Naïve Bayes
for Sentiment Analysis

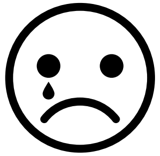
Note and Disclaimer

- 1. The PDF slides are part of a YouTube tutorial: https://youtu.be/YVC-L_lLQb4

- 2. The PDF slides and the video tutorial provided are for informational purposes only. Although the author has made efforts to ensure the quality of the slides and tutorial, their accuracy or completeness is not guaranteed. It is advisable to cross-verify information with other sources and exercise your own judgment when making decisions.

- 3. This tutorial is intended for personal use only. Therefore, please refrain from distributing it through any means.

	Reviews	Sentiment Class
Training Data	Great product	P
	A good product	P
	I like it	P
	Difficult to use	N
	complicated product	N



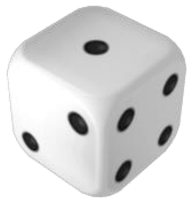
	Reviews	Sentiment Class
Test Data	I like the product	

?

Question:

Is “I like the product” positive or negative, based on the training data?

$$p(A|B) = \frac{p(AB)}{p(B)} = \frac{p(A)p(B|A)}{p(B)}$$



Event A: Get an even number (2, 4, or 6)

Event B: Get a number greater than 3 (4, 5, or 6)

Probability of A given that B has occurred: $p(A|B) = \frac{2}{3}$

- $P(AB) = P(A \cap B) = \frac{2}{6}$
- $P(B) = \frac{1}{2}$
- $\frac{p(AB)}{p(B)} = \frac{2/6}{1/2} = \frac{2}{3}$

- $p(A) = \frac{1}{2}$
- $p(B|A) = \frac{2}{3}$
- $p(B) = \frac{1}{2}$
- $\frac{p(B|A)p(A)}{p(B)} = \frac{\frac{2}{3} \times \frac{1}{2}}{\frac{1}{2}} = \frac{2}{3}$

$$p(A|B) = \frac{p(A)p(B|A)}{p(B)}$$

B: "I like the product"

A: Positive (P) or negative (N)

$$\left\{ \begin{array}{l} \dfrac{p(P)p(B|P)}{p(B)} \\ \\ \dfrac{p(N)p(B|N)}{p(B)} \end{array} \right.$$

$$p(A|B) \left\{ \begin{array}{l} p(\mathrm{P}\ |\text{"}\textit{I like the product}\text{"}) \\ p(\mathrm{N}\ |\text{"}\textit{I like the product}\text{"}) \end{array} \right.$$

Question:

Which one is greater, $p(\mathrm{P}\ |\text{"}\textit{I like the product}\text{"})$ or $p(\mathrm{N}\ |\text{"}\textit{I like the product}\text{"})$?

proportional

$$p(A|B) = \frac{p(A)p(B|A)}{p(B)} \propto p(A)p(B|A)$$

Prior Likelihood

B: “I like the product”

A: Positive (P) or negative (N)

Posterior

$$p(A|B) \propto p(A)p(B|A)$$

$$p(P)p(\text{"}I\ like\ the\ product\text{"}|P)$$

$$p(N)p(\text{"}I\ like\ the\ product\text{"}|N)$$

?

	Reviews	Sentiment
Training Data	Great product	P
	A good product	P
	I like it	P
	Difficult to use	N
	complicated product	N

$$p(P) = \frac{3}{5}$$

$$p(N) = \frac{2}{5}$$

$$p(P) \times p(\text{"}I\ like\ the\ product\text{"}|P)$$

$$p(N) \times p(\text{"}I\ like\ the\ product\text{"}|N)$$

- The "naive" assumption in naive Bayes is that the **features used to describe an observation** are assumed to be **conditionally independent given the class label**.

  - Despite its simplicity, the naive Bayes classifier often performs well in practice, especially in situations where the independence assumption is not severely violated.

$$p(\text{"}I\ like\ the\ product\text{"}|P) = p(\text{"I"}|P) \times p(\text{"}like\text{"}|P) \times p(\text{"}the\text{"}|P) \times p(\text{"}product\text{"}|P)$$

$$p(\text{"}I\ like\ the\ product\text{"}|N) = p(\text{"I"}|N) \times p(\text{"}like\text{"}|N) \times p(\text{"}the\text{"}|N) \times p(\text{"}product\text{"}|N)$$

$$p(\text{"}I\ like\ the\ product\text{"}|P) = p(\text{"I"}|P) \times p(\text{"}like\text{"}|P) \times p(\text{"}the\text{"}|P) \times p(\text{"}product\text{"}|P)$$

$$p(\text{"}I\ like\ the\ product\text{"}|N) = p(\text{"I"}|N) \times p(\text{"}like\text{"}|N) \times p(\text{"}the\text{"}|N) \times p(\text{"}product\text{"}|N)$$

