Naïve Bayes Classifier

Bayesian classification:
For new example \( x \), assign label \( c_j \) that maximizes the following:

\[
\arg\max_{c_j} P(c_j) \prod_{i=1}^{d} P(x_j \mid c_j)
\]

Naïve Bayes assumption:

Notation:
- \( n_{ij} \) = number of examples with \( x_i = 1 \) and \( y = c_j \)
- \( N_j \) = number of examples of class \( c_j \)

Compute \( P(x_j \mid c_j) \) as follows:

\[
P(x \mid c_j) = \prod_{i=1}^{d} \left( \frac{n_{ij}}{N_j} \right)^{x_i} \left( 1 - \frac{n_{ij}}{N_j} \right)^{1-x_i}
\]

Laplace smoothing:

Add 1 to numerators and 2 to denominators

\[
P(x \mid c_j) = \prod_{i=1}^{d} \left( \frac{n_{ij} + 1}{N_j + 2} \right)^{x_i} \left( 1 - \frac{n_{ij} + 1}{N_j + 2} \right)^{1-x_i}
\]