

REGRESSION ANALYSIS  
NONPARAMETRIC STATISTICS

kk.  
MPM 22/04  
Har  
e

# ESTIMASI KURVA REGRESI NONPARAMETRIK DENGAN METODE k-NN

## SKRIPSI



**BAMBANG HARIYANTO**

**JURUSAN MATEMATIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS AIRLANGGA  
SURABAYA**

**2004**

**ESTIMASI KURVA REGRESI NONPARAMETRIK DENGAN  
METODE k-NN**

**SKRIPSI**

**Sebagai Salah Satu Syarat Untuk Memperoleh  
Gelar Sarjana Sains Bidang Matematika Pada Fakultas Matematika dan  
Ilmu Pengetahuan Alam Universitas Airlangga**

Oleh :

**BAMBANG HARIYANTO**  
NIM. 089811832



Tanggal Lulus : 5 Februari 2004

Disetujui Oleh :

Pembimbing I,

Handwritten signature of the first supervisor, Drs. Suliyanto.

**Drs. Suliyanto, M.Si**  
NIP. 131933016

Pembimbing II,

Handwritten signature of the second supervisor, Nur Chamidah.

**Nur Chamidah, S.Si, M.Si**  
NIP. 132205653

## LEMBAR PENGESAHAN SKRIPSI

**Judul** : Estimasi Kurva Regresi Nonparametrik dengan Metode k-NN  
**Penyusun** : Bambang Hariyanto  
**NIM** : 089811832  
**Tanggal Ujian** : 5 Februari 2004

Disetujui Oleh :

**Pembimbing I,**



**Drs. Suliyanto, M.Si**  
**NIP. 131933016**

**Pembimbing II,**



**Nur Chamidah, S.Si, M.Si**  
**NIP. 132205653**

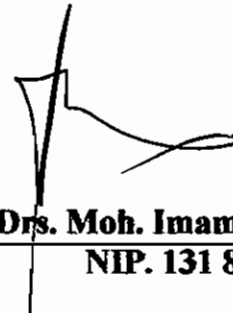
Mengetahui :

**Dekan Fakultas MIPA  
Universitas Airlangga**



**Drs. H. Abdul Latief Burhan, MS**  
**NIP. 131 286 709**

**Ketua Jurusan Matematika  
FMIPA Universitas Airlangga**



**Drs. Moh. Imam Utoyo, M.Si**  
**NIP. 131 801 397**

Bambang H, 2004. *Estimasi Kurva Regresi Nonparametrik dengan Metode k-NN*. Skripsi ini dibawah bimbingan Nur Chamidah S.Si,M.Si dan Drs. Suliyanti S.Si, Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam.

---

## ABSTRAK

Analisis regresi adalah salah satu alat di dalam statistika yang umumnya digunakan untuk menentukan hubungan antara dua variabel yaitu variabel respon Y dan variabel prediktor X untuk n pengamatan. Hubungan variabel tersebut dapat dinyatakan sebagai berikut :

$$y_i = m(x_i) + \varepsilon_i \quad i = 1, 2, \dots, n$$

dengan  $m(x)$  sebagai kurva regresi dan  $\varepsilon_i$  adalah error dengan mean 0 dan varians  $\sigma^2$ .

Tujuan dari penulisan ini adalah untuk mengestimasi kurva regresi  $m(x)$  dengan metode k-NN (*k- Nearest Neighborhood*). Secara umum susunan dari regresi penghalusan k-NN adalah sebagai berikut :

$$\hat{m}_k(x) = n^{-1} \sum_{i=1}^n W_{ki}(x) y_i$$

dimana  $\{W_{ki}(x)\}_{i=1}^n$  adalah barisan pembobot yang didefinisikan sebagai himpunan indek

$J_x = \{ i : x_i \text{ adalah salah satu dari } k \text{ persekitaran yang terdekat dengan } x \}$  dengan himpunan indek dari persekitaran pengamatan inilah barisan pembobot k-NN dibentuk yaitu :

$$W_{ki}(x) = \begin{cases} n/k & , \text{ jika } i \in J_x \\ 0 & , \text{ untuk yang lain} \end{cases}$$

dengan k adalah parameter penghalus yang mengatur kemulusan dari kurva yang diduga.

Dari penerapan model regresi nonparametrik pada data pertumbuhan anak yang diukur dari berat badan anak sebagai variabel respon dan umur anak sebagai variabel prediktor, maka diperoleh nilai estimatornya seperti dalam lampiran dengan k optimal berdasarkan kriteria GCV adalah 11.

*Kata Kunci* : Regresi Nonparametrik, Fungsi Kernel, Metode k-NN, GCV