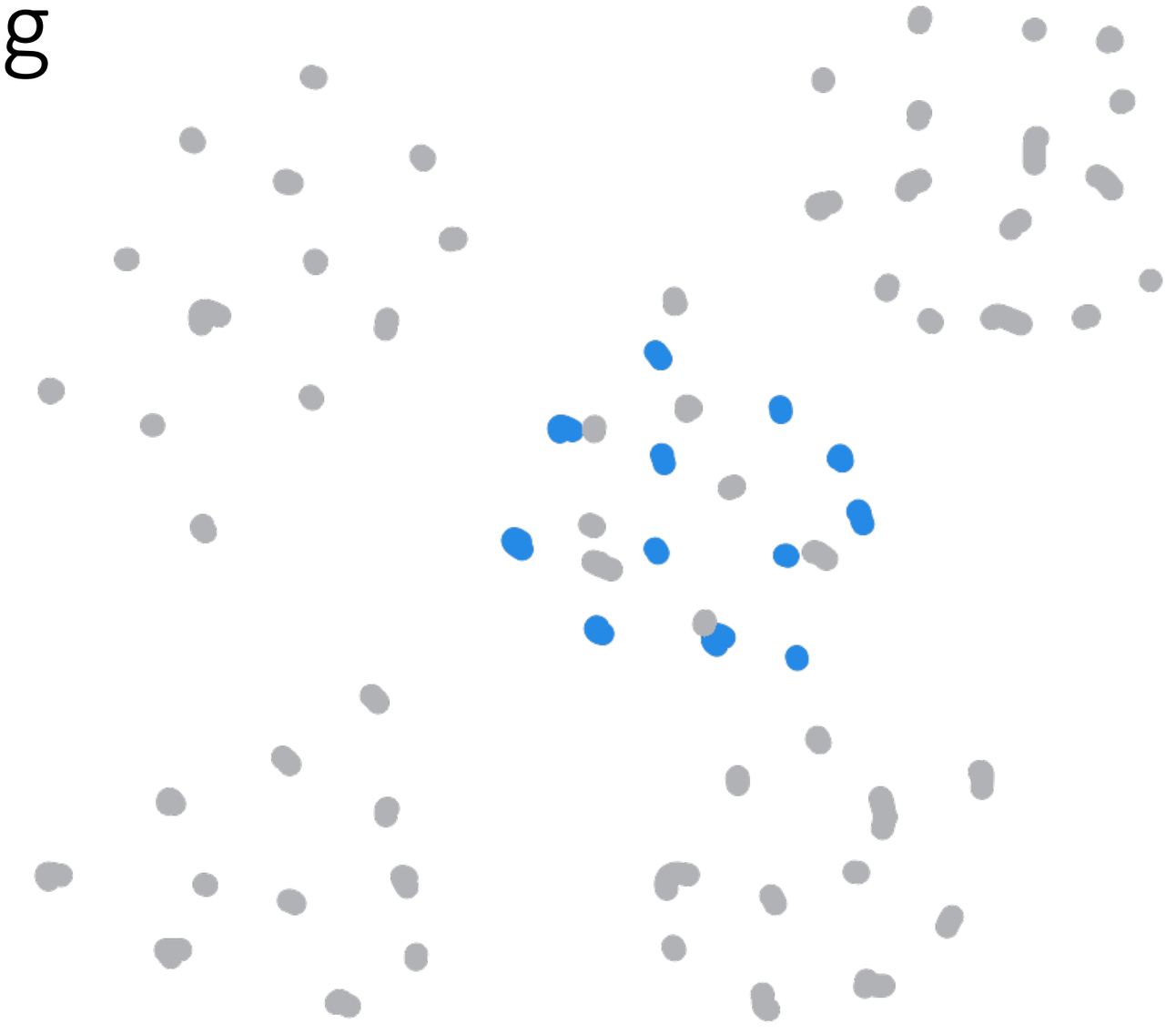


$w^+ = Pr(s=0)$

$w^- = Pr(s=1)$

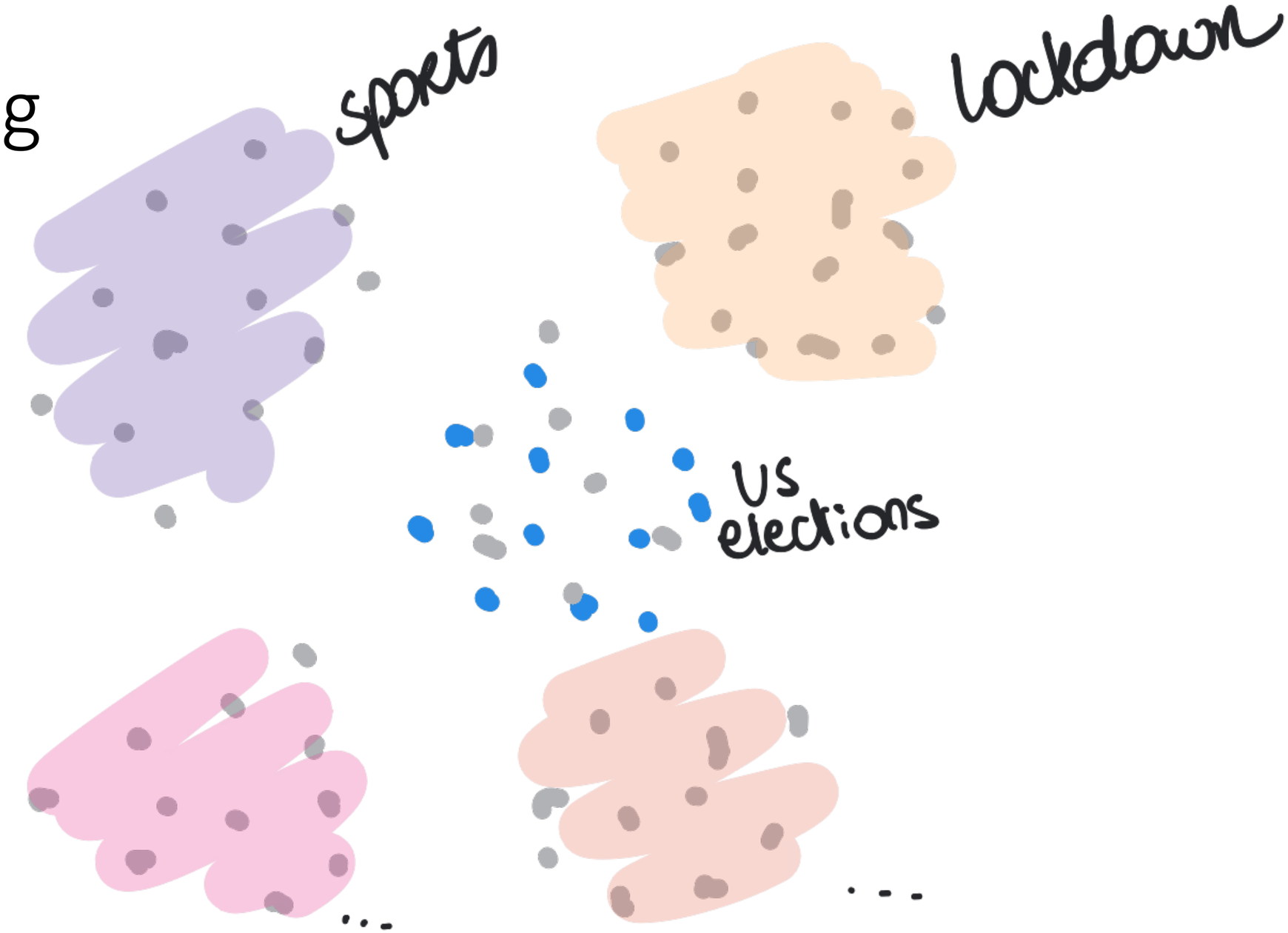
[1] Lee & Liu. Learning with positive and unlabeled examples using weighted logistic regression. ICML. 2003

