

Pemodelan Kemiskinan di Propinsi Jawa Timur dengan Pendekatan *Multivariate Adaptive*

Wahyuning Pintowati dan Bambang Widjanarko Otok

Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember (ITS)

Jl. Arief Rahman Hakim, Surabaya 60111

E-mail: bambang_wo@statistika.its.ac.id

Abstrak—Masalah kemiskinan merupakan salah satu permasalahan utama pemerintah yang diprioritaskan dalam menyusun strategi pembangunan setiap negara termasuk Indonesia. Selama ini telah banyak program-program atau kebijakan khusus yang dibuat oleh pemerintah guna menanggulangi masalah kemiskinan namun belum juga terpecahkan. Penelitian ini dilakukan untuk mengetahui model kemiskinan di Propinsi Jawa Timur dengan pendekatan *MARS ensemble*. Kemiskinan sendiri oleh BPS diukur dengan tiga indikator yaitu persentase penduduk miskin, indeks kedalaman kemiskinan, dan indeks keparahan kemiskinan. Ketiga indikator kemiskinan tersebut dijadikan sebagai variabel respon dalam penelitian ini yang dimodelkan dengan faktor yang diduga mempengaruhinya baik dari kualitas ekonomi, kualitas sumber daya manusia, dan kesehatan. Hasil penelitian menunjukkan bahwa pendekatan *MARS ensemble* memberikan hasil yang lebih baik dibandingkan dengan pendekatan *MARS* pada ketiga pemodelan tersebut.

Kata Kunci—persentase penduduk miskin, indeks kedalaman kemiskinan, indeks keparahan kemiskinan, *MARS, ensemble*

I. PENDAHULUAN

MASALAH kemiskinan merupakan salah satu permasalahan utama pemerintah yang diprioritaskan dalam menyusun strategi pembangunan setiap negara termasuk Indonesia. Meskipun pertumbuhan ekonomi negara Indonesia semakin meningkat, namun jumlah penduduk juga semakin bertambah sehingga beban ekonomi juga terus bertambah. Pesatnya pertumbuhan penduduk yang tidak seimbang dengan meningkatnya *Gross Domestic Product* (GDP) mengakibatkan semakin luasnya kesenjangan sosial yang terjadi di masyarakat. Ukuran kemiskinan merupakan salah satu indikator keberhasilan pembangunan dan menjadi isu global yang terungkap secara tegas dalam sasaran-sasaran pembangunan (MDGs). Tujuan yang ingin dicapai dalam penelitian ini adalah untuk memodelkan kemiskinan di Propinsi Jawa Timur dengan *MARS ensemble*. Penelitian tentang kemiskinan telah banyak dilakukan antara lain yaitu pemodelan kemiskinan di kabupaten/kota di Provinsi Jawa Tengah dengan menggunakan *Generalized Structured Component Analysis* (GSCA) [1] dan klasifikasi kesejahteraan rumah tangga di Jawa Timur [2].

Hasil studi akhir-akhir ini di bidang peramalan menunjukkan bahwa akurasi peramalan dapat ditingkatkan dengan menggabungkan solusi dari beberapa model dengan kombinasi linier daripada memilih satu solusi berdasarkan model yang

terbaik [3]. Metode penggabungan tersebut sering disebut sebagai pendekatan *ensemble*. Beberapa studi baik secara teori dan empiris menunjukkan bahwa menggabungkan beberapa model yang berbeda merupakan salah satu cara yang efektif untuk meningkatkan kemampuan dalam memprediksi [4]-[5].

Metode yang digunakan dalam penelitian ini adalah *Multivariate Adaptive Regression Spline* (*MARS ensemble*). Teknik *MARS* menjadi populer karena tidak mengasumsikan dan tidak menentukan tipe khusus seperti pada hubungan (linier, kuadratik, kubik) di antara variabel prediktor dan respon [6]. Penelitian tentang *MARS* sendiri pernah dilakukan untuk mengkaji faktor-faktor yang mempengaruhi pemberian ASI eksklusif pada rumah tangga miskin di Propinsi Sulawesi Tengah [7] serta pemodelan indeks harga konsumen kelompok bahan makanan [8].

II. LANDASAN TEORI

A. *Multivariate Adaptive Regression Splines*

Multivariate Adaptive Regression Splines merupakan pendekatan untuk regresi *multivariate* nonparametrik yang dikembangkan oleh Friedman. Model *MARS* difokuskan untuk mengatasi permasalahan dimensi yang tinggi, memiliki variabel banyak, serta ukuran sampel yang besar sehingga diperlukan perhitungan yang rumit. *MARS* merupakan pengembangan dari pendekatan *Recursive Partition Regression* (*RPR*) yang masih memiliki kelemahan dimana model yang dihasilkan tidak kontinu pada knot.

