

**BOOSTING MULTIVARIATE ADAPTIVE REGRESSION SPLINE (MARS)
BINARY RESPONSE UNTUK KLASIFIKASI KEMISKINAN
DI KABUPATEN JOMBANG**

Anna Apriana Hidayanti

Dosen Sekolah Tinggi Ilmu Ekonomi AMM Mataram

Email : anna.apriana89@gmail.com

Abstrac

Poverty is one of the main problems faced by the people of Indonesia since ancient to the present so that a variety of planning, policy and program development has been and will be held on the point is to reduce the number of poor households. Family expectations, conditional cash assistance program, launched of Indonesia Government In 2007, first time in Indonesia. This Program aims to improve the quality of human beings by providing conditional cash assistance in accessing health and education services. The determination of the target family expectations, conditional cash assistance program done by BPS, for the first time used data of year 2005 owned (based on the name and address) subsequently in 2008 PPLS use 14 metric of poverty if household certain worthy relief. Virtue of analysis descriptive verification of data about households in the county of Jombang done by BAPEDA Jombang in 2010 that analysis needs desired or expected by majority of poverty in the county of Jombang expects direct aid, aid and assistance of venture capital, cattle the average who expects blt is worth 35,73 and as much as 28,16 % households in the county of Jombang hope for the help of livestock 17,99 hope for the help of venture capital, physical relief house at 8.28 %, aid daily needs 6.51 % and help or health education by 2,03 %. In this study of predictor variables have a lot then the MARS method can be used as tools that are expected to get the proper classification accuracy levels and to improve the accuracy of classification models and resampling methods used boosting where boosting is one method of ensemble used to improve the accuracy of classification by means of models generate a combination of a model get the best models and the variables that affect the best model with the method of binary response and MARS get accuracy classification needs help households with very poor and poor status of the method of binary response and boosting MARS. Households needs help is in the form of money and not in the form of money. Results of the research there were show by the method of MARS seven variables that affect household models are very poor ownership of residential buildings has the greatest influence on the model and the reduction of the value of the General Cross Validation (GCV), then a variable that has the most influence on models and substractiongcv to the poor is household income every month. Total accuracy classifications MARS and boosting MARS for households with the status of very poor is 70,50 % and 61,93 %, Results with total accuracy classifications MARS and boosting MARS to poor households with status is 64,30 % and 46,34 %.

Kata kunci :*Multivariate Adaptive Regrression Spline, Boosting, Binary Response, Poverty*

1. PENDAHULUAN

1.1 Latar Belakang

Kemiskinan merupakan masalah klasik yang belum tuntas terselesaikan terutama di Negara berkembang, artinya kemiskinan menjadi masalah yang dihadapi dan menjadi perhatian di setiap Negara. Persoalan kemiskinan merupakan salah satu permasalahan pokok yang dihadapi bangsa Indonesia sejak dahulu hingga sekarang, sehingga berbagai perencanaan, kebijakan, serta program pembangunan yang telah dan akan dilaksanakan pada intinya adalah untuk mengurangi jumlah rumah tangga miskin.

Pada tahun 2007 pemerintah Indonesia meluncurkan Program Keluarga Harapan (PKH), program bantuan dana tunai bersyarat pertama di Indonesia. Program ini bertujuan meningkatkan kualitas manusia dengan memberikan bantuan dana tunai bersyarat dalam mengakses layanan kesehatan dan pendidikan tertentu. Penetapan sasaran PKH dilakukan oleh BPS untuk pertama kalinya menggunakan data tahun 2005 yang dimiliki (berdasarkan nama dan alamat) BPS melakukan survey pendidikan dan survey pelayanan kesehatan dan pendidikan guna mengidentifikasi rumah tangga miskin serta fasilitas pendidikan dan kesehatan. Daftar tahun 2005 memuat berada pada sebaran penghasilan terendah dan digunakan sebagai daftar untuk program Bantuan Langsung Tunai (BLT), kemudian pada tahun 2008 pendataan Program Perlindungan Sosial (PPLS) menggunakan 14 indikator kemiskinan apakah rumah tangga tertentu layak mendapatkan bantuan (Nazara, 2013).