Clustering



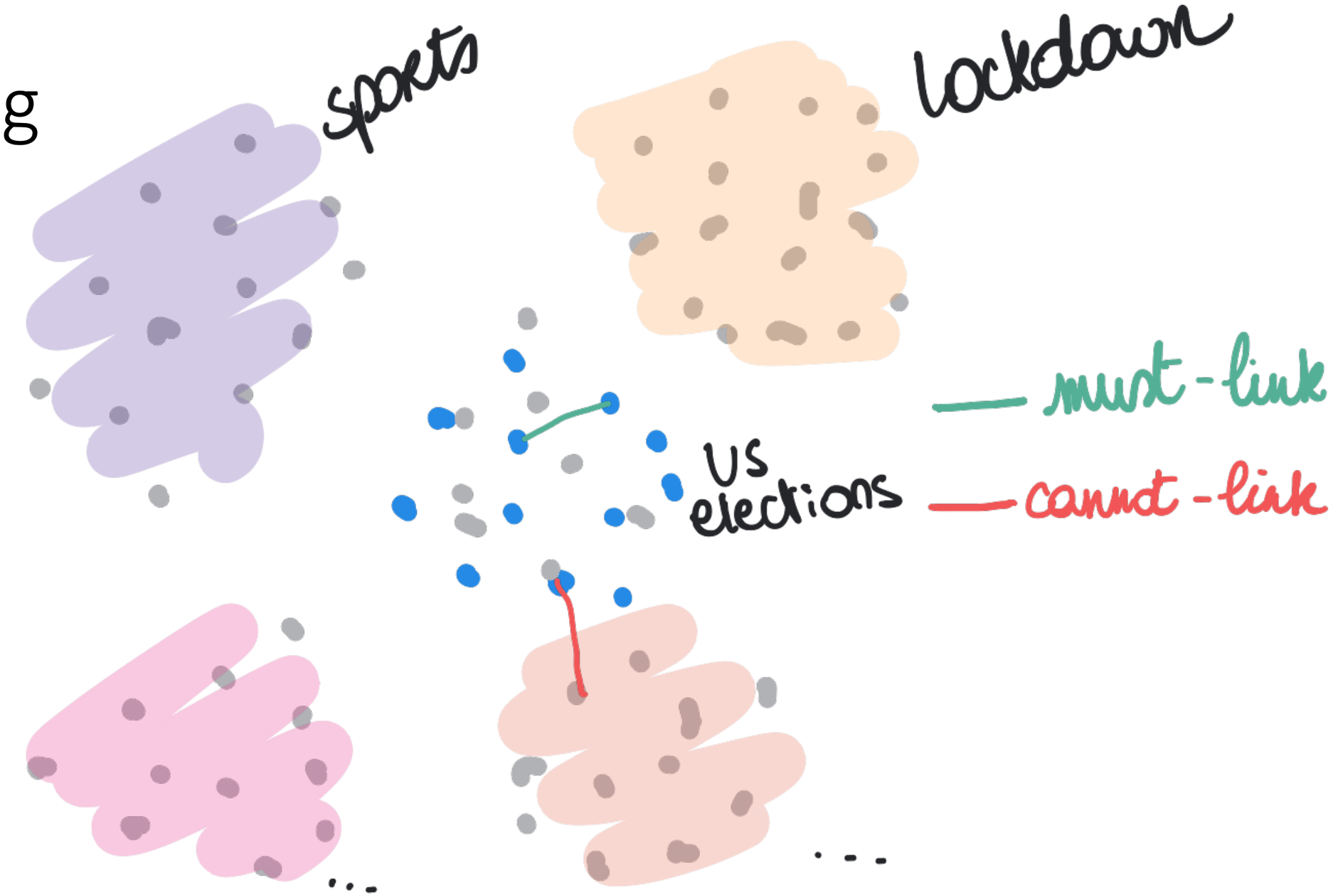
[1] Zhou et al. Learning with positive and unlabeled examples using topic-sensitive plsa. KDD. 2010