Beberapa hal yang perlu diperhatikan dalam menggunakan model *MARS* adalah (1) *Knot*, yaitu akhir dari sebuah garis regresi (*region*) dan awal dari sebuah garis regresi (*region*) yang lain. Di setiap titik *knot*, diharapkan adanya kontinuitas dari fungsi basis antar satu *region* dengan *region* lainnya. (2) *Basis Function*, yaitu suatu fungsi yang digunakan untuk menjelaskan hubungan antara variabel respon dan variabel prediktor. Fungsi basis ini merupakan fungsi parametrik yang didefinisikan pada tiap *region*. Pada umumnya fungsi basis yang dipilih adalah berbentuk polinomial dengan turunan yang kontinu pada setiap titik *knot*. Friedman menyarankan jumlah maksimum fungsi basis (BF) adalah 2-4 kali jumlah variabel prediktornya. Jumlah maksimum interaksi (MI) adalah 1, 2 dan 3. Minimum jarak antara *knot* atau minimum observasi antara *knot* (MO) sebesar 0, 1, 2, dan 3.

Rujukan [9] menyebutkan bahwa model umum persamaan *MARS* sebagai berikut.

dengan

- α_0 = fungsi basis induk (konstanta)
- α_m = koefisien dari fungsi basis ke- m
- M = banyaknya fungsi basis (*nonconstant basis function*)
- K_m = derajat interaksi
- s_{km} = nilainya 1 atau -1 jika data berada di sebelah kanan titik knot atau kiri titik knot.
- $x_{v(k,m)}$ = variabel prediktor
- t_{km} = nilai knots dari variabel respon $x_{v(k,m)}$

$$\hat{f}(x) = \alpha_0 + \sum_{m=1}^M \alpha_m \prod_{k=1}^{K_m} [s_{km} \cdot (x_{v(k,m)} - t_{km})] \quad (2.1)$$

Metode MARS menentukan knot secara otomatis oleh data dan menghasilkan model yang kontinu pada knot. Penentuan knot pada MARS menggunakan algoritma *forward stepwise* dan *backward stepwise*. Pemilihan model dengan menggunakan *forward stepwise* dilakukan untuk mendapatkan jumlah fungsi basis dengan kriteria pemilihan fungsi basis adalah meminimumkan *Average Sum of Square Residual* (ASR). Untuk memenuhi konsep parsimoni (model yang sederhana) dilakukan *backward stepwise* yaitu membuang fungsi basis yang memiliki kontribusi kecil terhadap respon dari *forward stepwise* dengan meminimumkan nilai *Generalized Cross Validation* (GCV). Pada MARS, beberapa hal yang perlu diperhatikan dalam pemilihan model yang paling optimum (terbaik) adalah jika nilai GCV dari model tersebut mempunyai nilai yang paling rendah (minimum) diantara model-model yang lain.

Fungsi GCV minimum didefinisikan sebagai berikut.

$$GCV(M) = \frac{ASR}{[1 - \frac{C(\hat{M})}{n}]^2} = \frac{\frac{1}{n} \sum_{i=1}^n [y_i - \hat{f}_M(x_i)]^2}{[1 - \frac{C(\hat{M})}{n}]^2} \quad (2.2)$$

dengan

- y_i = variabel respon
- $\hat{f}_M(x_i)$ = nilai taksiran variabel respon pada M fungsi basis
- n = banyaknya pengamatan
- $C(\hat{M}) = C(M) + dM$
- $C(M) = \text{Trace} [\mathbf{B}(\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T] + 1$
- d = nilai ketika setiap fungsi basis mencapai optimasi ($2 \leq d \leq 4$)

B. Pendekatan Ensemble

Akurasi peramalan dapat ditingkatkan dengan menggabungkan solusi dari beberapa model dengan kombinasi linier daripada memilih satu solusi berdasarkan model yang terbaik. Metode penggabungan tersebut sering disebut sebagai pendekatan *ensemble*. Salah satu metode pendekatan yang banyak menarik peneliti adalah pendekatan

dengan mengubah set data training dengan *resampling* atau disebut dengan *bootstrap aggregating* atau disingkat *bagging*. Metode *bagging* pertama kali digunakan oleh Breiman sebagai alat untuk membentuk *classifier* yang lebih stabil dengan menggunakan *bootstrap resampling*. *Bagging* banyak digunakan pada metode klasifikasi dan regresi untuk mereduksi variansi. Teknik ini dapat digunakan untuk memperbaiki stabilitas, meningkatkan akurasi dan kekuatan prediktif [10].

Sebuah *data set* E terdiri dari $\{(y_i, x_i), i=1, 2, \dots, n\}$ sehingga didapatkan $\{(y_i^*, x_i^*), i=1, 2, \dots, n\}$. Dilakukan replikasi *bootstrap* sehingga didapatkan E_i^* . Replikasi *bootstrap* dilakukan sebanyak B kali. Dari setiap data set hasil sampel *bootstrap* selanjutnya dimodelkan dengan MARS. Nilai GCV *bagging* didapatkan dari rata-rata nilai GCV pada setiap pemodelan sampel *bootstrap*.

