NAÏVE BAYES CLASSIFIER

Bayes Theorem

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

A Simple Example

- Consider a class on a certain subject
  - Female students in a class 300.
  - Male students in a class 200.
- 1% students have failed the course.
- We observe that
  - Among the students who failed, 50% are male and 50% are female.
- **Question:**
  - What is the probability that a student will fail given he is male ?

P(Female) = 300/500 = 0.6

P(Male) = 200/500 = 0.4

P(Fail) = 1% or 0.01

P(Male | Fail) = 50% or 0.5

P(Female | Fail) = 50% or 0.5

P(Fail | Male) = ?

$$= \frac{P(Male \mid Fail) * P(Fail)}{P(Male)}$$

$$= \frac{0.5 * 0.1}{0.4} = 0.0125$$

Naïve Bayes Classifier

- For a feature vector x, the Naïve Bayesian classifier will predict a class probability as follows.

$$P(C \mid x) = \frac{P(x \mid C) \times P(C)}{P(x)}$$

Naïve Bayes Classifier

- The probability $P(x|C)$ is simplified as,

$$P(x|C) = P(x_1, x_2, ..., x_n | C)$$

- It is assumed that the attributes are conditionally independent, which leads to,

$$P(x|C) = P(x_1, x_2, ..., x_n | C) = \prod_{i=1}^{n} P(x_i | C)$$

Email Classification Example

	Hello	Bob	Laptop	Win	Science	University	Student	Class
Email 1	0	0	0	1	1	1	0	Legitimate
Email 2	0	1	1	0	1	0	1	Legitimate
Email 3	1	0	0	0	0	1	1	Legitimate
Email 4	1	1	1	0	0	0	0	Spam
Email 5	1	1	1	1	1	0	0	Spam
New email	**1**	**0**	**1**	**1**	**0**	**0**	**0**	**?**

Naïve Bayesian Spam Filtering

- The probability for an email x being legitimate,

$$P(C \mid x) = P(C) \times \prod_{i=1}^{n} P(x_i \mid C) \Big/ P(x)$$

- The probability for an email x being spam,

$$P(C^c \mid x) = P(C^c) \times \prod_{i=1}^{n} P(x_i \mid C^c) \Big/ P(x)$$

- Two-way classification decisions
  - Legitimate: if $P(C \mid x) > P(C^c \mid x)$
  - Spam: Otherwise

Computing Class Probabilities

	Hello	Bob	Laptop	Win	Science	University	Student	Class
Email 1	0	0	0	1	1	1	0	Legitimate
Email 2	0	1	1	0	1	0	1	Legitimate
Email 3	1	0	0	0	0	1	1	Legitimate
Email 4	1	1	1	0	0	0	0	Spam
Email 5	1	1	1	1	1	0	0	Spam
New email	**1**	**0**	**1**	**1**	**0**	**0**	**0**	**?**

$$P(C) = \frac{|\,Legimate\ Emails\,|}{|\,Total\ Emails\,|} = 3/5 = 0.6$$

$$P(C^c) = \frac{|\,Spam\ Emails\,|}{|\,Total\ Emails\,|} = 2/5 = 0.4$$

Computing Likelihood

$$P(x \mid C) = \prod_{i=1}^{N} P(x_i \mid C)$$

$$= P(Hello \mid C) \times P(Laptop \mid C) \times P(Win \mid C)$$

$$= 1/3 \times 1/3 \times 1/3 = 0.037$$

$$P(x \mid C^c) = \prod_{i=1}^{N} P(x_i \mid C^c)$$

$$= P(Hello \mid C^c) \times P(Laptop \mid C^c) \times P(Win \mid C^c)$$

$$= 1 \times 1 \times 1/2 = 0.5$$

Classification Decision

$$P(Legimate \mid x) = P(C \mid x) = P(x \mid C) \times P(C) = \prod_{i=1}^{N} P(x_i \mid C) \times P(C)$$

$$= 0.037 \times 0.6 = 0.022$$

$$P(Spam \mid x) = P(C^c \mid x) = P(x \mid C^c) \times P(C^c) = \prod_{i=1}^{N} P(x_i \mid C^c) \times P(C^c)$$

$$= 0.5 \times 0.4 = 0.2$$

When attributes are Continuous

- When the attributes are continuous
  - We Consider underlying distributions
    - Such as Gaussain

$$P(x=v\,|\,C)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(v-\mu_c)}{2\sigma^2}}$$

$$P(C\,|\,x)=\prod_{i=1}^{N}P(x_i\,|\,C)\times P(C)$$