

Back to Factoring out Background Words

Text Mining Paper
d

... text mining...
is... clustering...
we... Text, the

$$P(w | \theta_d)$$

text 0.04
mining 0.035
association 0.03
clustering 0.005
...
the 0.000001

θ_d

$$p(\theta_d) + p(\theta_B) = 1$$

$$P(\theta_d) = 0.5$$

Topic
Choice

$$P(w | \theta_B)$$

the 0.03
a 0.02
is 0.015
we 0.01
food 0.003
...
text 0.000006

θ_B

$$P(\theta_B) = 0.5$$