1.

a.   (i) $4/5$

(ii) $4/5$

(iii) $2/3$

(iv) $2 \times \dfrac{P \times r}{P+r} = \dfrac{4/5 \times 2/3}{4/5 + 2/3} \times 2 = 8/11$

b.

		Predicted.	
		Men (0)	Women (1)
Actual	Men (0)	8	1
	Women (1)	2	4

c. $P(\text{actual} = \text{woman})$ means the probability of the customer being female which is the prior probability.

$P(\text{predict} = \text{women} \mid \text{actual} = \text{women})$ is the conditional probability that given the customer is female, how likely the algorithm will classify her as women.

d.   Recall

2.

A.
   (i) $1/3$

(ii) $1/3 \times 1/2 + 1/3 \times 1/2 = 1/3$

(iii) $1/3$

(iv) $P(y = N_o \mid x = bad) = \dfrac{P(x = bad \mid y = N_o) \cdot P(y = N_o)}{P(x = bad)}$

$$= \frac{\frac{1}{3} \times \frac{1}{2}}{\frac{1}{2} \times \frac{1}{3} + \frac{1}{2} \times \frac{1}{3}} = \frac{1}{2}$$

B.

$$p(\text{"love"}) = \frac{1}{6}$$

$$p(\text{"movie"}) = \frac{2}{3}$$

$$p(\text{"love"}, \text{"movie"}) = \frac{1}{6}$$

$$p(\text{"love"}) \cdot p(\text{"movie"}) = \frac{1}{6} \times \frac{2}{3} = \frac{1}{9} \neq \frac{1}{6}$$

No, they are not independent.

3.

a.

	trendy	jeans	old	blue	red	wool
product A	2	7/4	0	0	0	0
product B	0	7/4	2	2	0	0
product C	2	7/4	2	2	5/2	5/2
IDF	2	7/4	2	2	5/2	5/2

b. product B.

$$\cos(B, \text{query}) = 0.867$$

$$\cos(C, \text{query}) = 0.506$$

$$\cos(B, \text{query}) > \cos(C, \text{query})$$

4.

A.

	I	love	go	to	store	he	work	at	resteraunt	be	close	today	END
I	0	1	0	0	0	0	0	0	0	1	0	0	0
love	0	0	1	0	0	0	1	0	0	0	0	0	0
go	0	0	0	1	0	0	0	0	0	0	0	0	0
to	0	0	0	0	1	0	0	0	1	0	0	0	0
store	0	0	0	0	0	0	0	0	0	1	0	0	1
he	0	1	0	0	0	0	0	0	0	1	0	0	0
work	0	0	0	0	0	0	0	1	0	0	0	0	1
at	0	0	0	0	0	0	0	0	1	0	0	0	0
resteraunt	0	0	0	0	0	0	0	0	0	0	0	0	2
be	0	0	1	0	0	0	1	0	0	0	1	0	0
close	0	0	0	0	0	0	0	0	0	0	1	0	0
today	1	0	0	0	0	0	0	0	0	0	0	0	1
Start	1	0	0	0	1	2	0	0	0	0	0	1	0

B. $P(X_0 = Start, X_1 = I, X_2 = love, X_3 = working, X_4 = END)$

$$= \prod_{i=1}^{N+1} P(X_i \mid X_{i-1})$$

$$= 0.2 \times 0.5 \times 0.5 \times 0.5 = 0.025$$

c. These two sentences are not equally long.

It is not okay to compare them directly. We need to normalize them by using preplexity.