

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

Yani Puspita Kristiani¹, Diah Safitri², Dwi Ispriyanti³

¹Mahasiswa Jurusan Statistika FSM Universitas Diponegoro

^{2,3}Staf Pengajar Jurusan Statistika FSM Universitas Diponegoro

yanipus7@gmail.com, diahsafitri@gmail.com, ispriyanti.dwi@gmail.com

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

1. PENDAHULUAN

Kemiskinan merupakan masalah yang dihadapi oleh seluruh negara, terutama di negara berkembang seperti Indonesia. Menurut Haughton dan Khandker (2012), kemiskinan timbul 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. Menurut Badan Pusat Statistik (2014b), Provinsi Jawa Tengah memiliki jumlah penduduk miskin terbanyak di Indonesia. Kabupaten Blora tetap saja menjadi kabupaten miskin, karena kekuatan ekonominya lebih banyak ditopang oleh sektor pertanian seluas 40 persen wilayah (72.502.33 ha).

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.

Berdasarkan kuisioner 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 nonparametrik salah satunya adalah *Multivariate Adaptive Regression Spline* (MARS). Menurut Friedman (1991) metode MARS merupakan metode regresi nonparametrik multivariat dengan data dimensional tinggi.

Menurut Prasetyo (2012) *Fuzzy K-Nearest Neighbor* (FK-NN) adalah pengembangan dari teori *K-Nearest Neighbor* yang digabungkan dengan teori *fuzzy* dalam menyampaikan pemberian label kelas pada data uji yang diprediksi.

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.

2. TINJAUAN PUSTAKA

2.1 Kemiskinan

Menurut Dermoredjo dan Simatupang (2003), kemiskinan (*poverty*) merupakan masalah yang dihadapi oleh seluruh negara, terutama di negara berkembang seperti Indonesia. Hal ini dikarenakan kemiskinan itu bersifat multidimensional.

2.2 *Multivariate Adaptive Regression Spline* (MARS)

Menurut Friedman (1991) metode *Multivariate Adaptive Regression Spline* (MARS) merupakan pengembangan dari metode *Recursive Partitioning Regression* (RPR) dan fungsi *spline*. Metode MARS berguna untuk mengatasi permasalahan data berdimensi tinggi, yaitu data yang memiliki jumlah variabel prediktor sebanyak $3 \leq v \leq 20$, serta memiliki ukuran sampel sebesar $50 \leq N \leq 1000$. Model dalam metode MARS adalah sebagai berikut:

$$\hat{Y} = a_0 + \sum_{m=1}^M a_m \prod_{k=1}^{K_m} [S_{km} (X_{V(k,m)} - t_{km})]_+$$

dengan:

a_0 = fungsi basis induk (konstanta)

a_m = koefisien dari fungsi basis ke- m

M = banyaknya maksimum fungsi basis

K_m = derajat interaksi pada fungsi basis ke- m

S_{km} = bernilai +1 jika data terletak di kanan titik *knot*, dan bernilai -1 jika data terletak di kiri titik *knot*

$X_{V(k,m)}$ = variabel prediktor ke- v

t_{km} = nilai *knot* dari variabel prediktor $X_{V(k,m)}$

Menurut Pintowati dan Otok (2012) penentuan *knot* pada metode MARS dapat ditentukan secara otomatis oleh data dan menghasilkan model yang kontinu pada *knot*. Penentuan *knot* pada MARS menggunakan *forward stepwise* dan *backward stepwise*. Untuk mendapatkan model dengan fungsi basis maksimum digunakan *forward stepwise*. Sedangkan *backward stepwise* digunakan untuk memenuhi prinsip parsimoni (model yang sederhana) dengan cara mengeliminasi fungsi basis yang berkontribusi kecil sampai tidak ada fungsi basis yang dapat dieliminasi. Ukuran kontribusi pada *backward stepwise* ditentukan dengan kriteria *Generalized Cross Validation* (GCV). Bentuk fungsi GCV minimum dapat didefinisikan pada persamaan berikut:

$$GCV(M) = \frac{\frac{1}{N} \sum_{i=1}^N [y_i - \hat{f}_M(x_i)]^2}{\left[1 - \frac{C(\hat{M})}{N}\right]^2}$$

dengan,

N = ukuran sampel

y_i = variabel respon

x_i = variabel prediktor

M = banyaknya maksimum fungsi basis

$\hat{f}_M(x_i)$ = nilai taksiran variabel respon pada M fungsi basis di x_i

$C(\hat{M}) = C(M) + dM$; d = nilai ketika setiap fungsi basis mencapai nilai optimal ($2 \leq d \leq 4$)

$C(M) = \text{Trace}[\mathbf{B}(\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T] + 1$; \mathbf{B} adalah matriks dari M fungsi basis

Menurut Otok (2010) klasifikasi pada model MARS untuk nilai variabel respon biner didasarkan pada pendekatan analisis regresi logistik. Sehingga dapat digunakan model probabilitas dengan persamaan sebagai berikut:

$$\text{logit } \pi(x) = \ln\left(\frac{\pi(x)}{1 - \pi(x)}\right) = \alpha_0 + \sum_{m=1}^M \alpha_m \prod_{k=1}^{K_m} [s_{km}(x_{v(k,m)} - t_{km})]_+$$

Menurut Hosmer dan Lemeshow (1989) apabila variabel respon adalah biner maka dapat dikatakan sebagai regresi respon biner. Penentuan klasifikasi pada variabel respon biner (0 dan 1) dilakukan dengan menggunakan titik potong sebesar 0,5 dengan ketentuan jika estimasi peluang melebihi 0,5 maka hasil prediksi adalah masuk ke kelompok 1, jika estimasi peluang kurang dari atau sama dengan 0,5 maka hasil prediksi adalah masuk ke kelompok 0.

2.3 Fuzzy K-Nearest Neighbor

Menurut Prasetyo (2012) seperti halnya pada teori *fuzzy*, sebuah data mempunyai nilai keanggotaan pada setiap kelas yang artinya sebuah data bisa dimiliki oleh kelas yang berbeda dengan nilai derajat keanggotaan dalam interval $[0,1]$.

Formula yang digunakan adalah:

$$u_{(x,c_i)} = \frac{\sum_{k=1}^K u_{(xk,c_i)} \times d_{(x,x_k)}^{\frac{-2}{(m-1)}}}{\sum_{k=1}^K d_{(x,x_k)}^{\frac{-2}{(m-1)}}}$$

dengan:

$u_{(x,c_i)}$: nilai keanggotaan data x ke kelas c_i

K : jumlah tetangga terdekat yang digunakan

$u_{(xk,c_i)}$: nilai keanggotaan data tetangga dalam K tetangga pada kelas c_i , nilainya 1 jika data *training* x_k milik kelas c_i atau 0 jika bukan milik kelas c_i

$d_{(x,x_k)}$: jarak dari data x ke data x_k dalam K tetangga terdekat

m : bobot pangkat (*weight exponent*), yang besarnya : $m > 1$

Menurut Han and Kamber (2006) jarak *Euclidean* memberikan jarak lurus antara dua buah data dengan N dimensi. Rumus jarak *Euclidean* yang digunakan adalah sebagai berikut:

$$d_{(x_a, x_b)} = \sqrt{\sum_{v=1}^N (x_{av} - x_{bv})^2}$$

dengan:

- x_{av} = data testing ke- a pada variabel ke- v
- x_{bv} = data training ke- b pada variabel ke- v
- $d(x_a, x_b)$ = jarak *Euclidean*
- N = dimensi data variabel prediktor
- $(x_{av} - x_{bv})$ = selisih

Penghitungan nilai selisih tergantung pada tipe data yang digunakan. Menurut Prasetyo (2012), penghitungan nilai selisih berdasarkan tipe data untuk tiap variabel dapat diringkas seperti pada Tabel 1. Meskipun FK-NN menggunakan nilai keanggotaan untuk menyatakan keanggotaan data pada setiap kelas, untuk memberikan keluaran akhir hasil prediksi FK-NN memilih kelas dengan nilai keanggotaan terbesar pada data tersebut.