Dalam penelitian ini akan digunakan data sekunder yang diperoleh dari Badan Perencanaan Pembangunan Daerah (BAPPEDA) Kabupaten Jombang 2010. Kabupaten Jombang memiliki luas wilayah 1.159,50 Km². Kabupaten Jombang terbagi menjadi 21 kecamatan yang terdiri dari 302 desa dan 4 Kelurahan serta 1.258 dusun. Apabila ditinjau dari komposisi jumlah desa/kelurahan maka Kecamatan Sumobito memiliki jumlah desa terbanyak yaitu 21 buah.

Kabupaten Jombang berdasarkan data BPS 2011 (hasil **SUSENAS** Juli 2010) diketahui persentase penduduk miskin sebesar 13,845 atau sebesar 166.287 jiwa dari total jumlah penduduk Jombang 1.348.199 jiwa, dengan persentase tersebut Kabupaten Jombang masih diatas garis kemiskinan yang ditetapkan jika didasarkan pada tingkat kemiskinan Provinsi yang sebesar 15,26% dan tingkat kemiskinan nasional 13,33% (Indikator Kesejahteraan Provinsi Jawa Timur, TNP2K, 2011). Berdasarkan analisis deskriptif verifikasi data tentang RTM di Kabupaten Jombang yang dilakukan oleh BAPPEDA Jombang pada tahun 2010 bahwa analisis kebutuhan yang diinginkan atau diharapkan oleh sebagian besar RTM di Kabupaten Jombang mengharapkan bantuan langsung (BLT), bantuan ternak dan bantuan modal usaha, Rata-rata yang mengharapkan BLT adalah sebesar 35,73 dan sebesar 28,16% RTM di Kabupaten Jombang mengharapkan bantuan ternak, kemudian sebesar 17,99 mengharapkan bantuan modal usaha, bantuan fisik rumah sebesar 8,28%, bantuan kebutuhan sehari-hari 6,51% dan bantuan pendidikan /kesehatan sebesar 2,03%. Maka untuk mengetahui apakah tingkat keakuratan klasifikasi dari pemberian bantuan yang diharapkan penduduk miskin tersebut sudah tepat atau tidak maka dilakukan uji keakuratan klasifikasi, diketahui bahwa untuk menguji tingkat akurasi klasifikasi dapat menggunakan regresi parametrik dan regresi nonparametrik, karena pada penelitian ini diasumsikan bahwa hubungan variabel respon dan variabel prediktor memiliki hubungan yang tidak diketahui maka digunakan regresi non parametrik dimana dalam penelitian ini memiliki variabel yang berjumlah banyak maka metode MARS dapat digunakan sebagai *tools*.

MARS adalah salah satu model regresi nonparametrik, yaitu model yang tidak mengasumsikan bentuk hubungan fungsional antara variabel respon dan prediktor, dan mempunyai bentuk fungsional yang fleksibel, MARS baik digunakan ketika variabel prediktor yang digunakan berjumlah banyak dan data yang digunakan bersifat tidak linier (Munoz dan Felicimo, 2004 dalam Jalaluddin, 2009). Pada model MARS sendiri akan dilakukan tahap pengklasifikasian untuk mengetahui ketepatan atau tingkat akurasi klasifikasi dari model MARS. Tingkat akurasi dari suatu metode klasifikasi dapat ditingkatkan dengan tujuan memberikan hasil klasifikasi yang lebih baik dan menurunkan tingkat kesalahan klasifikasi maka dilakukan metode *resampling* dalam penyusunan modelnya untuk menurunkan tingkat kesalahan klasifikasi. Metode *resampling* yang digunakan adalah *boosting*. *Boosting* berasal dari kata *boost* yang berarti menaikkan atau menurunkan, *boosting* merupakan salah satu metode ensemble

