

Exercises

4.1 Assume the following likelihoods for each word being part of a positive or negative movie review, and equal prior probabilities for each class.

	pos	neg
I	0.09	0.16
always	0.07	0.06
I like	0.29	0.06
foreign	0.04	0.15
films	0.08	0.11

$$P(\text{pos}) = P(\text{neg}) = 0.5$$

$$C_{NB} = \underset{c \in C}{\operatorname{argmax}} P(c) \prod_{i \in \text{positions}} P(w_i | c)$$

$$\begin{aligned} C_{\text{pos}} &= P(\text{pos}) P(I | \text{pos}) P(\text{always} | \text{pos}) P(\text{I like} | \text{pos}) P(\text{foreign} | \text{pos}) P(\text{films} | \text{pos}) \\ &= (0.5)(0.09)(0.07)(0.29)(0.04)(0.08) \\ &= 2.92 \times 10^{-6} \end{aligned}$$

$$\begin{aligned} C_{\text{neg}} &= P(\text{neg}) P(I | \text{neg}) P(\text{always} | \text{neg}) P(\text{I like} | \text{neg}) P(\text{foreign} | \text{neg}) P(\text{films} | \text{neg}) \\ &= (0.5)(0.16)(0.06)(0.06)(0.15)(0.11) \\ &= 4.75 \times 10^{-6} \end{aligned}$$

$$C_{\text{pos}} < C_{\text{neg}}$$

Naive Bayes would assign class Neg