C. Kemiskinan

Kemiskinan sendiri merupakan bentuk ketidakmampuan untuk meraih kesejahteraan dipandang dari sisi ekonomi dalam memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur dari sisi pengeluaran. Penduduk miskin adalah penduduk yang memiliki rata-rata pengeluaran perkapita perbulan dibawah garis kemiskinan [11].

BPS mendefinisikan garis kemiskinan sebagai nilai rupiah yang harus dikeluarkan seseorang dalam sebulan agar dapat memenuhi kebutuhan dasar asupan kalori sebesar 2100 kkal/hari per kapita (garis kemiskinan makanan) ditambah dengan kebutuhan minuman non makanan yang merupakan kebutuhan seseorang yaitu papan, sandang, sekolah, transportasi dan kebutuhan individu rumahtangga dasar lainnya (garis kemiskinan non makanan). Badan Pusat Statistik menggunakan 3 indikator kemiskinan yaitu persentase penduduk miskin, indeks kedalaman kemiskinan, dan indeks keparahan kemiskinan.

III. METODOLOGI PENELITIAN

A. Sumber Data dan Variabel Penelitian

Data yang digunakan pada penelitian ini adalah data sekunder yang diperoleh dari hasil pendataan Survei Sosial Ekonomi Nasional (SUSENAS) Propinsi Jawa Timur tahun 2010. Unit observasi dalam penelitian ini adalah 38 kabupaten/kota di Propinsi Jawa Timur. Variabel yang digunakan dalam penelitian ini ada dua yaitu variabel respon dan prediktor. Variabel respon yang digunakan yaitu tiga indikator kemiskinan yaitu persentase penduduk miskin (Y_1), indeks kedalaman kemiskinan (Y_2), dan indeks keparahan kemiskinan (Y_3). Sedangkan variabel prediktor yang digunakan ada sepuluh variabel yaitu persentase pengeluaran per kapita untuk non makanan (X_1), persentase penduduk usia 15 tahun ke atas yang bekerja di sektor non pertanian (X_2), persentase penduduk usia 15 tahun ke atas yang bekerja di sektor formal (X_3), angka melek huruf penduduk usia 15-55 tahun (X_4), rata-rata lama sekolah (X_5), persentase penduduk yang tamat SD/SLTP/SLTA/ Perguruan Tinggi (X_6), persentase balita yang kelahirannya ditolong oleh tenaga kesehatan (X_7), angka harapan hidup (X_8),

persentase rumah tangga yang menggunakan jamban sendiri/bersama (X_9), persentase rumah tangga yang menggunakan air bersih (X_{10}).

B. Langkah Analisis

Langkah analisis yang dilakukan untuk mencapai tujuan penelitian ini adalah sebagai berikut.

1. Membuat plot antara variabel respon dengan variabel prediktor.
2. Mendapatkan model MARS terbaik untuk data set awal berdasarkan nilai GCV terkecil.
3. Mendapatkan variabel yang signifikan pada model MARS terbaik untuk data set awal.
4. Melakukan *bagging* dari pasangan variabel respon dan variabel prediktor yang signifikan pada model MARS terbaik untuk data set awal dengan $B = 200$ replikasi *bootstrap*.
5. Melakukan pemodelan MARS pada setiap pengambilan sampel B replikasi *bootstrap* dengan BF, MI, dan MO sama dengan BF, MI, dan MO pada model MARS terbaik untuk data set awal.
6. Mendapatkan nilai GCV pada setiap pemodelan sampel B replikasi *bootstrap*.
7. Mendapatkan GCV *bagging* dari rata-rata GCV pada pemodelan setiap sampel *bootstrap* sebanyak B replikasi. Model MARS *bagging* yang digunakan adalah model MARS terbaik untuk data set awal. Hal ini dikarenakan nilai knot yang berubah-ubah untuk setiap replikasi sehingga estimasi parameternya tidak bisa dirata-rata.

IV. HASIL DAN PEMBAHASAN

A. Plot Variabel Respon dengan Variabel Prediktor

Untuk memperoleh pola hubungan antara variabel respon dengan variabel prediktor dapat dilakukan dengan membuat plot data masing-masing variabel prediktor (X_i sampai dengan X_{10}) terhadap variabel respon.

Plot data antara ketiga variabel respon dengan sepuluh variabel yang diduga mempengaruhinya menunjukkan bahwa ada beberapa plot yang tidak menunjukkan kecenderungan membentuk pola tertentu. Oleh karena itu untuk memodelkan data tersebut tidak dapat digunakan pendekatan regresi parametrik, tetapi digunakan regresi nonparametrik. Pendekatan nonparametrik yang digunakan dalam penelitian ini yaitu MARS dengan alasan data yang digunakan berdimensi tinggi dan tidak ada informasi bentuk kurva regresinya.

