LMs for Topic Representation

B General Background English Text

the 0.03
a 0.02
is 0.015
we 0.01
...
food 0.003
computer 0.00001
text 0.000006
…

Background LM: $p(w|B)$

C Computer Science Papers

the 0.032
a 0.019
is 0.014
we 0.011
...
computer 0.004
software 0.0001
text 0.00006
…

Collection LM: $p(w|C)$

d **Text mining paper**

the 0.031
…
text 0.04
mining 0.035
association 0.03
clustering 0.005
computer 0.0009
…
food 0.000001

Document LM: $p(w|d)$