Tabel 1. Formula Jarak Dua Data dengan Satu Atribut

Tipe Atribut	Formula Jarak
Nominal	$d_{ab} = \begin{cases} 0 & \text{Jika } x_a = x_b \\ 1 & \text{Jika } x_a \neq x_b \end{cases}$
Ordinal	$d_{ab} = x_a - x_b / (n - 1)$ n adalah jumlah pengkategorian dalam x
Interval atau Rasio	$d_{ab} = x_a - x_b $

2.4 Evaluasi Kesalahan Hasil Klasifikasi

Menurut Johnson dan Wichern (2007), *Apparent Error Rate* (APER) adalah prosedur evaluasi yang digunakan untuk melihat kesalahan klasifikasi yang dilakukan oleh suatu fungsi klasifikasi. Nilai APER menunjukkan nilai proporsi sampel yang salah diklasifikasikan pada fungsi klasifikasi. Pada umumnya, pengukuran klasifikasi dilakukan dengan matriks konfusi (*confusion matrix*). Jika y_0 dan y_1 merupakan subjek pengklasifikasian, maka bentuk matriks konfusi dapat dilihat pada tabel berikut:

Tabel 2. Matriks Konfusi untuk Klasifikasi Dua Kelas

Hasil Observasi (Actual Class)	Taksiran (Predicted Class)	
	y_0	y_1
y_0	f_{00}	f_{01}
y_1	f_{10}	f_{11}

Dengan f_{00} adalah jumlah data dalam kelas 0 yang dipetakan secara benar ke kelas 0, f_{01} adalah jumlah data dalam kelas 0 yang dipetakan secara salah ke kelas 1, f_{11} adalah jumlah data dalam kelas 1 yang secara benar dipetakan ke kelas 1, f_{10} adalah jumlah data dalam kelas 1 yang dipetakan secara salah ke kelas 0. Untuk menghitung APER digunakan rumus:

$$APER = \frac{f_{01} + f_{10}}{f_{00} + f_{01} + f_{10} + f_{11}} \times 100\%$$

3. METODE PENELITIAN

Data yang digunakan dalam penelitian ini merupakan data sekunder berupa hasil Survei Sosial Ekonomi Nasional (SUSENAS) tahun 2014 di Kabupaten Blora. Data ini diperoleh dari Badan Pusat Statistik Provinsi Jawa Tengah.

Variabel dalam penelitian ini terdiri dari variabel prediktor dan variabel respon. Variabel respon pada penelitian adalah kelompok rumah tangga berdasarkan pengeluaran per kapita per bulan yang sudah dikonversikan berdasarkan garis kemiskinan, dimana pada penelitian ini dikelompokkan kedalam dua kategori yaitu:

1. Kelompok 0, yaitu kelompok rumah tangga miskin.
2. Kelompok 1, yaitu kelompok rumah tangga tidak miskin.

Variabel prediktor pada penelitian ini adalah

1. Jenis kelamin kepala rumah tangga
2. Umur kepala rumah tangga
3. Jumlah anggota rumah tangga
4. Ijazah tertinggi yang dimiliki kepala rumah tangga
5. Kegiatan utama kepala rumah tangga
6. Lapangan usaha utama kepala rumah tangga
7. Status pekerjaan utama kepala rumah tangga
8. Status penguasaan tempat tinggal
9. Jenis atap tempat tinggal
10. Jenis dinding rumah tinggal
11. Jenis lantai rumah tinggal
12. Penggunaan fasilitas tempat buang air besar
13. Tempat pembuangan akhir tinja
14. Sumber air minum
15. Bahan bakar energi utama untuk memasak
16. Pengalaman membeli raskin selama tiga bulan terakhir
17. Ada anggota rumah tangga yang memiliki jaminan pembiayaan/asuransi kesehatan
18. Ada anggota rumah tangga yang dapat menggunakan telepon selular
19. Ada anggota rumah tangga yang menguasai penggunaan komputer desktop

