Naive Bayes Classifier

07 July 2024 19:59

Based on Bayes Theorem (statistical concept)

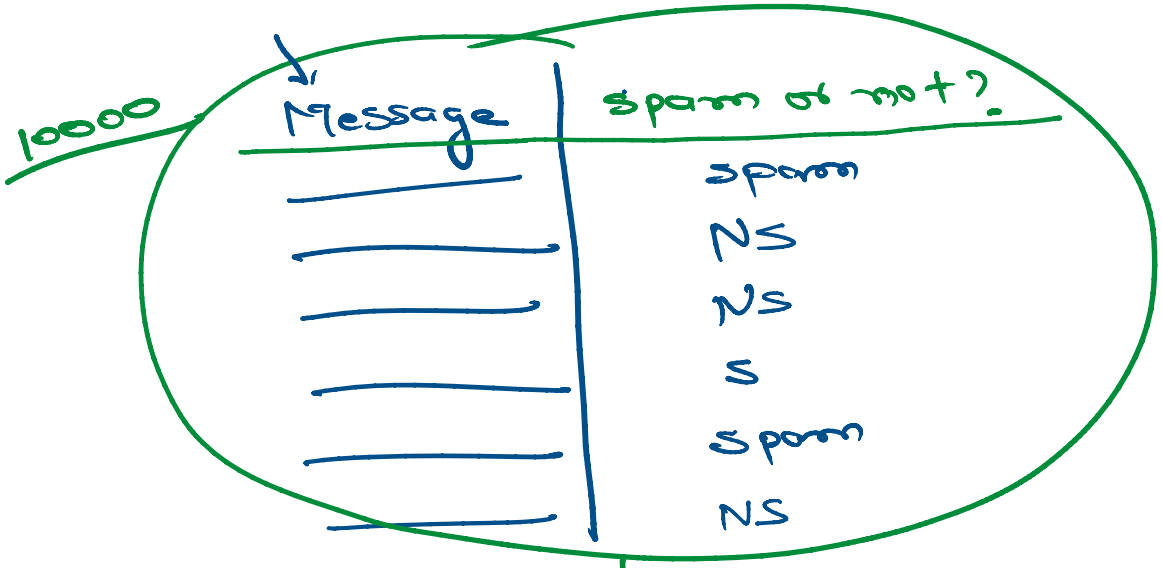
Naive Bayes

↓

Used specially for spam detection/classification

Genuine → spam → very dangerous

Spam → Genuine



Truecaller

Naive Bayes

↓

Learn the patterns in spam and Non-spam msgs.

↓

Future msgs and classify them on the basis of patterns we have learned.

Male 0

Male		0
F		1
M	→ Encoding →	0
F		1
F		1
M		0

Example →

Can we have a quick meeting tomorrow at 5 PM?

↓

Tokenization

(separate each word from the msg and make a list of words)

[can, we, have, a, quick, meeting, tomorrow, at, 5, PM]

↓

Remove stopwords (a, too, of, the, at etc)

