term frequency vs vocabulary size 1-grams 1e-01 -1e-03 relative word frequency 1e-05 **-**1e-07 -1e+03 1e+05 1e+01 vocab size