

Training Corpus
(annotated with
correct classes)

Shakespeare:

...not to be, that is the question.

Whether 'tis nobler in the mind...

Definition
of feature
vector

f1: Number of proper nouns

following comma.

f2: sentence contains two nouns.

Document to be
classified.

When in the course of
human events...

Feature vectors

(class, f1, f2),

(class, f1, f2),

(class, f1, f2),

...

Observed features

f1=3 seen twice

f1=2 seen 6 times

...

f2 seen 47 times

$P(\text{features}|\text{class})$

$P(f1=2|\text{bacon})=0.112$

...

$P(f2|\text{bacon})=0.787$

$P(f2|\text{shakespeare})=0.821$

Bayes' Theorem

$P(\text{class}|\text{observed features})$

$P(\text{shakespeare}|\text{observed})=0.3$

$P(\text{bacon}|\text{observed})=0.7$

Decision Rule

MAP

Answer

Bacon wrote it.