# Naïve Bayes Classifier

The Naïve Bayes Classifier is a simple probabilistic classifier based on applying Bayes' theorem. It has been successfully applied to many problems, such as:

1. Spam detection: Is an email message SPAM or HAM?
2. Optical character recognition: What letter or digit did a person write?
3. Medical diagnosis: Is a tumor malignant or benign?
4. Fraud detection: Does unusual credit card activity indicate fraud?
5. Lending: Will a borrower default on a loan?
6. Essay grading: What score should be assigned to an essay?
7. Customer service email routing: Which department should handle a request?

Bayes' theorem states:

$$P(y|x) = \frac{P(y)\, P(x|y)}{P(x)}.$$

Suppose that an email message contains the word "cash". We want to know if it is spam or ham.

$$P(spam|cash) = \frac{P(spam)\, P(cash|spam)}{P(cash)}$$

$$P(ham|cash) = \frac{P(ham)\, P(cash|ham)}{P(cash)}.$$

These equations can be expressed as

$$posterior = \frac{prior \times likelihood}{evidence}.$$

The prior is the base probability of spam, the likelihood is the fraction of spam messages that contain the word "cash", and the evidence is the fraction of all messages that contain the word "cash". The posterior is the updated probability of spam after weighing the evidence.

Given a new observation, we predict the class that maximizes the posterior probability. Notice that the denominator can be ignored, because it is the same for all classes.

Naïve Bayes assumes that the features are *independent*. This assumption allows us to multiply probabilities. For example, if an email contains the words "cash" and "lottery", then we might calculate as follows:

$$P(cash, lottery|spam) = P(cash|spam)\, P(lottery|spam).$$

## Example: Predicting Titanic Survivors

- 34% of the passengers on board the Titanic survived.
    - 32% of survivors were male; 68% were female.
    - 43% were in 1st class, 26.5% were in 2nd class, and 30.5% were in 3rd class.
- 66% of the passengers on board the Titanic died.
    - 82% of those who died were male; 18% were female.
    - 15% were in 1st class, 19% were in 2nd class, and 66% were in 3rd class.

A female in 1st class is predicted to survive:

- Survival: $0.34 \times 0.68 \times 0.43 = 0.099$
- Death: $0.66 \times 0.18 \times 0.15 = 0.018$

A male in 1st class is predicted to die:

- Survival: $0.34 \times 0.32 \times 0.43 = 0.047$
- Death: $0.66 \times 0.82 \times 0.15 = 0.082$

## Notes:

1. Naïve Bayes is resistant to over-fitting, and it is often useful when there is a large number of features (e.g. text processing).
2. If the predictor variables are continuous, then we use probability densities instead of probabilities. One often uses a Gaussian distribution or a kernel density estimate.
3. Multiplying many small numbers together is likely to produce an underflow. This can be avoided by adding the logarithms.
4. Bayesian networks can be used when there are conditional dependencies among the variables.