

# TF-IDF

By : Ahmed Ibrahim

# Example

- Sen 1 : good boy
- Sen 2 : good girl
- Sen 3 : boy gilr good
- First we calculate TF  $\rightarrow$   $TF = \frac{\text{No.of rep in a sen}}{\text{No.of words in a sen}}$

$$IDF = \log\left(\frac{\text{no. of sentences}}{\text{No.of sentences containing words}}\right)$$

	sen1	sen2	sen3
Good	1/2	1/2	1/3
Boy	1/2	0	1/3
girl	0	1/2	1/3

x

words	IDF
Good	Log(3/3)=0
Boy	Log(3/2)
girl	Log(3/2)

	good	boy	girl	output
Sen 1	0	$\frac{1}{2} * \log(3/2)$	0	
Sen 2	0	0	$\frac{1}{2} * \log(3/2)$	
Sen 3	0	$\frac{1}{3} * \log(3/2)$	$\frac{1}{3} * \log(3/2)$	