B. Pembentukan Model MARS

Dalam penelitian ini metode MARS akan diterapkan dalam pemodelan antara ketiga indikator kemiskinan dengan variabel-variabel yang diduga mempengaruhinya baik dari aspek kualitas ekonomi, sumber daya manusia (SDM), serta kesehatan. Tahap selanjutnya adalah menentukan maksimum jumlah fungsi basis (FB), maksimum interaksi (MI), dan jumlah minimal pengamatan antar knot atau minimum observasi (MO). Fungsi basis merupakan fungsi yang

didefinisikan dari subregion. Fungsi basis umum yang digunakan adalah dua sampai empat kali jumlah variabel prediktor. Variabel prediktor yang digunakan dalam penelitian ini ada sebanyak sepuluh variabel sehingga banyaknya fungsi basis yang digunakan adalah sebanyak 20, 30, dan 40.

Tahapan pembentukan model MARS sendiri dilakukan dengan mencobakan semua kombinasi nilai FB, MI, dan MO yang telah ditentukan sebelumnya. Banyaknya model yang mungkin berdasarkan kombinasi tersebut yaitu sebanyak 36 model. Hasil pemodelan MARS persentase penduduk miskin dapat dilihat dalam Tabel 1.

Tabel 1.
Hasil Pemodelan MARS pada Persentase Penduduk Miskin

No. Model	BF	MI	MO	GCV	Variabel yang Masuk Model
1	20	1	0	10,66	X_2, X_5
2	20	1	1	10,983	X_1, X_2, X_5, X_9
3	20	1	2	9,764	X_1, X_2, X_5, X_9
4	20	1	3	10,845	X_5
5	20	2	0	10,509	X_2, X_5
6	20	2	1	10,861	X_2, X_4, X_5
7*	20	2	2	7,948	X_1, X_2, X_3, X_4, X_5
8	20	2	3	13,711	X_1, X_5, X_6
9	20	3	0	10,689	X_2, X_5
10	20	3	1	10,500	X_2, X_4, X_5
11	20	3	2	10,855	X_2, X_4, X_5, X_9
12	20	3	3	12,951	X_1, X_5, X_6

)* model terbaik

Dari semua kemungkinan model yang telah dicobakan didapatkan model terbaiknya dengan kriteria model yang memiliki nilai GCV terkecil yaitu model ke-7 dengan jumlah FB = 20, MI = 2, dan MO = 2 dengan nilai GCV sebesar 7,948. Model MARS terbaik yang didapatkan yaitu sebagai berikut.

$$Y_1 = 14,414 + 5,392 * BF2 - 1,019 * BF4 + 0,331 * BF6 - 0,202 * BF9 - 0,106 * BF109 \tag{4.1}$$

- BF1 = max (0, $X_5 - 6,170$)
- BF2 = max (0, $6,170 - X_5$)
- BF4 = max (0, $93,190 - X_4$) * BF1
- BF6 = max (0, $43,290 - X_2$)
- BF7 = max (0, $X_1 - 34,890$)
- BF9 = max (0, $13,090 - X_3$) * BF7
- BF10 = max (0, $X_4 - 94,220$) * BF7

Interpretasi model MARS yang tertulis pada persamaan 4.1 adalah sebagai berikut.

- 1) BF2 = max (0, $6,170 - X_5$)
Artinya, koefisien BF1 akan bermakna jika nilai X_5 lebih kecil dari 6,170 maka setiap kenaikan satu fungsi basis (BF2) dapat meningkatkan persentase penduduk miskin sebesar 5,392 pada kabupaten/kota dengan rata-rata lama sekolah kurang dari 6,170 tahun.
- 2) BF4 = max (0, $93,190 - X_4$) * BF1
BF1 = max (0, $X_5 - 6,170$)
Artinya, koefisien BF4 akan bermakna jika nilai X_4 lebih kecil dari 93,190 dan X_5 lebih besar dari 6,170 maka setiap kenaikan satu fungsi basis (BF4) dapat mengurangi persentase penduduk miskin sebesar 1,019 pada kabupaten/kota dengan angka melek huruf penduduk usia

15-55 tahun kurang dari 93,190 dan rata-rata lama sekolah lebih dari 6,170 tahun.