Adapun tahapan yang dilakukan dalam penelitian ini, yaitu :

1. Mempersiapkan data klasifikasi rumah tangga miskin dan rumah tangga tidak miskin di Kabupaten Blora yang terdiri dari variabel prediktor dan variabel respon.
2. Membagi data tersebut menjadi data *training* dan data *testing* dengan proporsi 70:30.
3. Tahapan analisis data yang dilakukan untuk mengklasifikasi kelompok rumah tangga di Kabupaten Blora menggunakan metode *Multivariate Adaptive Regression Spline* (MARS) adalah sebagai berikut:
 - a. Mengkombinasikan banyaknya fungsi basis (*Basis Function* (BF)), *Maximum Interaction* (MI) dan *Minimum Observation* (MO). Fungsi basis yang digunakan adalah antara 2 - 4 kali banyaknya variabel prediktor, maka fungsi basis yang digunakan adalah 38 - 76. *Maximum Interaction* (MI) yang digunakan adalah 1, 2, dan 3, serta *Minimum Observation* (MO) yang digunakan yaitu 0, 1, 2 dan 3.

- b. Membentuk model kelompok rumah tangga di Kabupaten Blora dengan metode pendugaan kuadrat terkecil.
 - c. Memilih model terbaik dengan nilai GCV minimal.
 - d. Menginterpretasi model terbaik yang sudah didapatkan.
 - e. Menghitung klasifikasi berdasarkan model terbaik.
4. Tahapan analisis data yang dilakukan untuk mengklasifikasi kelompok rumah tangga di Kabupaten Blora menggunakan metode *Fuzzy K-Nearest Neighbor* (FK-NN) adalah sebagai berikut:
- a. Menentukan data *testing* dan data *training* dari data SUSENAS tahun 2014 di Kabupaten Blora.
 - b. Menghitung jarak (kedekatan lokasi) antara data *testing* dengan data *training* dengan jarak *Euclidean*.
 - c. Menentukan nilai K.
 - d. Menghitung nilai keanggotaan.
 - e. Menentukan kelas prediksi kelompok rumah tangga di Kabupaten Blora sesuai dengan nilai keanggotaan.
 - f. Menghitung klasifikasi kelompok rumah tangga di Kabupaten Blora.

Setelah hasil klasifikasi dari metode *Multivariate Adaptive Regression Spline* (MARS) dan *Fuzzy K-Nearest Neighbor* (FK-NN) didapatkan, langkah selanjutnya adalah membuat tabel matriks konfusi serta menghitung kesalahan hasil klasifikasi dari kedua metode tersebut dengan menggunakan APER untuk menguji metode manakah yang lebih baik digunakan dalam klasifikasi kelompok rumah tangga di Kabupaten Blora.

4. HASIL DAN PEMBAHASAN

4.1 Metode *Multivariate Adaptive Regression Spline* (MARS)

Berdasarkan kombinasi BF 38, 57, 76 dengan MI 1, 2, 3 serta MO 0, 1, 2, dan 3, model terbaik yang diperoleh yaitu hasil kombinasi BF=76, MI=3, MO=1 karena memiliki nilai GCV minimum sebesar 0,068 dan ketepatan klasifikasi sebesar 92,89 %. Model MARS terbaiknya adalah sebagai berikut:

$$\hat{f}(x) = 0,989 - 0,669 * BF_5 - 0,119 * BF_9 + 0,354 * BF_{11} - 0,649 * BF_{14} \\ + 0,306 * BF_{16} - 0,260 * BF_{26} - 0,004 * BF_{32} + 0,140 * BF_{55} \\ - 0,131 * BF_{59}$$