Clustering



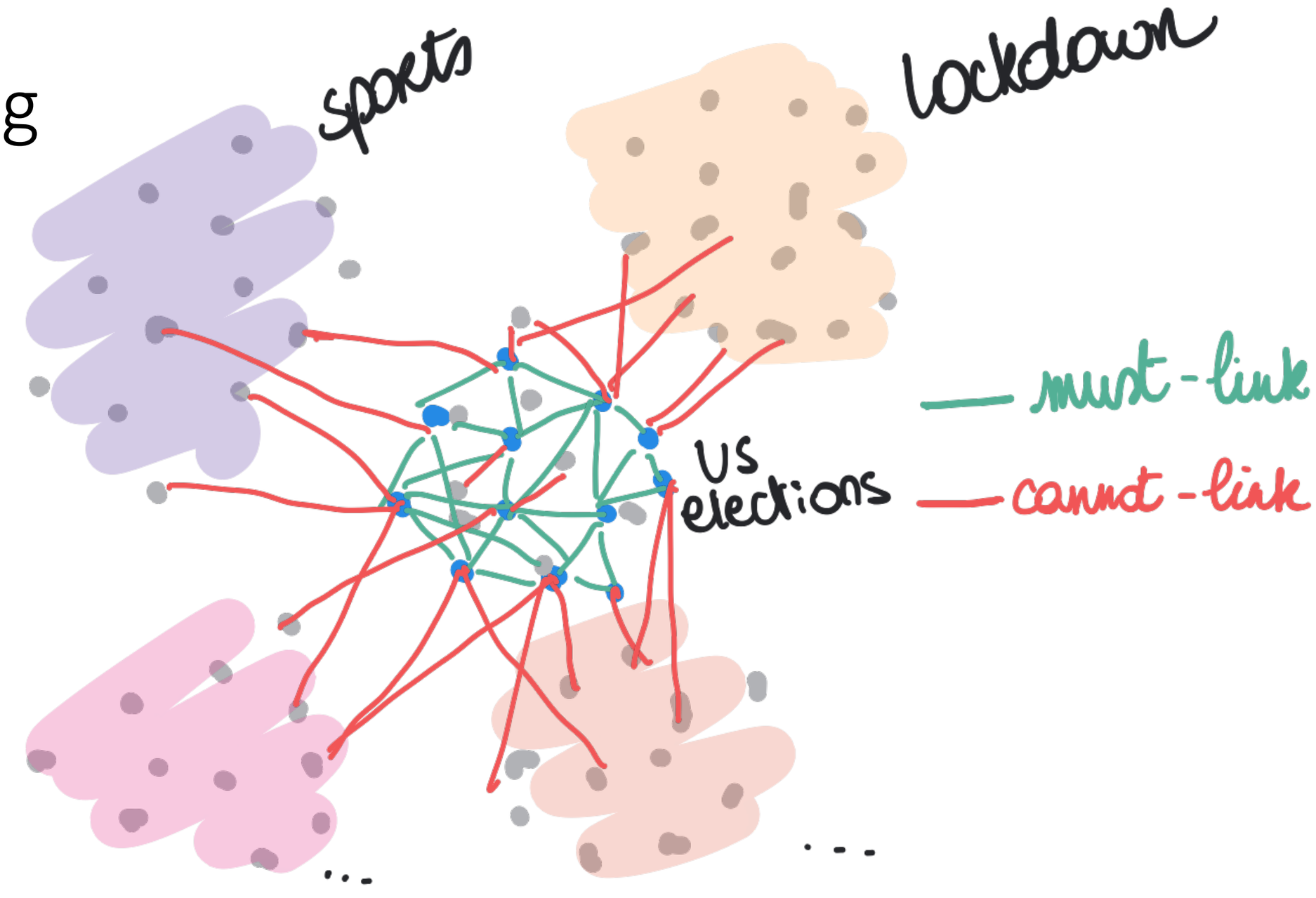
[1] Zhou et al. Learning with positive and unlabeled examples using topic-sensitive plsa. KDD. 2010

Clustering



[1] Zhou et al. Learning with positive and unlabeled examples using topic-sensitive plsa. KDD. 2010

Clustering



[1] Zhou et al. Learning with positive and unlabeled examples using topic-sensitive plsa. KDD. 2010

- Different penalty costs +/-
- Different example weights +/-
- Different constraint weights

How to set penalties/importance?

- F_1' score for tuning $F_1' = \frac{pr}{\Pr(y=1)} = \frac{r^2}{\Pr(\hat{y}=1)}$
- So that a balanced classifier is learned
- Based on true class prior α

Up next...