- 3) $BF6 = \max(0, 43,290 - X_2)$
Artinya, koefisien BF6 akan bermakna jika nilai X_2 lebih kecil dari 43,290 maka setiap kenaikan satu fungsi basis (BF6) dapat meningkatkan persentase penduduk miskin sebesar 0,331 pada kabupaten/kota dengan persentase penduduk usia 15 tahun ke atas yang bekerja di sektor non pertanian kurang dari 43,290 persen.
- 4) $BF9 = \max(0, 13,090 - X_3) * BF7$
 $BF7 = \max(0, X_7 - 34,890)$
Artinya, koefisien BF9 akan bermakna jika nilai X_3 lebih kecil dari 13,090 dan X_7 lebih besar dari 34,890 maka setiap kenaikan satu fungsi basis (BF9) dapat mengurangi persentase penduduk miskin sebesar 0,202 pada kabupaten/kota dengan persentase penduduk usia 15 tahun ke atas yang bekerja di sektor formal kurang dari 13,090 persen dan persentase pengeluaran per kapita untuk non makanan lebih dari 34,890 persen.
- 5) $BF10 = \max(0, X_4 - 94,220) * BF7$
 $BF7 = \max(0, X_7 - 34,890)$
Artinya, koefisien BF10 akan bermakna jika nilai X_4 lebih besar dari 94,220 dan X_7 lebih besar dari 34,890 maka setiap kenaikan satu fungsi basis (BF10) dapat mengurangi persentase penduduk miskin sebesar 0,106 pada kabupaten/kota dengan angka melek huruf penduduk usia 15-55 tahun lebih dari 94,220 dan persentase pengeluaran per kapita untuk non makanan lebih dari 34,890 persen.

Tabel 2 menunjukkan tingkat kepentingan variabel prediktor pada fungsi pengelompokan, yang ditaksir oleh kenaikan nilai GCV karena berpindahnya variabel-variabel yang dipertimbangkan tersebut dari model. Persentase pengeluaran per kapita untuk non makanan (X_7) merupakan variabel terpenting pada pemodelan persentase penduduk miskin dengan tingkat kepentingan 100%. Nilai GCV menunjukkan bahwa apabila suatu variabel dimasukkan dalam model maka GCV akan berkurang sebesar nilai GCV pada variabel tersebut. Tingkat kepentingan untuk kesepuluh variabel prediktor yang digunakan dalam pemodelan dapat dilihat pada Tabel 2 berikut.

Tabel 2

Tingkat Kepentingan Variabel Prediktor pada Pemodelan Y_1		
Variabel	Tingkat Kepentingan	GCV
X_7	100.000%	18.703
X_4	86.198%	15.940
X_5	79.924%	14.819
X_2	68.936%	13.059
X_3	28.094%	8.797
X_6	0.000%	7.949
X_7	0.000%	7.949
X_8	0.000%	7.949
X_9	0.000%	7.949
X_{10}	0.000%	7.949

Dari model MARS terbaik untuk indikator tersebut didapatkan variabel yang berpengaruh signifikan ada lima yaitu persentase pengeluaran per kapita untuk non makanan (X_7), persentase penduduk usia 15 tahun ke atas yang bekerja di sektor non pertanian (X_2), persentase penduduk usia 15

tahun ke atas yang bekerja di sektor formal (X_3), angka melek huruf penduduk usia 15-55 tahun (X_4), dan rata-rata lama sekolah (X_5).

Hasil pemodelan MARS indeks kedalaman kemiskinan dapat dilihat dalam Tabel 3.

Tabel 3.
Hasil Pemodelan MARS pada Indeks Kedalaman Kemiskinan

No. Model	BF	MI	MO	GCV	Variabel yang Masuk Model
1	20	1	0	0,423	X_2, X_3, X_4, X_5
2	20	1	1	0,421	X_7, X_3, X_4, X_5
3	20	1	2	0,361	$X_7, X_2, X_3, X_5, X_6, X_7$
4	20	1	3	0,349	$X_2, X_3, X_4, X_5, X_6, X_7$
5*	20	2	0	0,340	X_2, X_3, X_4, X_5, X_9
6	20	2	1	0,434	X_2, X_3, X_5
7	20	2	2	0,478	X_5
8	20	2	3	0,414	X_2, X_4, X_5, X_8
9	20	3	0	0,340	X_2, X_3, X_4, X_5, X_9
10	20	3	1	0,434	X_7, X_2, X_5
11	20	3	2	0,460	X_5
12	20	3	3	0,360	X_2, X_4, X_5, X_6, X_9

)* model terbaik

Dari 36 kemungkinan model yang telah dicobakan didapatkan model terbaiknya dengan kriteria model yang memiliki nilai GCV terkecil yaitu model ke-5 dengan jumlah $BF = 20$, $MI = 2$, dan $MO = 0$ dengan nilai GCV sebesar 0,340. Model MARS terbaik yang didapatkan yaitu sebagai berikut.

$$Y_2 = 2,020 + 0,038 * BF3 + 0,007 * BF6 - 0,004 * BF7 \tag{4.2}$$

- $BF2 = \max(0, 6,650 - X_5)$
- $BF3 = \max(0, X_9 - 32,160) * BF2$
- $BF4 = \max(0, X_3 - 4,590)$
- $BF6 = \max(0, 38,470 - X_2) * BF4$
- $BF7 = \max(0, X_4 - 93,900) * BF4$

Interpretasi model MARS yang tertulis pada persamaan 4.2 adalah sebagai berikut.