yang digunakan untuk meningkatkan akurasi dari model klasifikasi dengan cara membangkitkan kombinasi dari suatu model. Dalam penelitian, diharapkan mampu menjelaskan dan mendapatkan tingkat keakuratan klasifikasi secara lebih tepat dan spesifik untuk rumah tangga dengan status sangat miskin dan rumah tangga dengan status miskin dengan metode MARS respon biner maupun dengan metode *boosting* MARS.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dikemukakan, maka permasalahan dalam penelitian ini dapat dirumuskan sebagai berikut :

1. Bagaimana mendapatkan model terbaik bantuan yang diharapkan rumah tangga miskin berdasarkan indikator-indikator kemiskinan di Kabupaten Jombang ?
2. Bagaimana perbandingan tingkat akurasi klasifikasi dari metode MARS respon biner dan Boosting MARS respon biner ?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah di atas, maka tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut :

1. Mendapatkan model terbaik bantuan yang diharapkan rumah tangga miskin berdasarkan indikator-indikator kemiskinan di Kabupaten Jombang
2. Membandingkan tingkat akurasi klasifikasi dari metode MARS respon biner dan Boosting MARS respon biner

1.4 Manfaat Penelitian

Manfaat yang ingin dicapai dari hasil penelitian ini adalah sebagai berikut :

1. Memberikan kontribusi dengan pengembangan metode *boosting* MARS respon biner.
2. Sebagai informasi tambahan yang dapat digunakan BAPPEDA dalam pemberian bantuan terhadap rumah tangga miskin di Kabupaten Jombang

II. TINJAUAN PUSTAKA

2.1 Multivariate Adaptive Regression Spline (MARS)

MARS adalah salah satu metode regresi nonparametik yang diperkenalkan oleh **Friedman (1991)**. Model MARS digunakan untuk mengatasi kelemahan *Recursive Partitioning Regression (RPR)* yaitu menghasilkan model yang kontinu pada knot. RPR digunakan untuk mengatasi kesulitan yang timbul saat penentuan knot pada regresi *dummy* untuk dimensi data yang rendah, penentuan knot tidak akan mengalami kesulitan, tetapi untuk data yang berdimensi tinggi terdapat kesulitan. Untuk mengatasi hal tersebut digunakan model RPR karena penentuan knot tergantung (otomatis) dari data. Tetapi model ini masih terdapat kelemahan yaitu model yang dihasilkan tidak kontinu pada knot, dan untuk mengatasinya digunakan model MARS.

2.2 Klasifikasi pada MARS Respon Biner

Pada model MARS, masalah klasifikasi didasarkan pada pendekatan analisis regresi logistik. Jika variabel respon terdiri dari dua nilai, maka dikatakan sebagai regresi dengan respon biner, sehingga dapat digunakan model probabilitas dengan persamaan berikut (**Otok, 2010**) :

$$\begin{aligned}
 P(Y = 1 | X = x) &= \pi(x) = \frac{e^{f(x)}}{1 + e^{f(x)}} \\
 P(Y = 0 | X = x) &= (1 - \pi(x)) = \frac{1}{1 + e^{f(x)}}
 \end{aligned}
 \tag{2.1}$$

Lemma 1 : (Otok, 2008) Jika hubungan dengan model logistik , $\sigma_L : \mathbb{R} \rightarrow (0,1)$ $\sigma_L = \pi(x) = \left[\frac{e^z}{(1+e^z)} \right]$ maka inver dari σ_L dapat dikatakan sebagai transformasi logit, $\log \text{it } \pi(x) = \ln \left[\frac{\pi(x)}{1-\pi(x)} \right] = z$

