

$X$ : text repr. (feat vec)

$y$ : doc label

Log Reg:  $P(y|x; w) = \frac{1}{1 + e^{-w^T x}}$

Training time,  $Y^{tr} = y_1, y_2 \dots y_D$   
 $X^{tr} = x_1, x_2 \dots x_D$

MLE training

avg max  $\vec{w}$   $\log P(Y^{tr} | X^{tr}; w)$

Naive Bayes  $P(y) =$  prior prob. of label

$P(x|y) =$  Unigram LM for  $y$

$P(y|x) =$  use Bayes rule for  $y$

MLE training

avg max  $\theta$   $P(X^{tr}, Y^{tr}; \theta) = \text{avg max}_{\theta} P(X^{tr} | Y^{tr}) P(Y^{tr})$

$$P(y=1 | x) = \frac{1}{1 + \dots}$$

$$X_3 = 1 \text{ (Hilf A200) }$$

$$y_3 = 0$$

$$P(y = y_3 | X_3) = P(y = 0 | X_3)$$