Loss function

- Hemming loss:
  $\frac{1}{L} \sum_{i \in [1,2,\ldots L]} (\hat{y}_i \text{ xor } y_i)$
- exact loss
  0 if $\hat{y} = y$ else 1
- trade off
  $\sum_{i \in \hat{Y}} [1 - a - y_i]$

Algorithm

We can consider each "superarm" of the label powerset as a seperate class and then from there we can apply multiclass method.Let k be the cardinality of the labelset
for i = 1,2,....T:
1)input x
2)calculate $haty = argmax(W^T x)$ 3)calculate
$p_r = (1 - \gamma)|r = \hat{y}| + \frac{\gamma}{k}$
4)predict $\tilde{y}$ according to the distribution p and recieve $\tilde{y}$ intersection Y 5)update w according to $w = w - \bigtriangledown L$ where L is the loss

Related works

randomised k set : randomly break down the entire powerset of labels into m labelsets which have size k. Train m classifiers and then take the union of their outputs.