TF-IDF mathematical intuition

TF = Term frequency = $\frac{\#\text{ times word appears in sentence}}{\text{total \# words in sentence}}$

IDF = Inverse document freq = $\log\left(\frac{\text{total \# of sentences}}{\#\text{ sentences word appears in}}\right)$

TF-IDF score for each word in sentence: TF * IDF

Example

s_1: cute cat
s_2: cute dog
s_3 cute cat dog

TF

	s_1	s_2	s_3
cute	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$
cat	$\frac{1}{2}$	0	$\frac{1}{3}$
dog	0	$\frac{1}{2}$	$\frac{1}{3}$

IDF

words	IDF
cute	$\log(\frac{3}{3})$
cat	$\log(\frac{3}{2})$
dog	$\log(\frac{3}{2})$

	cute	cat	dog
s_1	0	$\frac{1}{2}\log(\frac{3}{2})$	0
s_2	0	0	$\frac{1}{2}\log(\frac{3}{2})$
s_3	0	$\frac{1}{3}\log(\frac{3}{2})$	$\frac{1}{3}\log(\frac{3}{2})$

x_{11} s_1 vector: $[0, \frac{1}{2}\log\frac{3}{2}, 0]$

x_{21} s_2 vector: $[0, 0, \frac{1}{2}\log(\frac{3}{2})]$

x_{31} s_3 vector: $[0, \frac{1}{3}\log[\frac{3}{2}, \frac{1}{3}\log(\frac{3}{2})]$

Important terms get higher weights.

$$P(Y=1|h) = \frac{e^{b_0+b_1(\text{TF-IDF of cute})+b_2(\text{TF-IDF cat})+b_3(\text{TF-IDF dog})}}{1+e^{b_0+b_1(\cdot)+b_2(\cdot)+b_3(\cdot)}}$$

logistic reg function: $P(Y=1|h=h) = \frac{e^{B_0+B_1x_{i1}}}{1+e^{B_0+B_1x_{i1}}}$