Selanjutnya jika Lemma 1 terpenuhi dan $z = \hat{f}(x)$ yaitu,

$z = \hat{f}(x)$ yaitu

$$z = \hat{f}(x) = a_0 + \sum_{m=1}^M a_m \prod_{k=1}^{k_m} S_{km} \cdot (x_{v(k,m)} - t_{K_m})$$

Variabel respon Y merupakan variabel respon biner (0 dan 1) dengan p banyaknya variabel prediktor x (x_1, x_2, \dots, x_p) sehingga model MARS untuk klasifikasi dapat dinyatakan dapat ditulis dalam model :

$$\log \text{it } \pi(x) = \ln \left(\frac{\pi(x)}{1-\pi(x)} \right) = a_0 + \sum_{m=1}^M a_m \prod_{k=1}^{k_m} S_{K_m} \cdot (x_{v(k,m)} - t_{K_m})$$

dan dalam bentuk matriks,

$$\log \text{it } \pi(x) = \mathbf{B}a \tag{2.2}$$

dengan

$$a = (a_0, a_1, \dots, a_n)^T \quad ; \quad \text{dan matriks pada persamaan (2.9)}$$

2.3 Ketepatan Klasifikasi

Untuk menghitung ketepatan klasifikasi pada hasil pengelompokan digunakan *Apparent Error Rate* (APER). Nilai APER menyatakan representasi proporsi sampel yang salah diklasifikasikan. Dalam penelitian kali ini digunakan respon biner sehingga penentuan kesalahan klasifikasi dapat dihitung dari tabel klasifikasi berikut :

Tabel 2.1 Tabel Klasifikasi

| Hasil Observasi | Taksiran | |
|-----------------|-----------------|-----------------|
| | y ₁ | y ₂ |
| y ₁ | n ₁₁ | n ₁₂ |
| y ₂ | n ₂₁ | n ₂₂ |

Keterangan :

n₁₁ = jumlah observasi dari y₁ yang tepat diklasifikasikan sebagai y₁

n₂₂ = jumlah observasi dari y₂ yang tepat diklasifikasikan sebagai y₂

n₂₁ = jumlah observasi dari y₂ yang salah diklasifikasikan sebagai y₁

n₁₂ = jumlah observasi dari y₁ yang salah diklasifikasikan sebagai y₂

n = jumlah observasi

$$APER(\%) = \frac{\text{jumlah prediksi salah}}{\text{jumlah total prediksi}} = \frac{n_{21} + n_{12}}{n} \times 100\% \tag{2.3}$$

$$\text{Total accuracy rate (100\%)} = \frac{\text{jumlah prediksi benar}}{\text{jumlah total prediksi}} = \frac{n_{11} + n_{22}}{n} \times 100\% . \tag{2.4}$$

2.4 Boosting

Boosting merupakan salah satu metode *ensemble* yang digunakan dalam meningkatkan akurasi pada proses klasifikasi dan prediksi dengan cara membangkitkan kombinasi dari suatu model, Ide dasar dari *boosting* adalah pada bobot pada proses *learning* dimana setiap sampel pada proses *training* diatur memiliki bobot *nonnegative* (Okun, 2011) .

Berikut ini adalah langkah-langkah dari algoritma Adaboost:

Input:

- D, suatu himpunan yang terdiri dari sebanyak d class-labeled training tuple
- T, banyaknya classifier yang akan dibangkitkan
- Suatu classification learning scheme

Output:

- Suatu model komposit M^*

2.5 Kemiskinan

Badan Pusat Statistik (BPS) menetapkan indikator kemiskinan di Indonesia. Indikator tersebut adalah sebagai berikut :

- 1) Luas lantai rumah;
- 2) Jenis lantai rumah;
- 3) Jenis dinding rumah;
- 4) Fasilitas tempat buang air besar;
- 5) Sumber air minum;
- 6) Penerangan yang digunakan;
- 7) Bahan bakar yang digunakan;
- 8) Frekuensi makan dalam sehari;
- 9) Kebiasaan membeli daging/ayam/susu;
- 10) Kemampuan membeli pakaian;
- 11) Kemampuan berobat ke puskesmas/poliklinik;
- 12) Lapangan pekerjaan kepala rumah tangga;
- 13) Pendidikan kepala rumah tangga; dan
- 14) Kepemilikan aset