Dengan,

$$BF_1 = (X_{18} = 1)$$

$$BF_2 = (X_{18} = 2)$$

$$BF_3 = \max\{0, X_3 - 4\} * BF_2$$

$$BF_5 = (X_6 = 3) * BF_2$$

$$BF_7 = (X_{11} = 5) * BF_1$$

$$BF_9 = \max\{0, X_3 - 1\} * BF_7$$

$$BF_{11} = \max\{0, 3 - X_3\} * BF_5$$

$$BF_{12} = (X_{12} = 2) * BF_2$$

$$BF_{14} = (X_{14} = 6 \text{ OR } X_{14} = 7) * BF_{12}$$

$$BF_{16} = (X_1 = 1) * BF_7$$

$$BF_{19} = (X_{13} = 3 \text{ OR } X_{13} = 4 \text{ OR } X_{13} = 5)$$

$$BF_{21} = (X_{15} = 1 \text{ OR } X_{15} = 6 \text{ OR } X_{15} = 7 \text{ OR } X_{15} = 9) * BF_{19}$$

$$BF_{26} = (X_{14} = 2 \text{ OR } X_{14} = 5 \text{ OR } X_{14} = 6) * BF_{21}$$

$$BF_{32} = \max\{0, X_2 - 22\} * BF_3$$

$$BF_{55} = \max\{0, X_2 - 50\} * BF_7$$

$$BF_{59} = \max\{0, X_2 - 48\} * BF_7$$

Berdasarkan model yang telah diperoleh dengan metode MARS. Hasil klasifikasi dapat dilihat pada Tabel 3.

Tabel 3. Hasil Klasifikasi dengan Metode MARS

Hasil Observasi (<i>Actual Class</i>)	Hasil Prediksi (<i>Predicted Class</i>)	
	Kelas 0 (RT Miskin)	Kelas 1 (RT Tidak Miskin)
Kelas 0 (RT Miskin)	1	17
Kelas 1 (RT Tidak Miskin)	0	150

4.2 Metode *Fuzzy K-Nearest Neighbor* (FK-NN)

Hasil laju *error* untuk percobaan nilai K pada *K-Nearest Neighbor* (K-NN) disajikan pada Tabel 4.

Tabel 4. Hasil Laju *Error* untuk Percobaan Nilai K pada FK-NN

Nilai K	Laju <i>Error</i>
1	0,1726
2	0,1726
3	0,1250
5	0,1012
9	0,0952
10	0,0952
11	0,0952
12	0,0952
13	0,1012
22	0,1012
23	0,1012
24	0,1071
25	0,1071
26	0,1071
27	0,1071
28	0,1071
29	0,1071
30	0,1071

Berdasarkan Tabel 4, nilai K yang menghasilkan laju *error* terkecil adalah $K = 9, 10, 11$ dan 12, maka nilai K yang digunakan peneliti dalam penelitian ini adalah $K = 9$, dipilih 9 jarak terdekat. Berdasarkan perhitungan yang telah dilakukan, hasil klasifikasi menggunakan metode *Fuzzy K-Nearest Neighbor* (FK-NN) dapat dilihat pada Tabel 5.

Tabel 5. Hasil Klasifikasi dengan Metode FK-NN

Hasil Observasi (<i>Actual Class</i> s)	Hasil Prediksi (<i>Predicted Class</i>)	
	Kelas 0 (RT Miskin)	Kelas 1 (RT Tidak Miskin)
Kelas 0 (RT Miskin)	2	16
Kelas 1 (RT Tidak Miskin)	0	150

4.3 Evaluasi Kesalahan Hasil Klasifikasi

Ukuran yang digunakan untuk mengevaluasi kesalahan hasil klasifikasi kelompok rumah tangga di Kabupaten Blora adalah APER (*Apparent Error Rate*). Berdasarkan hasil klasifikasi menggunakan metode *Multivariate Adaptive Regression Spline* (MARS), besarnya nilai APER (*Apparent Error Rate*) ditunjukkan pada Tabel 6.

