Giải:

Gọi x_i là văn bản thuộc lớp c_j , với i ∈ [1, m]; m là số lượng văn bản thuộc lớp j

$\Rightarrow \sum_{i=1}^{m} x_i^j = Count(c_j)$

Xét các biến x_i là độc lập

Theo phân phối Categorical (trường hợp tổng quát của phân phối Bermoulli):

$$P(x_i|\hat{P}(c_j) = \prod_{i=1}^{m} \hat{P}(c_j)^{x_i}$$

Đánh giá $\hat{P}(c_j)$ dựa trên Maximum log-likelihood:

$$\mathscr{L}\hat{P}(c_j) = \arg\max[P(x_1, x_2, x_3..., x_m|\hat{P}(c_j)] = \arg\max \prod_{i=1}^{m} P(x_i|\hat{P}(c_j)) = \arg\max \prod_{i=1}^{m}\prod_{j=1}^{k} \hat{P}(c_j)^{x_i}$$

$$= \arg\max \prod_{j=1}^{k} \hat{P}(c_j)^{\sum_{i=1}^{m} x_i} = \arg\max \prod_{j=1}^{k} \hat{P}(c_j)^{Count(c_j)} = \arg\max \sum_{j=1}^{k} Count(c_j) \log(\hat{P}(c_j))$$

Điều kiện chuẩn :$\sum_{j=1}^{k} \hat{P}(c_j) = 1$

⇒ Áp dụng phương pháp nhân tử Lagrange: