

**KLASIFIKASI KELOMPOK RUMAH TANGGA  
DI KABUPATEN BLORA MENGGUNAKAN  
*MULTIVARIATE ADAPTIVE REGRESSION SPLINE (MARS) DAN  
FUZZY K-NEAREST NEIGHBOR (FK-NN)***



---

---

**SKRIPSI**

---

---

**Oleh :**

**YANI PUSPITA KRISTIANI**

**24010211120018**

**JURUSAN STATISTIKA  
FAKULTAS SAINS DAN MATEMATIKA  
UNIVERSITAS DIPONEGORO  
SEMARANG**

**2015**

**KLASIFIKASI KELOMPOK RUMAH TANGGA  
DI KABUPATEN BLORA MENGGUNAKAN  
*MULTIVARIATE ADAPTIVE REGRESSION SPLINE (MARS)*  
DAN *FUZZY K-NEAREST NEIGHBOR (FK-NN)***

**Disusun Oleh :**

**YANI PUSPITA KRISTIANI**

**24010211120018**

Diajukan Sebagai Salah Satu Syarat Memperoleh Gelar Sarjana Sains pada  
Jurusan Statistika Fakultas Sains dan Matematika Universitas Diponegoro

**JURUSAN STATISTIKA  
FAKULTAS SAINS DAN MATEMATIKA  
UNIVERSITAS DIPONEGORO**

**SEMARANG**

**2015**

## HALAMAN PENGESAHAN I

Judul Skripsi : Klasifikasi Kelompok Rumah Tangga di Kabupaten Blora  
Menggunakan *Multivariate Adaptive Regression Spline*  
(MARS) dan *Fuzzy K-Nearest Neighbor* (FK-NN)

Nama Mahasiswa : Yani Puspita Kristiani

NIM : 24010211120018

Telah diujikan pada sidang Tugas Akhir tanggal 14 September 2015 dan dinyatakan lulus pada tanggal 29 September 2015

Semarang, September 2015

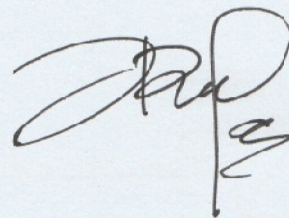
Mengetahui,  
Ketua Jurusan Statistika  
Fakultas Sains dan Matematika  
Universitas Diponegoro



Dia. Dwi Ispriyanti, M.Si

NIP.195709141986032001

Panitia Penguji Tugas Akhir  
Ketua,



Triastuti Wuryandari, S.Si, M.Si

NIP.197109061998032001

## HALAMAN PENGESAHAN II

Judul Skripsi : Klasifikasi Kelompok Rumah Tangga di Kabupaten Blora  
Menggunakan *Multivariate Adaptive Regression Spline*  
(MARS) dan *Fuzzy K-Nearest Neighbor* (FK-NN)

Nama Mahasiswa : Yani Puspita Kristiani

NIM : 24010211120018

Telah diujikan pada sidang Tugas Akhir tanggal 14 September 2015

Semarang, September 2015

Pembimbing I



Diah Safitri, S.Si, M.Si

NIP.197510082003122001

Pembimbing II



Dra. Dwi Ispriyanti, M.Si

NIP.195709141986032001

## KATA PENGANTAR

Puji syukur kepada Allah SWT yang telah melimpahkan karunia-Nya sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul **“Klasifikasi Kelompok Rumah Tangga di Kabupaten Blora Menggunakan *Multivariate Adaptive Regression Spline* (MARS) dan *Fuzzy K-Nearest Neighbor* (FK-NN)”**.

Tugas Akhir ini tidak akan terselesaikan dengan baik tanpa adanya dukungan dan bantuan dari berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih kepada:

1. Ibu Dra. Dwi Ispriyanti, M.Si selaku Ketua Jurusan Statistika Fakultas Sains dan Matematika Universitas Diponegoro dan dosen pembimbing II.
2. Ibu Diah Safitri, S.Si., M.Si selaku dosen pembimbing I.
3. Dosen-dosen Jurusan Statistika Fakultas Sains dan Matematika Universitas Diponegoro.
4. Semua pihak yang telah membantu kelancaran penyusunan Tugas Akhir ini, yang tidak dapat penulis sebutkan satu per satu.

Penulis menyadari bahwa dalam penyusunan Tugas Akhir ini masih jauh dari sempurna. Oleh karena itu, penulis mengharapkan kritik dan saran demi kesempurnaan penulisan selanjutnya.

Semarang, September 2015

Penulis

## ABSTRAK

Metode klasifikasi yang baik akan menghasilkan sedikit kesalahan klasifikasi. Metode klasifikasi sangat berkembang dan dua di antara metode klasifikasi yang telah ada yaitu metode *Multivariate Adaptive Regression Spline* (MARS) dan metode *Fuzzy K-Nearest Neighbor* (FK-NN). Penelitian ini bertujuan untuk membandingkan pengklasifikasian kelompok rumah tangga miskin dan kelompok rumah tangga tidak miskin berdasarkan pengeluaran perkapita yang sudah dikonversikan berdasarkan garis kemiskinan antara metode MARS dan FK-NN. Penelitian ini menggunakan data sekunder berupa hasil Survei Sosial Ekonomi Nasional (SUSENAS) di Kabupaen Blora tahun 2014. Evaluasi kesalahan hasil klasifikasi menggunakan APER. Hasil klasifikasi terbaik menggunakan metode MARS adalah jika menggunakan kombinasi BF=76, MI=3, MO=1 karena menghasilkan nilai *Generalized Cross Validation* (GCV) terkecil dan diperoleh nilai APER sebesar 10,119 %. Hasil klasifikasi terbaik metode FK-NN adalah jika menggunakan K=9 karena menghasilkan laju *error* terkecil dan diperoleh nilai APER sebesar 9,523%. Berdasarkan nilai APER menunjukkan bahwa pengklasifikasian kelompok rumah tangga di Kabupaten Blora menggunakan metode FK-NN lebih baik dibandingkan dengan metode MARS.

**Kata kunci:** Klasifikasi, MARS, FK-NN, APER, SUSENAS, Blora

## ABSTRACT

Good classification method will result on less classification error. Classification method developed rapidly. Two of the existing classification methods are Multivariate Adaptive Regression Spline (MARS) and Fuzzy K-Nearest Neighbor (FK-NN). This research aims to compare the classification of poor household and prosperous household based on per capita income which has been converted according to the poverty line between MARS and FK-NN method. This research used secondary data in the form of result of National Economy and Social Survey (SUSENAS) in Blora subdistrict in 2014. The result of the classification was evaluated using APER. The best classification result using MARS method is by using the combination of BF= 76, MI= 3, MO= 1 because it will result on the smallest Generalized Cross Validation (GCV) and the APER is 10,119%. The best classification result using FK-NN method is by using K=9 because it will result on the smallest error and the APER is 9,523%. The APER calculation shows that the classification of household in Blora subdistrict using FK-NN method is better than using MARS method.

**Keywords:** Classification, MARS, FK-NN, APER, SUSENAS, Blora

## DAFTAR ISI

|   | Halaman |
|---|---------|
| HALAMAN JUDUL .....   | i       |
| LEMBAR PENGESAHAN I.....  | ii      |
| LEMBAR PENGESAHAN II.....                                       | iii     |
| KATA PENGANTAR .....  | iv      |
| ABSTRAK .....   | v       |
| <i>ABSTRACT</i> .....   | vi      |
| DAFTAR ISI .....  | vii     |
| DAFTAR TABEL.....   | ix      |
| DAFTAR LAMPIRAN .....   | x       |
| BAB I PENDAHULUAN   |         |
| 1.1 Latar Belakang.....   | 1       |
| 1.2 Rumusan Masalah.....  | 4       |
| 1.3 Batasan Masalah .....                                       | 5       |
| 1.4 Tujuan Penelitian .....                                     | 5       |
| BAB II TINJAUAN PUSTAKA   |         |
| 2.1 Kemiskinan .....  | 6       |
| 2.2 <i>Multivariate Adaptive Regression Spline (MARS)</i> ..... | 9       |
| 2.2.1 MARS berkembang dari RPR.....                             | 9       |
| 2.2.2 Model MARS.....   | 10      |
| 2.2.3 Estimasi Parameter.....                                   | 14      |
| 2.3 Klasifikasi pada MARS.....                                  | 17      |
| 2.4 <i>Fuzzy</i> .....  | 18      |

|                                |   |    |
|--------------------------------|---|----|
| 2.5                            | <i>Fuzzy K-Nearest Neighbor</i> .....                             | 19 |
| 2.6                            | Evaluasi Kesalahan Hasil Klasifikasi .....                        | 21 |
| BAB III METODOLOGI PENELITIAN  |   |    |
| 3.1                            | Sumber Data .....   | 22 |
| 3.2                            | Variabel Penelitian .....   | 22 |
| 3.3                            | Metode Analisis Data .....  | 26 |
| BAB IV ANALISIS DAN PEMBAHASAN |   |    |
| 4.1                            | Metode <i>Multivariate Adaptive Regression Spline</i> (MARS)..... | 30 |
| 4.1.1                          | Pemodelan MARS.....   | 30 |
| 4.1.2                          | Pendugaan Klasifikasi Menggunakan Metode MARS...40                |    |
| 4.1.3                          | Hasil Klasifikasi Menggunakan Metode MARS .....                   | 44 |
| 4.2                            | Metode <i>Fuzzy K-Nearest Neighbor</i> (FK-NN) .....              | 45 |
| 4.2.1                          | Pendugaan Klasifikasi Menggunakan Metode FK-NN..45                |    |
| 4.2.2                          | Hasil Klasifikasi Menggunakan Metode FK-NN.....                   | 49 |
| 4.3                            | Evaluasi Kesalahan Hasil Klasifikasi .....                        | 50 |
| BAB V KESIMPULAN .....         |   | 52 |
| DAFTAR PUSTAKA .....           |   | 54 |
| LAMPIRAN.....                  |   | 56 |

## DAFTAR TABEL

|  | Halaman |
|--|---------|
| Tabel 1 Formula Jarak Dua Data dengan Satu Atribut .....                 | 20      |
| Tabel 2 Matriks Konfusi untuk Klasifikasi Dua Kelas .....                | 21      |
| Tabel 3 Variabel Penelitian.....   | 22      |
| Tabel 4 Penentuan Model Terbaik .....                                    | 30      |
| Tabel 5 Hasil Klasifikasi dengan Metode MARS .....                       | 44      |
| Tabel 6 Hasil Laju <i>Error</i> untuk Percobaan Nilai K pada FK-NN ..... | 46      |
| Tabel 7 Hasil Klasifikasi dengan Metode FK-NN .....                      | 49      |
| Tabel 8 Perhitungan APER untuk Metode MARS .....                         | 50      |
| Tabel 9 Perhitungan APER untuk Metode FK-NN.....                         | 50      |

## DAFTAR LAMPIRAN

|            | Halaman   |
|------------|---|
| Lampiran 1 | Data Kelompok Rumah Tangga serta Faktor-Faktor yang Menjadi Penciri Kelompok Rumah Tangga dari 562 Rumah Tangga di Kabupaten Blora ..... 56                                   |
| Lampiran 2 | <i>Output</i> dari Metode <i>Multivariate Adaptive Regression Spline</i> (MARS) Menggunakan <i>Software</i> MARS 2.0 ..... 57   |
| Lampiran 3 | Nilai GCV untuk Seluruh Pengombinasian BF, MI dan MO ..... 72   |
| Lampiran 4 | <i>Syntax</i> Metode <i>Fuzzy K-Nearest Neighbor</i> Menggunakan <i>Software</i> Matlab R2009a ..... 73   |
| Lampiran 5 | <i>Output</i> Hasil Klasifikasi Menggunakan Metode <i>Fuzzy K-Nearest Neighbor</i> Menggunakan <i>Software</i> Matlab R2009a dengan K=9 ..... 78                              |
| Lampiran 6 | Hasil Perhitungan Jarak <i>Euclidean</i> dengan Menggunakan Data <i>Testing</i> adalah Data ke-398 dan telah Dilakukan Pengurutan dari yang Terkecil hingga Terbesar ..... 84 |
| Lampiran 7 | Hasil Perhitungan Jarak <i>Euclidean</i> dengan Menggunakan Data <i>Testing</i> adalah Data ke-413 dan telah Dilakukan Pengurutan dari yang Terkecil hingga Terbesar ..... 85 |
| Lampiran 8 | Hasil Perhitungan Data <i>Testing</i> Menggunakan Metode <i>Multivariate Adaptive Regression Spline</i> (MARS) ..... 86   |

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Kemiskinan merupakan masalah yang dihadapi oleh seluruh negara, terutama di negara berkembang seperti Indonesia. Menurut Haughton dan Khandker (2012), kemiskinan timbul<sup>1</sup> apabila masyarakat tidak memiliki kemampuan-kemampuan utama, pendapatan, pendidikan yang memadai, kondisi kesehatan yang baik, rasa aman, kepercayaan diri yang tinggi dan hak seperti kebebasan berbicara. Dengan demikian, kemiskinan merupakan sebuah fenomena multidimensional dan tidak dapat diatasi dengan solusi yang sederhana.

Menurut Badan Pusat Statistik (2014a), penduduk miskin di Indonesia menunjukkan tren menurun selama periode 2010-2014. Tahun 2010, jumlah penduduk miskin sebesar 31,02 juta jiwa. Pemerintah terus berupaya menekan angka kemiskinan, dan tahun 2014 jumlah penduduk miskin sudah berkurang menjadi 28,28 juta jiwa. Jika ditinjau berdasarkan daerah tempat tinggal, penduduk miskin lebih banyak tinggal di daerah pedesaan dibanding daerah perkotaan. Tahun 2010, jumlah penduduk miskin yang tinggal di daerah pedesaan sebesar 19,93 juta jiwa dan di tahun 2014 berkurang menjadi 17,77 juta jiwa. Sementara itu, jumlah penduduk miskin di daerah perkotaan tahun 2010 sebesar 11,10 juta jiwa dan di tahun 2014 menurun menjadi 10,51 juta jiwa.

Menurut Badan Pusat Statistik (2014b), Provinsi Jawa Tengah memiliki jumlah penduduk miskin terbanyak di Indonesia. Di Jawa Tengah sendiri kemiskinan di pedesaan lebih banyak daripada kemiskinan di perkotaan. Sekitar 57,16 persen penduduk miskin di Jawa Tengah berada di pedesaan. Wilayah Blora

seluas 1.820,59 km<sup>2</sup> atau 5,5 persen dari luas propinsi menjadikan Kabupaten Blora sebagai kabupaten terluas nomor tiga di Jawa Tengah setelah Kabupaten Cilacap dan Grobogan.

Menurut Badan Pusat Statistik (2014a), kelompok rumah tangga berdasarkan pengeluaran per kapita yang sudah dikonversikan berdasarkan garis kemiskinan terdiri dari dua kelompok yaitu kelompok rumah tangga miskin dan kelompok rumah tangga tidak miskin. Garis Kemiskinan digunakan sebagai batas untuk mengelompokkan rumah tangga miskin dan rumah tangga tidak miskin. Rumah tangga miskin adalah rumah tangga yang memiliki rata-rata pengeluaran perkapita per bulan di bawah Garis Kemiskinan (GK).

Melihat masih tingginya tingkat kemiskinan maka perlu adanya suatu cara untuk mengidentifikasi kelompok rumah tangga. Salah satunya yaitu dengan melihat faktor-faktor yang dapat menjadi ciri dari kelompok rumah tangga tersebut. Berdasarkan kuisisioner Survei Sosial Ekonomi Nasional (SUSENAS) tahun 2014 yang diperoleh dari Badan Pusat Statistik, faktor-faktor yang diduga mempengaruhi kelompok rumah tangga adalah jenis kelamin kepala rumah tangga, umur kepala rumah tangga, jumlah anggota rumah tangga, ijazah tertinggi kepala rumah tangga, kegiatan utama kepala rumah tangga, lapangan usaha utama kepala rumah tangga, status pekerjaan utama kepala rumah tangga, status penguasaan tempat tinggal, jenis atap tempat tinggal, jenis dinding rumah tinggal, jenis lantai rumah tinggal, penggunaan fasilitas tempat buang air besar, tempat pembuangan akhir tinja, sumber air minum, bahan bakar energi utama untuk memasak, pengalaman membeli beras raskin, ada anggota rumah tangga yang memiliki jaminan pembiayaan/asuransi kesehatan, ada anggota rumah tangga

yang dapat menggunakan telepon selular dan ada anggota rumah tangga yang menguasai penggunaan komputer desktop.

