

	Doc	Words	Author
Training	1	W1 W2 W3 W4 W5	C ( <a href="#">Christopher Marlowe</a> )
	2	W1 W1 W4 W3	C ( <a href="#">Christopher Marlowe</a> )
	3	W1 W2 W5	C ( <a href="#">Christopher Marlowe</a> )
	4	W5 W6 W1 W2 W3	W ( <a href="#">William Stanley</a> )
	5	W4 W5 W6	W ( <a href="#">William Stanley</a> )
	6	W4 W6 W3	F ( <a href="#">Francis Bacon</a> )
	7	W2 W2 W4 W3 W5 W5	F ( <a href="#">Francis Bacon</a> )
Test	8 (Hamlet)	W1 W4 W6 W5 W3	?

$$P(C) = N_c / N$$

$$= \text{Number of class C} / \text{Total number of classes} = 3/7$$

$$P(W) = N_w / N$$

$$= \text{Number of class W} / \text{Total number of classes} = 2/7$$

$$P(F) = N_f / N$$

$$= \text{Number of class F} / \text{Total number of classes} = 2/7$$

$P(W1|C)$  = The probability that the word W1 appears on the class C document

$$= (\text{count}(W1, C) + 1) / (\text{count}(C) + |V|)$$

$$= (4 + 1) / (12 + 6)$$

$$= 5/18$$

$P(W1|W)$  = The probability that the word W1 appears on the class W document

$$= (\text{count}(W1, W) + 1) / (\text{count}(W) + |V|)$$

$$= (1 + 1) / (8 + 6)$$

$$= 2/14$$

$P(W1|F)$  = The probability that the word W1 appears on the class F document

$$= (\text{count}(W1, F) + 1) / (\text{count}(F) + |V|)$$

$$= (0 + 1) / (9 + 6)$$

$$= 1/15$$

$P(W3|C)$  = The probability that the word W3 appears on the class C document

$$= (\text{count}(W3, C) + 1) / (\text{count}(C) + |V|)$$

$$= (2 + 1) / (12 + 6)$$

$$= 3/18$$

$P(W3|W)$  = The probability that the word W3 appears on the class W document

$$= (\text{count}(W3, W) + 1) / (\text{count}(W) + |V|)$$

$$= (1 + 1) / (8 + 6)$$

$$= 2/14$$

**P(W3|F)** = The probability that the word W3 appears on the class F document  

$$= (\text{count}(W3, F) + 1) / (\text{count}(F) + |V|)$$

$$= (2 + 1) / (9 + 6)$$

$$= 3/15$$

**P(W4|C)** = The probability that the word W4 appears on the class C document  

$$= (\text{count}(W4, C) + 1) / (\text{count}(C) + |V|)$$

$$= (2 + 1) / (12 + 6)$$

$$= 3/18$$

**P(W4|W)** = The probability that the word W4 appears on the class W document  

$$= (\text{count}(W4, W) + 1) / (\text{count}(W) + |V|)$$

$$= (1 + 1) / (8 + 6)$$

$$= 2/14$$

**P(W4|F)** = The probability that the word W4 appears on the class F document  

$$= (\text{count}(W4, F) + 1) / (\text{count}(F) + |V|)$$

$$= (2 + 1) / (9 + 6)$$

$$= 3/15$$

**P(W5|C)** = The probability that the word W5 appears on the class C document  

$$= (\text{count}(W5, C) + 1) / (\text{count}(C) + |V|)$$

$$= (2 + 1) / (12 + 6)$$

$$= 3/18$$

**P(W5|W)** = The probability that the word W5 appears on the class W document  

$$= (\text{count}(W5, W) + 1) / (\text{count}(W) + |V|)$$

$$= (2 + 1) / (8 + 6)$$

$$= 3/14$$

**P(W5|F)** = The probability that the word W5 appears on the class F document  

$$= (\text{count}(W5, F) + 1) / (\text{count}(F) + |V|)$$

$$= (2 + 1) / (9 + 6)$$

$$= 3/15$$

**P(W6|C)** = The probability that the word W6 appears on the class C document  

$$= (\text{count}(W6, C) + 1) / (\text{count}(C) + |V|)$$

$$= (0 + 1) / (12 + 6)$$

$$= 1/18$$

**P(W6|W)** = The probability that the word W6 appears on the class W document  

$$= (\text{count}(W6, W) + 1) / (\text{count}(W) + |V|)$$

$$= (2 + 1) / (8 + 6)$$

$$= 3/14$$

**$P(W_6|F)$**  = The probability that the word  $W_6$  appears on the class F document  
 $= (\text{count}(W_6, F) + 1) / (\text{count}(F) + |V|)$   
 $= (1 + 1) / (9 + 6)$   
 $= 2/15$

$P(C)$	$P(W)$	$P(F)$
$3/7$	$2/7$	$2/7$
$P(W_1 C)$	$P(W_1 W)$	$P(W_1 F)$
$5/18$	$2/14$	$1/15$
$P(W_2 C)$	$P(W_2 W)$	$P(W_2 F)$
$3/18$	$2/14$	$3/15$
$P(W_3 C)$	$P(W_3 W)$	$P(W_3 F)$
$3/18$	$2/14$	$3/15$
$P(W_4 C)$	$P(W_4 W)$	$P(W_4 F)$
$3/18$	$2/14$	$3/15$
$P(W_5 C)$	$P(W_5 W)$	$P(W_5 F)$
$3/18$	$3/14$	$3/15$
$P(W_6 C)$	$P(W_6 W)$	$P(W_6 F)$
$1/18$	$3/14$	$2/15$

**$P(C|d_8)$**  =  $P(C) * P(W_1|C) * P(W_4|C) * P(W_6|C) * P(W_5|C) * P(W_3|C)$   
 $= 3/7 * 5/18 * 3/18 * 1/18 * 3/18 * 3/18$   
 $= 0.00003061924$

**$P(W|d_8)$**  =  $P(W) * P(W_1|W) * P(W_4|W) * P(W_6|W) * P(W_5|W) * P(W_3|W)$   
 $= 2/7 * 2/14 * 2/14 * 3/14 * 3/14 * 2/14$   
 $= 0.00003824936$

**$P(F|d_8)$**  =  $P(F) * P(W_1|F) * P(W_4|F) * P(W_6|F) * P(W_5|F) * P(W_3|F)$   
 $= 2/7 * 1/15 * 3/15 * 2/15 * 2/15 * 3/15$   
 $= 0.00001354497$

- **Does d8 belong to C or W or F?**

Compare  $P(C|d8)$ ,  $P(W|d8)$ ,  $P(F|d8)$ :

$$P(C|d8) = 0.00003061924$$

$$P(W|d8) = 0.00003824936 \Rightarrow \text{Largest}$$

$$P(F|d8) = 0.00001354497$$

**So, d8 belongs to W.**