- 1) $BF3 = \max(0, X_9 - 32,160) * BF2$
 $BF2 = \max(0, 6,650 - X_5)$
Artinya, koefisien BF3 akan bermakna jika nilai X_9 lebih besar dari 32,160 dan X_5 lebih kecil dari 6,650 maka setiap kenaikan satu fungsi basis (BF3) dapat meningkatkan indeks kedalaman kemiskinan sebesar 0,038 pada kabupaten/kota dengan persentase rumah tangga yang menggunakan jamban sendiri/bersama lebih dari 32,160 dan rata-rata lama sekolah kurang dari 6,650 tahun.
- 2) $BF6 = \max(0, 38,470 - X_2) * BF4$
 $BF4 = \max(0, X_3 - 4,590)$
Artinya, koefisien BF6 akan bermakna jika nilai X_2 lebih kecil dari 38,470 dan X_3 lebih besar dari 4,590 maka setiap kenaikan satu fungsi basis (BF6) dapat meningkatkan indeks kedalaman kemiskinan sebesar 0,007 pada kabupaten/kota dengan persentase penduduk usia 15 tahun ke atas yang bekerja di sektor non pertanian kurang dari 38,470 dan persentase penduduk usia 15 tahun ke atas yang bekerja di sektor formal lebih dari 4,590 persen.
- 3) $BF7 = \max(0, X_4 - 93,900) * BF4$
 $BF4 = \max(0, X_3 - 4,590)$

Artinya, koefisien BF7 akan bermakna jika nilai X_4 lebih besar dari 93,900 dan X_3 lebih besar dari 4,590 maka setiap kenaikan satu fungsi basis (BF7) dapat menurunkan indeks kedalaman kemiskinan sebesar 0,004 pada kabupaten/kota dengan angka melek huruf penduduk usia 15-55 tahun lebih dari 93,900 dan persentase penduduk usia 15 tahun ke atas yang bekerja di sektor formal lebih dari 4,590 persen.

Tingkat kepentingan untuk kesepuluh variabel prediktor yang digunakan dalam pemodelan dapat dilihat pada Tabel 4 berikut.

Tabel 4
Tingkat Kepentingan Variabel Prediktor pada Pemodelan Y_2

Variabel	Tingkat Kepentingan	GCV
X_5	100,000%	1,140
X_9	100,000%	1,140
X_3	55,661%	0,588
X_2	31,190%	0,418
X_4	29,795%	0,411
X_1	0,000%	0,340
X_6	0,000%	0,340
X_7	0,000%	0,340
X_8	0,000%	0,340
X_{10}	0,000%	0,340

Dari model MARS terbaik untuk indikator tersebut didapatkan variabel yang berpengaruh signifikan ada lima yaitu persentase penduduk usia 15 tahun ke atas yang bekerja di sektor non pertanian (X_2), persentase penduduk usia 15 tahun ke atas yang bekerja di sektor formal (X_3), angka melek huruf penduduk usia 15-55 tahun (X_4), rata-rata lama sekolah (X_5), dan persentase rumah tangga yang menggunakan jamban sendiri/bersama (X_9).

Hasil pemodelan MARS indeks keparahan kemiskinan dapat dilihat dalam Tabel 5.

Tabel 5
Hasil Pemodelan MARS pada Indeks Keparaharan Kemiskinan

No. Model	BF	MI	MO	GCV	Variabel yang Masuk Model
1	20	1	0	0,041	X_2, X_5
2	20	1	1	0,037	$X_1, X_2, X_3, X_5, X_7, X_9$
3	20	1	2	0,042	X_1, X_2, X_5
4*	20	1	3	0,030	$X_1, X_2, X_3, X_5, X_7, X_9$
5	20	2	0	0,038	X_2, X_5, X_9
6	20	2	1	0,067	X_2, X_3, X_5, X_6
7	20	2	2	0,049	X_5
8	20	2	3	0,045	X_5, X_{10}
9	20	3	0	0,036	X_2, X_3, X_5, X_9
10	20	3	1	0,048	X_1, X_2, X_3, X_5
11	20	3	2	0,048	X_2, X_5
12	20	3	3	0,045	X_5, X_{10}

)* model terbaik

Dari 36 kemungkinan model yang telah dicobakan didapatkan model terbaiknya dengan kriteria model yang memiliki nilai GCV terkecil yaitu model ke-4 dengan jumlah FB = 20, MI = 1, dan MO = 3 dengan nilai GCV sebesar 0,030. Model MARS terbaik yang didapatkan yaitu sebagai berikut.

$$Y_3 = 0,548 - 0,104 * BF1 + 0,577 * BF2 - 0,024 * BF3 - 0,051 * BF4 + 0,120 * BF5 - 0,013 * BF8 \quad (4.3)$$

$$BF1 = \max(0, X_5 - 6,910)$$

$$BF2 = \max(0, 6,910 - X_5)$$

$$BF3 = \max(0, X_2 - 20,130) * BF2$$

$$BF4 = \max(0, 20,130 - X_2) * BF2$$

$$BF5 = \max(0, X_1 - 52,970)$$

$$BF8 = \max(0, 13,460 - X_3) * BF3$$

Interpretasi model MARS yang tertulis pada persamaan 4.1 adalah sebagai berikut.