Tabel 6. Perhitungan APER untuk Metode MARS

Hasil Observasi (<i>Actual Class</i> s)	Hasil Prediksi (<i>Predicted Class</i>)	
	Kelas 0 (RT Miskin)	Kelas 1 (RT Tidak Miskin)
Kelas 0 (RT Miskin)	1	17
Kelas 1 (RT Tidak Miskin)	0	150

$$\text{APER} = \frac{17}{1+0+17+150} \times 100 \% = 10,119 \%$$

Berdasarkan hasil klasifikasi menggunakan metode *Fuzzy K-Nearest Neighbor* (FK-NN), besarnya nilai APER (*Apparent Error Rate*) ditunjukkan pada Tabel 7.

Tabel 7. Perhitungan APER untuk Metode FK-NN

Hasil Observasi (<i>Actual Class</i> s)	Hasil Prediksi (<i>Predicted Class</i>)	
	Kelas 0 (RT Miskin)	Kelas 1 (RT Tidak Miskin)
Kelas 0 (RT Miskin)	2	16
Kelas 1 (RT Tidak Miskin)	0	150

$$\text{APER} = \frac{16}{2+0+16+150} \times 100 \% = 9,523 \%$$

5. KESIMPULAN

Berdasarkan hasil dan pembahasan yang telah dipaparkan sebelumnya, maka dapat diambil kesimpulan sebagai berikut:

1. Hasil klasifikasi data *testing* menggunakan metode MARS menghasilkan nilai *Apparent Error Rate* (APER) sebesar 10,119 %.
2. Hasil klasifikasi data *testing* menggunakan metode FK-NN menghasilkan nilai *Apparent Error Rate* (APER) sebesar 9,523 %.
3. Berdasarkan perhitungan *Apparent Error Rate* (APER) menunjukkan bahwa pengklasifikasian menggunakan metode FK-NN lebih baik dibandingkan dengan metode MARS dalam mengklasifikasikan kelompok rumah tangga di Kabupaten Blora berdasarkan rata-rata pengeluaran perkapita perbulan.

6. DAFTAR PUSTAKA

- Badan Pusat Statistik. 2014a. *Indikator Kesejahteraan Rakyat 2014*. Jakarta: Badan Pusat Statistik
- Badan Pusat Statistik. 2014b. *Perhitungan Dan Analisis Kemiskinan Makro Indonesia Tahun 2014*. Jakarta: Badan Pusat Statistik
- Dermoredjo, S. K., dan Pantjar Simatupang. 2003, *Produksi Domestik Bruto, Harga, dan Kemiskinan, Media Ekonomi dan Keuangan Indonesia*, Hal. 191 - 324, Vol. 51, No. 3
- Friedman. J. H. 1991. *Multivariate Adaptive Regression Splines*. The Annals of Statistics, Vol. 19
- Han, J., Kamber, M. 2006. *Data Mining: Concepts and Techniques. Second Edition*. San Fransisco: Morgan Kaufmann Publisers
- Haughton, J. dan Khandker, S. R. 2012. *Pedoman Tentang Kemiskinan dan Ketimpangan*. Jakarta: Salemba Empat
- Hosmer. D. W., Lemeshow. S. 1989. *Applied Logistic Regression*. John Wiley & Sons, Inc
- Johnson. R. A., Wichern. D. W. 2007. *Applied Multivariate Statistical Analysis. Sixth Edition*. New Jersey: Prentice Hall International, Inc
- Otok, B. W. 2010. *Pendekatan Multivariate Adaptive Regression Spline (MARS) pada Pengelompokan Zona Musim Suatu Wilayah*. Jurnal Statistika, Vol. 10, No. 2
- Pintowati, W., Otok, B.W. 2012. *Pemodelan Kemiskinan di Propinsi Jawa Timur dengan Pendekatan Multivariate Adaptive*. Jurnal Sains dan Seni ITS, Vol.1, No.1
- Prasetyo, E. 2012. *Data Mining Konsep dan Aplikasi Menggunakan MATLAB*. Yogyakarta: ANDI