Untuk menjelaskan pola hubungan antara kelompok rumah tangga dengan faktor-faktor yang diduga mempengaruhi kelompok rumah tangga dapat digunakan pendekatan kurva regresi. Menurut Eubank (1999) pendekatan dalam menentukan kurva regresi dapat menggunakan dua cara yaitu pendekatan parametrik dan pendekatan nonparametrik. Apabila kurva regresi membentuk suatu pola hubungan tertentu seperti linier, kuadratik dan kubik serta memenuhi asumsi-asumsi klasik seperti normalitas, kesamaan varian, independensi ataupun nonautokorelasi maka dapat menggunakan pendekatan parametrik. Namun apabila asumsi dari pendekatan parametrik tidak terpenuhi maka pendugaan dapat dilakukan dengan pendekatan nonparametrik. Regresi nonparametrik memiliki fleksibilitas yang tinggi dalam mengestimasi kurva regresi. Dalam pandangan regresi nonparametrik, data diharapkan mencari sendiri estimasi kurva regresi, tanpa dipengaruhi oleh faktor subyektifitas dari perancang penelitian.

Salah satu metode regresi nonparametrik adalah *Multivariate Adaptive Regression Spline* (MARS). Menurut Friedman (1991) metode MARS merupakan metode regresi nonparametrik multivariat dengan data dimensional tinggi, yaitu data yang memiliki jumlah variabel prediktor sebesar  $3 \leq v \leq 20$  dan sampel data yang berukuran  $50 \leq N \leq 1000$ .

*Fuzzy K-Nearest Neighbor* adalah pengembangan dari teori *K-Nearest Neighbor* yang digabungkan dengan teori *fuzzy* dalam menyampaikan pemberian label kelas pada data uji yang diprediksi. Metode *Fuzzy K-Nearest Neighbor* dapat

melakukan prediksi secara tegas pada uji berdasarkan tetangga terdekat. Selain itu juga memprediksi kelas dengan memberikan nilai keanggotaan (Prasetyo, 2012).

Berdasarkan penjelasan di atas, dilakukan perbandingan metode MARS dan metode FK-NN dalam pengklasifikasian kelompok rumah tangga di Kabupaten Blora. Metode yang menghasilkan nilai *Apparent Error Rate* (APER) terkecil merupakan metode terbaik dalam pengklasifikasian kelompok rumah tangga.

## **1.2 Rumusan Masalah**

Berdasarkan uraian latar belakang tersebut, permasalahan yang akan dibahas pada penelitian ini adalah sebagai berikut:

1. Bagaimana hasil klasifikasi dan kesalahan hasil klasifikasi kelompok rumah tangga di Kabupaten Blora berdasarkan faktor-faktor yang diduga mempengaruhi kelompok rumah tangga dengan menggunakan metode MARS.
2. Bagaimana hasil klasifikasi dan kesalahan hasil klasifikasi kelompok rumah tangga di Kabupaten Blora berdasarkan faktor-faktor yang diduga mempengaruhi kelompok rumah tangga dengan menggunakan metode FK-NN.
3. Bagaimana perbandingan kesalahan hasil klasifikasi kelompok rumah tangga di Kabupaten Blora dengan menggunakan metode MARS dan FK-NN.

### 1.3 Batasan Masalah

Data yang digunakan pada penelitian ini adalah data hasil Survei Sosial dan Ekonomi (SUSENAS) tahun 2014 di Kabupaten Blora. Pengolahannya menggunakan dua metode, yaitu metode *Multivariate Adaptive Regression Spline* (MARS) dan metode *Fuzzy K-Nearest Neighbor*, dimana masing-masing metode hanya mencari kesalahan hasil klasifikasi.

### 1.4 Tujuan Penelitian

Tujuan yang ingin dicapai dari penelitian ini adalah sebagai berikut:

1. Mengklasifikasi dan menghitung kesalahan hasil klasifikasi kelompok rumah tangga di Kabupaten Blora berdasarkan faktor-faktor yang diduga mempengaruhi kelompok rumah tangga dengan menggunakan metode MARS.
2. Mengklasifikasi dan menghitung kesalahan hasil klasifikasi kelompok rumah tangga di Kabupaten Blora berdasarkan faktor-faktor yang diduga mempengaruhi kelompok rumah tangga dengan menggunakan metode FK-NN.
3. Membandingkan kesalahan hasil klasifikasi kelompok rumah tangga di Kabupaten Blora dengan menggunakan metode MARS dan FK-NN.