Naive Bayes Classifier:

Bayesian classifier:

want to find $P(C|A_1, A_2 \ldots A_n)$

Can we estimate $P(C|A_1, A_2 \ldots A_n)$ directly from data.

$$P(C|A_1 A_2 \ldots A_n) = \frac{P(A_1 A_2 \ldots A_n|C)\,P(C)}{P(A_1, A_2 \ldots A_n)}$$

Assume Independence among attributes $A_1 \ldots A_n$.

$$P(A_1, A_2 \ldots A_n|C) = P(A_1|C)\,P(A_2|C) \ldots P(A_n|C)$$

New point is classified to C_j if $P(C_j)\prod P(A_i|C_j)$ is maximal

If one of the conditional probability is zero, then the entire expression is zero

Original: $P(A_i|C) = \frac{N_{ic}}{N_c}$

Laplace: $P(A_i|C) = \frac{N_{ic}+1}{N_c+c}$

m-estimate: $P(A_i|C) = \frac{N_{ic}+mp}{N_c+m}$

c: Number of classes

p: prior probability

m: parameter