1) $BF1 = \max(0, X_5 - 6,910)$

Artinya, koefisien BF1 akan bermakna jika nilai X_5 lebih besar dari 6,910 maka setiap kenaikan satu fungsi basis (BF1) dapat menurunkan indeks keparahan kemiskinan sebesar 0,104 pada kabupaten/kota dengan rata-rata lama sekolah lebih dari 6,910 tahun.

2) $BF2 = \max(0, 6,910 - X_5)$

Artinya, koefisien BF2 akan bermakna jika nilai X_5 lebih kecil dari 6,910 maka setiap kenaikan satu fungsi basis (BF2) dapat meningkatkan indeks keparahan kemiskinan sebesar 0,577 pada kabupaten/kota dengan rata-rata lama sekolah kurang dari 6,910 tahun.

3) $BF3 = \max(0, X_2 - 20,130) * BF2$

$$BF2 = \max(0, 6,910 - X_5)$$

Artinya, koefisien BF3 akan bermakna jika nilai X_2 lebih besar dari 20,130 dan X_5 lebih kecil dari 6,910 maka setiap kenaikan satu fungsi basis (BF3) dapat menurunkan indeks keparahan kemiskinan sebesar 0,024 pada kabupaten/kota dengan persentase penduduk usia 15 tahun ke atas yang bekerja di sektor non pertanian lebih dari 20,130 dan rata-rata lama sekolah kurang dari 6,910 tahun.

4) $BF4 = \max(0, 20,130 - X_2) * BF2$

$$BF2 = \max(0, 6,910 - X_5)$$

Artinya, koefisien BF4 akan bermakna jika nilai X_2 lebih kecil dari 20,130 dan X_5 lebih kecil dari 6,910 maka setiap kenaikan satu fungsi basis (BF4) dapat menurunkan indeks keparahan kemiskinan sebesar 0,051 pada kabupaten/kota dengan persentase penduduk usia 15 tahun ke atas yang bekerja di sektor non pertanian kurang dari 20,130 dan rata-rata lama sekolah kurang dari 6,910 tahun.

5) $BF8 = \max(0, 13,460 - X_3) * BF3$

$$BF3 = \max(0, X_2 - 20,130) * BF2$$

$$BF2 = \max(0, 6,910 - X_5)$$

Artinya, koefisien BF8 akan bermakna jika nilai X_3 lebih kecil dari 13,460 dan X_2 lebih besar dari 20,130 serta X_5 lebih kecil dari 6,910 maka setiap kenaikan satu fungsi basis (BF8) dapat menurunkan indeks keparahan kemiskinan sebesar 0,013 pada kabupaten/kota dengan persentase penduduk usia 15 tahun ke atas yang bekerja di sektor formal kurang dari 13,460, persentase penduduk usia 15 tahun ke atas yang bekerja di sektor non pertanian lebih dari 20,130 dan rata-rata lama sekolah kurang dari 6,910 tahun.

Tingkat kepentingan untuk kesepuluh variabel prediktor yang digunakan dalam pemodelan dapat dilihat pada Tabel 6 berikut.

Tabel 6.
Tingkat Kepentingan Variabel Prediktor pada Pemodelan Y_3

Variabel	Tingkat Kepentingan	GCV
X_5	100,000%	0,131
X_2	18,711%	0,051
X_1	13,168%	0,050
X_3	8,689%	0,049
X_4	0,000%	0,048
X_6	0,000%	0,048
X_7	0,000%	0,048
X_8	0,000%	0,048
X_9	0,000%	0,048
X_{10}	0,000%	0,048

Dari model MARS terbaik untuk indikator tersebut didapatkan variabel yang berpengaruh signifikan ada enam yaitu persentase pengeluaran per kapita untuk non makanan (X_1), persentase penduduk usia 15 tahun ke atas yang bekerja di sektor non pertanian (X_2), persentase penduduk usia 15 tahun ke atas yang bekerja di sektor formal (X_3), rata-rata lama sekolah (X_5), persentase balita yang kelahirannya ditolong oleh tenaga kesehatan (X_7), dan persentase rumah tangga yang menggunakan jamban sendiri/bersama (X_9).

C. Pemodelan MARS Bagging

Variabel yang berpengaruh signifikan pada model MARS terbaik pada ketiga model tersebut selanjutnya dilakukan *bagging* dan dimodelkan kembali dengan 200 kali pengulangan. Nilai GCV yang didapatkan pada tiap-tiap model kemudian dirata-rata untuk mendapatkan GCV *bagging*, Nilai ini selanjutnya dibandingkan dengan nilai dan GCV model MARS terbaik untuk mendapatkan model terbaik. Hasil perbandingan nilai GCV pada model MARS terbaik dengan MARS *bagging* untuk ketiga model dapat dilihat pada Tabel 7 berikut.

