Logistic Regression

Il-Chul Moon
Dept. of Industrial and Systems Engineering
KAIST
icmoon@kaist.ac.kr

Naïve Bayes vs. Logistic Regression

- Naïve Bayes classifier
  - $P(Y|X) = \frac{1}{1+\exp(-\frac{1}{2(\sigma_1^i)^2}\sum_{1\le i\le d}\left\{2(\mu_2^i-\mu_1^i)X_i+{\mu_2^i}^2-{\mu_2^i}^2\right\}+log\pi_2-log\pi_1)}$
  - Assumption to get this formula
    - Naïve Bayes assumption, Same variance assumption between classes
    - Gaussian distribution for P(X|Y)
    - Bernoulli distribution for P(Y)
  - # of parameters to estimate = 2*2*d+1=4d+1
    - With the different variances between classes
- Logistic Regression
  - $P(Y|X) = \frac{1}{1+e^{-\dot{\theta}^T x}}$
  - Assumption to get this formula
    - Fitting to the logistic function
  - # of parameters to estimate = d+1
- Who is the winner?
  - Really??? There is no winner... Why?

Together, modeling joint prob.

KAIST

Generative-Discriminative Pair

- Generative model, $P(Y|X)=P(X,Y)/P(X)=P(X|Y)P(Y)/P(X)$
  - Full probabilistic model of all variables
    - Estimate the parameters of $P(X|Y)$, $P(Y)$ from the data
  - Characteristics: Bayesian, Prior, Modeling the joint probability
  - Naïve Bayes Classifier
- Discriminative model, $P(Y|X)$
  - Do not need to model the distribution of the observed variables
    - Estimate the parameters of $P(Y|X)$ from the data
  - Characteristics: Modeling the conditional probability
  - Logistic Regression
- Pros and Cons [Ng & Jordan, 2002]
  - Logistic regression is less biased
  - Probably approximately correct learning: Naïve Bayes learns faster