

# Anatomy of Naïve Bayes Classifier

Two categories:  $\theta_1$  and  $\theta_2$

$$\text{score}(d) = \log \frac{p(\theta_1 | d)}{p(\theta_2 | d)} = \log \frac{p(\theta_1) \prod_{w \in V} p(w | \theta_1)^{c(w,d)}}{p(\theta_2) \prod_{w \in V} p(w | \theta_2)^{c(w,d)}}$$

$$= \underbrace{\log \frac{p(\theta_1)}{p(\theta_2)}}_{\text{Category bias } (\beta_0) \text{ doesn't depend on } d!} + \sum_{w \in V} \underbrace{c(w,d)}_{\text{Sum over all words (features } \{f_i\})} \underbrace{\log \frac{p(w | \theta_1)}{p(w | \theta_2)}}_{\text{Weight on each word (feature) } \beta_i}$$

Feature value:  $f_i = c(w,d)$



Generalize

$$d = (f_1, f_2, \dots, f_M), \quad f_i \in \mathcal{R}$$

$$\text{score}(d) = \beta_0 + \sum_{i=1}^M f_i \beta_i \quad \beta_i \in \mathcal{R}$$

= Logistic Regression!