Tabel 7
Perbandingan Nilai GCV pada Model MARS dengan MARS *Bagging*

Model	Y1	Y2	Y3
MARS	7.948	0.340	0.030
MARS <i>Bagging</i>	5.233	0.279	0.020

Tabel 7 menunjukkan bahwa untuk ketiga model yang dibentuk, nilai GCV yang didapatkan dari model MARS *bagging* lebih kecil dibandingkan dengan model MARS. Karena tidak ada ketentuan atau batasan nilai tertentu untuk dapat menyimpulkan suatu model lebih baik dibandingkan dengan model yang lain berdasarkan kriteria GCV, begitu nilai GCV MARS *bagging* lebih kecil secara umum disimpulkan bahwa model ini lebih baik dibandingkan dengan model MARS meskipun penurunan nilai GCV dalam penelitian ini relatif kecil. Penurunan nilai GCV pada penelitian ini relatif kecil terutama pada pemodelan Y_3 . Hal ini dimungkinkan karena skala dari data yang kecil. Dengan adanya skala data yang kecil, maka penurunan nilai GCV yang terjadi tampak kurang signifikan.

V. KESIMPULAN

Kesimpulan yang dapat diambil dari penelitian ini adalah pemodelan kemiskinan dengan tiga indikator yaitu persentase penduduk miskin, indeks kedalaman kemiskinan, dan indeks keparahan kemiskinan dengan faktor yang mempengaruhinya baik dari segi ekonomi, pendidikan, kesehatan, dan

lingkungan menunjukkan bahwa pendekatan MARS *ensemble* dalam kasus ini dengan menggunakan *bagging* memberikan hasil yang lebih baik dibandingkan dengan pendekatan MARS. Nilai GCV yang didapatkan dari model MARS *bagging* lebih kecil dibandingkan dengan model MARS. Hal tersebut menunjukkan bahwa *bagging* dalam penelitian ini dapat meningkatkan akurasi dan kekuatan prediktif.

UCAPAN TERIMA KASIH

Penulis W. P. mengucapkan terima kasih kepada Badan Pusat Statistik (BPS) Provinsi Jawa Timur yang telah menyediakan data dalam penelitian ini, Eka Tjipta Foundation yang telah memberikan dukungan finansial bagi penulis selama kuliah, serta semua pihak yang telah memberikan bantuan dalam penyelesaian penelitian ini.

DAFTAR PUSTAKA

- [1] D. F. Ekasari, "Pemodelan SEM dengan Generalized Structured Component Analysis (GSCA)," Tesis Institut Teknologi Sepuluh Nopember, Surabaya (2012).
- [2] D. A. P. Pratama, "Klasifikasi Kesejahteraan Rumah Tangga di Jawa Timur dengan Pendekatan MARS Bagging," Tugas Akhir Institut Teknologi Sepuluh Nopember, Surabaya (2011).
- [3] J. M. Bates dan C. W. J. Granger, "The combination of forecasts," *Operational Research Quarterly* 20 (1969) 451–468.
- [4] A. Krogh dan J. Vedelsby, "Neural Network Ensembles, Cross Validation, and Active Learning," *Advances in Neural Information Processing* 7 (1995) 231-238.
- [5] M. P. Perron dan L. N. Cooper, "When Network Disagree: Ensemble Methods For Hybrid Neural Network," *Neural Network For Speech and Image Processing*. (1993) 126-142.
- [6] I. N. Budiantara, S. Guritno, B. W. Otok, dan F. Suryadi, "Pemodelan B-Spline dan MARS pada Nilai Ujian Masuk Terhadap IPK Mahasiswa Jurusan Disain Komunikasi Visual UK Petra Surabaya," *Jurnal Teknik Industri* vol 8, Surabaya (2006).
- [7] A. Santoso, "Faktor-faktor yang Mempengaruhi Pemberian ASI Eksklusif pada Rumah Tangga Miskin di Propinsi Sulawesi Tengah dengan Pendekatan MARS Bagging," Tesis Institut Teknologi Sepuluh Nopember, Surabaya (2009).
- [8] R. Andriani, "Pemodelan Indeks Harga Konsumen Kelompok Bahan Makanan menggunakan Metode Intervensi dan Regresi Spline," Tugas Akhir Institut Teknologi Sepuluh Nopember, Surabaya (2011).
- [9] J. H. Friedman, "Multivariate Adaptive Regression Splines," *Tech Report 102 Rev, Department of Statistics Stanford University Stanford, California* (1990).
- [10] L. Breiman, "Bagging Prediktor," *Technical Report No. 421. Department of Statistics University of California* (1994).
- [11] BPS Provinsi Jawa Timur, "Survei Sosial Ekonomi Nasional tahun 2010 Provinsi Jawa Timur," BPS Provinsi Jawa Timur, Surabaya (2011).