

Tf Idf Representation

Tf - term frequency ← like bag of words
Idf - inverse document frequency ← weighting by how often word occurs in corpus

Quiz: would you weight common words higher,
or rare words?



common



rare