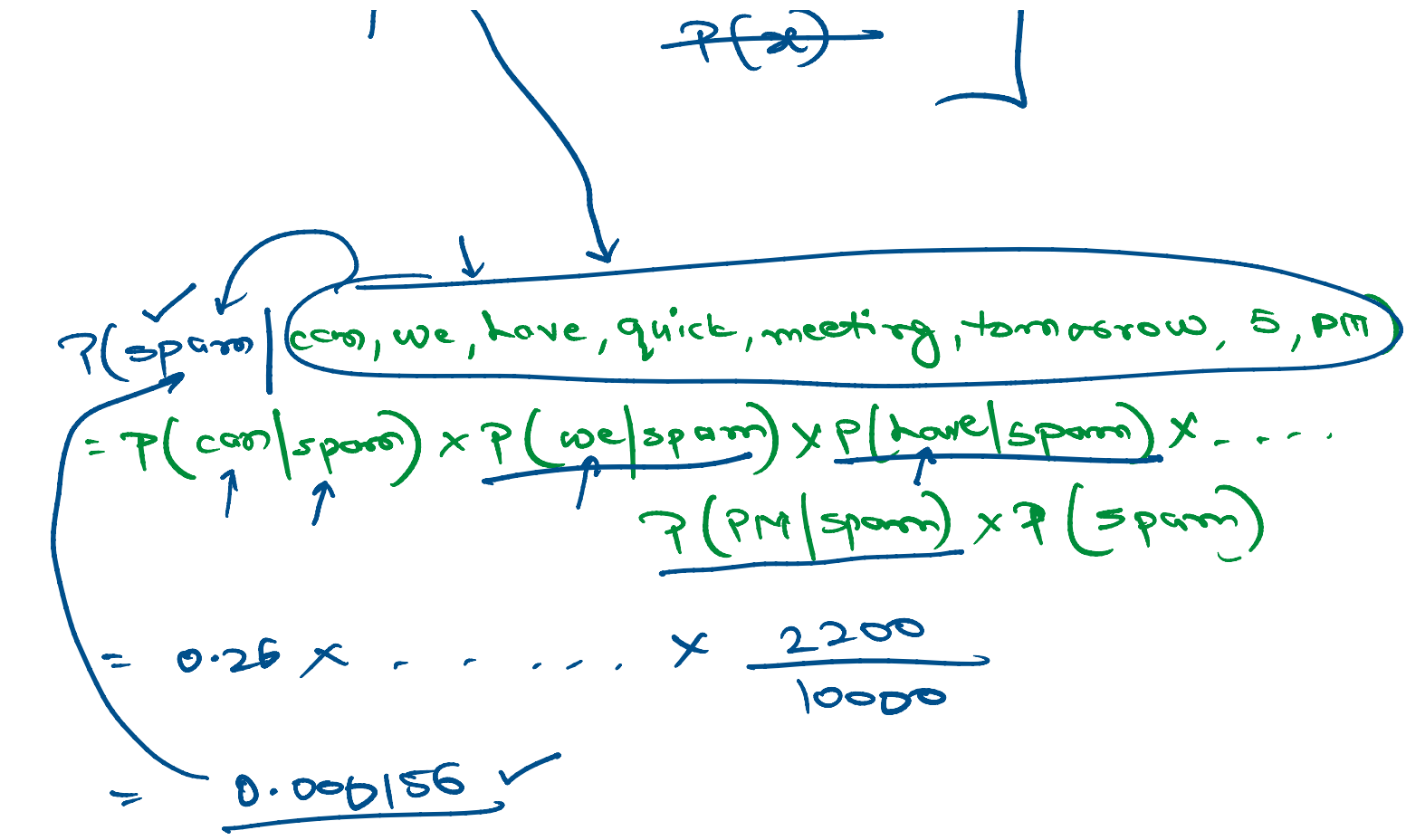
↓

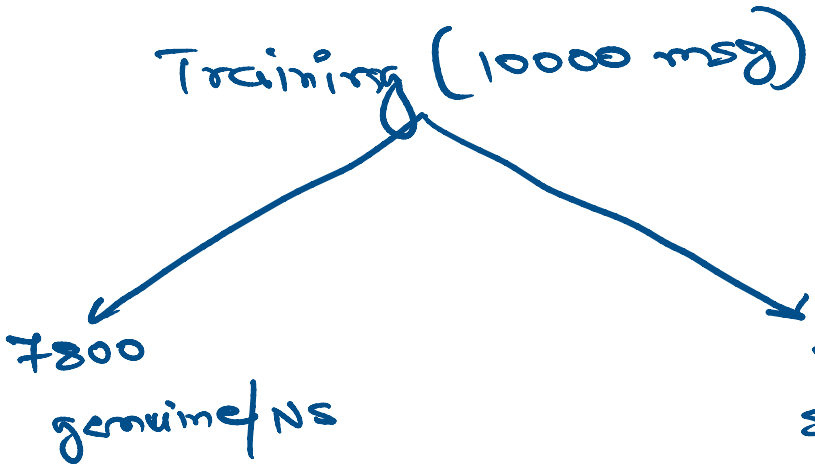
Vectorization

(converting text to numbers)

Naive Bayes →

$$P(c|x) = \frac{P(x|c) \times P(c)}{P(x)}$$





$P(NS) = \frac{7800}{10000}$

$= 0.78$

$= 78\%$.

$P(spam) = \frac{2200}{10000}$

$= 0.22$

$= 22\%$.

Let's take spam msgs in the training data.

$\frac{2200}{\text{Training data } (10,000)}$

2200 — training data 10,000

$$P(can|spam) = \frac{576}{2200}$$

$$= 0.26$$

$$P(we|spam) = \frac{180}{2200}$$

$$=$$

$$P(Non\,spam|can, we \ldots) = 1 - P(spam|\ldots)$$

$$= 1 - 0.18$$

$$= 0.82$$

$$= 82\%$$

Example →

Congratulation, you have a pre-approved loan of 50000 USD. No documentation required. Fill the form to claim the loan.

$$P(spam|congratulation, you, have, pre\text{-}approved, loan) =$$

$$= P(congratulation|spam)$$

our email

Regular


Regular