	Reviews	Sentiment Class
Training Data	Great product	P
	A good product	P
	I like it	P
	Difficult to use	N
	complicated product	N

	Word	P	N
1	Great	1	
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
	Total		

$$p(\text{"}I\ like\ the\ product\text{"}|P) = p(\text{"I"}|P) \times p(\text{"}like\text{"}|P) \times p(\text{"}the\text{"}|P) \times p(\text{"}product\text{"}|P)$$

$$p(\text{"}I\ like\ the\ product\text{"}|N) = p(\text{"I"}|N) \times p(\text{"}like\text{"}|N) \times p(\text{"}the\text{"}|N) \times p(\text{"}product\text{"}|N)$$

	Reviews	Sentiment Class
Training Data	Great product	P
	A good product	P
	I like it	P
	Difficult to use	N
	complicated product	N

	Word	P	N
1	Great	1	
2	product	1+1=2	1
3	a	1	
4	good	1	
5	I	1	
6	Like	1	
7	it	1	
8	Difficult		1
9	To		1
10	use		1
11	complicated		1
	Total	8	5

$$p(\text{"}I\ like\ the\ product\text{"}|P) = p(\text{"I"}|P) \times p(\text{"}like\text{"}|P) \times p(\text{"}the\text{"}|P) \times p(\text{"}product\text{"}|P)$$

$$p(\text{"}I\ like\ the\ product\text{"}|N) = p(\text{"I"}|N) \times p(\text{"}like\text{"}|N) \times p(\text{"}the\text{"}|N) \times p(\text{"}product\text{"}|N)$$

	Word	P	N
1	Great	1	
2	product	1+1=2	1
3	a	1	
4	good	1	
5	I	1	
6	Like	1	
7	it	1	
8	Difficult		1
9	To		1
10	use		1
11	complicated		1
	Total	8	5

- p(“I” | P) = $\frac{1}{8}$
- p(“I” | N) = $\frac{0}{5}$
- p(“like” | P) = $\frac{1}{8}$
- p(“like” | N) = $\frac{0}{5}$
- p(“product” | P) = $\frac{2}{8}$
- p(“product” | N) = $\frac{1}{5}$

	Word	P	N
1	Great	1	
2	product	1+1=2	1
3	a	1	
4	good	1	
5	I	1	
6	Like	1	
7	it	1	
8	Difficult		1
9	To		1
10	use		1
11	complicated		1
	Total	8	5

Laplace smoothing: $p(w_i|c) = \frac{N_{w_i} + k}{N_{words\ in\ c} \ + k|V|}$

k typically it is 1.

$|V|$ represents all unique words across classes.

- **p(“I” | P) =** $\frac{1+1}{8+1*11} = \frac{2}{19}$
- **p(“like” | P) =** $\frac{1+1}{8+1*11} = \frac{2}{19}$
- **p(“product” | P) =** $\frac{2+1}{8+1*11} = \frac{3}{19}$
- **p(“I” | N) =** $\frac{0+1}{5+1*11} = \frac{1}{16}$
- **p(“like” | N) =** $\frac{0+1}{5+1*11} = \frac{1}{16}$
- **p(“product” | N) =** $\frac{1+1}{5+1*11} = \frac{2}{16}$

$$p(\text{P}) \times \text{p(“I” | P)} \times \text{p(“like” | P)} \times \text{p(“product” | P)} = \frac{3}{5} \times \frac{2 \times 2 \times 3}{19^3} = 1.0 \times 10^{-3}$$

$$p(\text{N}) \times \text{p(“I” | N)} \times \text{p(“like” | N)} \times \text{p(“product” | N)} = \frac{2}{5} \times \frac{1 \times 1 \times 2}{16^3} = 0.2 \times 10^{-3}$$

$$
\begin{cases}
p(\text{P} \mid \text{"}\textit{I like the product}\text{"}) \propto \\
p(\text{P}) \times \text{p}(\text{“I”} \mid \text{P}) \times \text{p}(\text{“like”} \mid \text{P}) \times \text{p}(\text{“product”} \mid \text{P}) = \frac{3}{5} \times \frac{2 \times 2 \times 3}{19^3} = 1.0 \times 10^{-3} \\
\\
p(\text{N} \mid \text{"}\textit{I like the product}\text{"}) \propto \\
p(\text{N}) \times \text{p}(\text{“I”} \mid \text{N}) \times \text{p}(\text{“like”} \mid \text{N}) \times \text{p}(\text{“product”} \mid \text{N}) = \frac{2}{5} \times \frac{1 \times 1 \times 2}{16^3} = 0.2 \times 10^{-3}
\end{cases}
$$

Since $1.0 \times 10^{-3} > 0.2 \times 10^{-3}$,
we conclude that “*I like the product*” is likely positive.