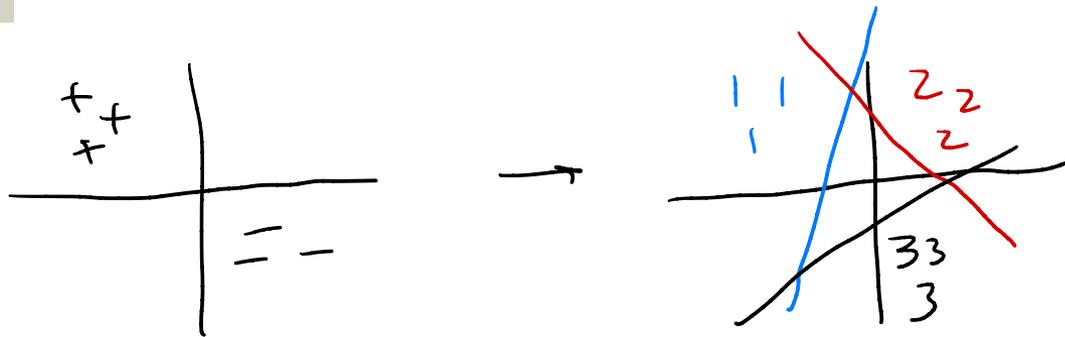
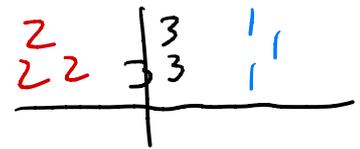


Multiclass Classification



$$y = \{1, 2, 3\}$$

one-vs-all: n binary classifiers



Two techniques

One weight vector per class (different weights, DW)
 or different features (DF) per class

$$DW: \operatorname{argmax}_{y \in \mathcal{Y}} \underbrace{\overline{w}_y^T}_{NN} f(\overline{x})$$

$$DF: \operatorname{argmax}_{y \in \mathcal{Y}} \overline{w}^T \underbrace{f(\overline{x}, y)}_{\text{hypothesized } y}$$

Topic Classification

\bar{x} = too many drug trials, too few patients

$y = \{\text{health, sports, science}\}$

$f(\bar{x}) = \text{bag-of-unigrams}$ [drug, patients, baseball]

$f(\bar{x}) = [1, 1, 0]$

DW: $\bar{w}_{\text{health}} = [2, 5.6, -3]$
 $\Rightarrow 7.6$

$\bar{w}_{\text{sports}} = [1.2, -3.1, 5.7]$
 $\Rightarrow -1.9$

DF: $f(\bar{x}, y) = f(\bar{x})$ replicated for each class

$f(\bar{x}, y = \text{Health}) = [1, 1, 0, 0, 0, 0, 0, 0, 0]$

$f(\bar{x}, y = \text{Sports}) = [0, 0, 0, 1, 1, 0, 0, 0, 0]$

Indicator (sent
contains word i
 $\wedge y = \text{Sports}$)

$\bar{w} = [2, 5.6, -3, 1.2, -3.1, 5.7, \dots]$