Axioms of Probability

$$0 \le P(A) \le 1$$

P(true) = 1   P(false) = 0.

$$P(A \vee B) = P(A) + P(B) - P(A \wedge B)$$

$$P(A \mid B) = \frac{P(A \wedge B)}{P(B)}$$

If A and B are independent:

$$P(A \wedge B) = P(A)P(B)$$

$$P(H \mid E) = \frac{P(E \mid H)P(H)}{P(E)}$$

Bayesian categorization:

$$P(c_i \mid E) = \frac{P(c_i)P(E \mid c_i)}{P(E)}$$

Naïve Bayes:

$$P(E \mid c_i) = P(e_1 \wedge e_2 \wedge \cdots \wedge e_m \mid c_i) = \prod_{j=1}^{m} P(e_j \mid c_i)$$

Laplace Smoothing:

$$P(e_j \mid c_i) = \frac{n_{ij} + mp}{n_i + m}$$

Single Link clustering:

$$sim(c_i, c_j) = \max_{x \in c_i, y \in c_j} sim(x, y)$$

$$sim((c_i \cup c_j), c_k) = \max(sim(c_i, c_k), sim(c_j, c_k))$$

Complete Link clustering:

$$sim(c_i, c_j) = \min_{x \in c_i, y \in c_j} sim(x, y)$$

$$sim((c_i \cup c_j), c_k) = \min(sim(c_i, c_k), sim(c_j, c_k))$$

Group average clustering:

$$sim(c_i, c_j) = \frac{1}{\left|c_i \cup c_j\right|(\left|c_i \cup c_j\right| - 1)} \sum_{\vec{x} \in (c_i \cup c_j)} \sum_{\vec{y} \in (c_i \cup c_j): \vec{y} \neq \vec{x}} sim(\vec{x}, \vec{y})$$

$$\vec{s}(c_j) = \sum_{\vec{x} \in c_j} \vec{x}$$

$$sim(c_i, c_j) = \frac{(\vec{s}(c_i) + \vec{s}(c_j)) \bullet (\vec{s}(c_i) + \vec{s}(c_j)) - (|c_i| + |c_i|)}{(|c_i| + |c_i|)(|c_i| + |c_i| - 1)}$$

Centroid:

$$\vec{\mu}(\mathrm{c}) = \frac{1}{|c|} \sum_{\vec{x} \in c} \vec{x}$$