Pivoted Document Length Normalization
vector space model
the VS model doesn’t specify...
empirical distribution of words
Word Ranks in 'Oliver Twist'
VS heuristics
term frequency
inverse document frequency
document length normalization
DLN: pivoted length normalization

Singhal, Buckley, and Mitra (SIGIR 1996)
\[
\frac{TF \cdot IDF}{(1 - slope) + \text{slope} \cdot \frac{\text{old norm}}{\text{avg old norm}}}
\]
$(1 - \text{slope}) \cdot \text{pivot} + \text{slope} \cdot \text{num\_unique}$
\[ TF = \frac{1 + \log(tf)}{1 + \log(\text{avg}(tf))} \]
\[
\frac{1+\log(tf)}{1+\log(\text{avg}(tf))} \\
(1 - \text{slope}) \cdot \text{avg}(\text{num\_unique}) + \text{slope} \cdot \text{num\_unique}
\]