- BAYES METHODS
- KERNEL ESTIMATOR

ESTIMASI MODEL REGRESI LOGISTIK NONPARAMETRIK ADITIF BERDASARKAN ESTIMATOR KERNEL

SKRIPSI

MARISA RIFADA

DEPARTEMEN MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS AIRLANGGA
SURABAYA
2009

Marisa Rifada, 2008. **Estimasi Model Regresi Logistik Nonparametrik Aditif Berdasarkan Estimator Kernel.** Skripsi dibawah bimbingan Nur Chamidah, S.Si, M.Si dan Ir. Elly Ana, M.Si., Departemen Matematika, Fakultas Sains dan Teknologi, Universitas Airlangga, Surabaya

ABSTRAK

Tujuan dari penulisan skripsi ini adalah untuk mengestimasi model regresi logistik nonparametrik aditif berdasarkan estimator Kernel. Secara umum bentuk model regresi logistik nonparametrik aditif adalah : $\log\left(\frac{\mu_i}{1-\mu_i}\right) = \sum_{j=1}^{p} m_j(x_{ji})$, dengan $m_j(.)$ adalah fungsi regresi yang diestimasi berdasarkan estimator Kernel. Dalam mengestimasi model regresi logistik nonparametrik aditif berdasarkan estimator Kernel digunakan algoritma *Local scoring*. Algoritma ini terdiri dari dua *loop* yaitu langkah *Scoring* (*outer loop)* yang diiterasikan sampai nilai rata-rata *deviance* konvergen dan langkah *Backfitting* terboboti *(inner loop)* yang diiterasikan sampai nilai rata-rata *Residual Sum of Square* (RSS) konvergen. Bentuk estimasi model regresi logistik nonparametrik aditif berdasarkan estimator Kernel dengan menggunakan algoritma *Local scoring* adalah sebagai berikut :

$$\log\left(\frac{\hat{\mu}_i}{1-\hat{\mu}_i}\right) = \sum_{j=1}^{p} \hat{m}_j(x_{ji})$$

$$= \sum_{j=1}^{p}\left([C^T W_{ji} C]^{-1} C^T W_{ji} \left\{ z - \sum_{k=1}^{j-1} \hat{m}_k^{(s+1)}(X_k) - \sum_{k=j+1}^{p} \hat{m}_k^{(s)}(X_k) \right\} \right)$$

dengan $W_{ji} = \text{diag}\left(K_h(x_{ji} - x_{j1}), K_h(x_{ji} - x_{j2}), \ldots, K_h(x_{ji} - x_{jn})\right)$

$K_h(.)$ adalah fungsi Kernel

Data yang digunakan untuk penerapan yaitu data kejadian penyakit Pneumonia pada balita di Rumah Sakit dr. Syaiful Anwar Malang tahun 2002 dengan jumlah sampel sebanyak 60 balita. Variabel responnya adalah terjadinya Pneumonia ($Y = 1$) atau tidak terjadinya Pneumonia ($Y = 0$). Sedangkan variabel prediktornya adalah usia balita (X_1) dan jumlah protein dalam darah balita (X_2). Dari hasil analisa data diperoleh bahwa semakin bertambahnya usia dan jumlah protein dalam darah balita maka peluang balita menderita Pneumonia akan semakin kecil. Seorang balita berpeluang besar menderita Pneumonia pada usia kurang dari 15 bulan dan jumlah protein dalam darah kurang dari 6 gr/dl.

Kata Kunci : *Regresi Logistik, Estimator Kernel, Algoritma Local Scoring, Pneumonia pada Balita*

Marisa Rifada, 2008. **The Estimation of Additive Nonparametric Logistic Regression Model based on Kernel Estimator**. This *Skripsi* is under advised by Nur Chamidah, S.Si, M.Si and Ir. Elly Ana, M.Si., Mathematics Department, Faculty of Sains and Technology, Airlangga University, Surabaya

ABSTRACT

The purpose of this *Skripsi* is to estimate additive nonparametric logistic regression model based on Kernel estimator. In general, the model of additive nonparametric logistic regression is : $\log\left(\frac{\mu_i}{1-\mu_i}\right)=\sum_{j=1}^{p} m_j(x_{ji})$, with $m_j(.)$ is a regression function which is estimated based on Kernel estimator. Local scoring algorithm is used in estimating additive nonparametric logistic regression model based on kernel estimator. This algorithm consists of two loops. First loop is Scoring step (outer loop) which is iterated until the deviance's average is convergence. The second loop is Backfitting weighted step (inner loop) which is iterated until the average of Residual Sum of Square (RSS) is convergence. The estimation of additive nonparametric logistic regression model based on kernel estimator with Local scoring algorithm is :

$$\log\left(\frac{\hat{\mu}_i}{1-\hat{\mu}_i}\right)=\sum_{j=1}^{p}\hat{m}_j(x_{ji})$$

$$=\sum_{j=1}^{p}\left([C^T W_{ji} C]^{-1} C^T W_{ji}\left\{z-\sum_{k=1}^{j-1}\hat{m}_k^{(s+1)}(X_k)-\sum_{k=j+1}^{p}\hat{m}_k^{(s)}(X_k)\right\}\right)$$

with $W_{ji}=\text{diag}\left(K_h(x_{ji}-x_{j1}), K_h(x_{ji}-x_{j2}),\ldots, K_h(x_{ji}-x_{jn})\right)$

$K_h(.)$ is Kernel function

The data which are used for applying are the data of Pneumonia desease identified to baby under five years old at dr. Syaiful Anwar hospital of Malang in 2002 with 60 baby under five years old as the sample. While the response variable is the suffer as well as identification of Pneumonia (Y = 1) or unidentification of Pneumonia (Y = 0). Whereas the predictor variables are the age (X_1) and the amount of protein contained in blood of baby under five years old (X_2). Based on the data analysis obtained that the more age on baby under five years old and the amount of protein contained in blood the less the probability of a baby under five years old to suffer Pneumonia. A baby under five years old gets more probability to suffer Pneumonia on the age of less than 15 months and the amount of protein contained in blood is less than 6 gr/dl.

Key Words : *Logistic Regression, Kernel Estimator, Local Scoring Algorithm, Pneumonia on Baby Under Five Years Old*