

Discriminative Classifier 1: Logistic Regression

Binary Response Variable: $Y \in \{0,1\}$

Predictors: $X = (x_1, x_2, \dots, x_M)$, $x_i \in \mathfrak{R}$

$$Y = \begin{cases} 1 & \text{category}(d) = \theta_1 \\ 0 & \text{category}(d) = \theta_2 \end{cases}$$

Modeling $p(Y|X)$ directly

Allow many other features than words!

$$\log \frac{p(\theta_1 | d)}{p(\theta_2 | d)} = \log \frac{p(Y = 1 | X)}{p(Y = 0 | X)} = \log \frac{p(Y = 1 | X)}{1 - p(Y = 1 | X)} = \beta_0 + \sum_{i=1}^M x_i \beta_i \quad \beta_i \in \mathfrak{R}$$

$$p(Y = 1 | X) = \frac{e^{\beta_0 + \sum_{i=1}^M x_i \beta_i}}{e^{\beta_0 + \sum_{i=1}^M x_i \beta_i} + 1}$$