III. METODE PENELITIAN

3.1 Data dan Variabel Penelitian

Data yang digunakan dalam penelitian adalah data sekunder yang diperoleh dari Badan Perencanaan Pembangunan Daerah (BAPPEDA) Kabupaten Jombang tahun 2010. Dalam penelitian sebagai unit observasi adalah unit rumah tangga dengan status sangat miskin dan miskin di Kabupaten/Kecamatan Jombang .

Variabel respon (Y) merupakan variabel yang berisi kelas yang terdiri dari dua kategori ialah kategori (-1) yaitu, kebutuhan yang diperlukan rumah tangga miskin berupa uang dimana bantuan berupa uang berdasarkan jumlah rumah tangga yang mengharapkan bantuan berupa : uang, BLT, dan modal usaha dan kategori (1) yaitu, kebutuhan yang diperlukan rumah tangga miskin berupa non uang dimana bantuan berupa tidak dalam bentuk uang (non uang) : hewan ternak, alat pertanian dan bantuan kesehatan, sedangkan variabel-variabel prediktor menyangkut semua indikator yang termaksud kedalam indikator kesehatan, SDM, dan ekonomi. Variabel-variabel prediktor dijelaskan dalam Tabel 3.2

3.2 Variabel Penelitian

Variabel yang digunakan dalam penelitian ini terbagi menjadi 2 yaitu variabel respon (Y) dan variabel prediktor (X).

Tabel 3.1 : Variabel Prediktor Penelitian

| Variabel | Definisi Operasional | Kategori |
|----------|---|--|
| X_1 | Status penguasaan bangunan tempat tinggal | a: milik sendiri b: kontrak/sewa c: bebas sewa d: rumah dinas |

| | | |
|-----------------|--|--|
| | | e: rumah milik orang tua/saudara f: rumah milik orang tua g: lainnya |
| X ₂ | Luas kavling termasuk bangunan (m ²) | - |
| X ₃ | Luas lantai (m ²) | - |
| X ₄ | Jenis atap terluas | a: beton b: genteng c: kayu sirap d: seng e: asbes f: ijuk/rumbia |
| X ₅ | Jenis dinding terluas | a: tembok b: kayu c: bambu d: lainnya |
| X ₆ | Jenis lantai terluas | a: keramik/marmer/granit b: ubin/tegel/teraso c: semen/bata merah d: kayu/papan e: bambu f: tanah g: lainnya |
| X ₇ | Tempat buang air besar (jamban) | a: milik sendiri b: milik bersama c: umum d: tidak ada |
| X ₈ | Tempat pembuangan air tinja | a: septictank b: kolam/sawah c: sungai/waduk d: lubang tanah e: tanah lapang/kebun f: lainnya |
| X ₉ | Sumber penerangan utama | a: listrik PLN meteran b: listrik PLN bukan meteran (menumpang, dsb.) c: listrik non PLN d: bukan listrik |
| X ₁₀ | Sumber air minum | a: air dalam kemasan b: ledeng c: pompa d: sumur e: mata air f: air sungai g: lainnya |
| X ₁₁ | Bahan bakar yang digunakan dalam memasak | a: listrik b: gas/elpiji c: minyak tanah d: arang kayu/tempurung e: kayu bakar f: lainnya |
| X ₁₂ | Tempat pengobatan | a: RS/puskesmas/pustu b: praktik dokter c: praktik paramedik d: praktik pengobatan tradisional e: lainnya |
| X ₁₃ | Ijazah terakhir kepala keluarga | 1: tidak punya 2: SD/setara 3: SLTP/setara 4: SLTA/setara 5: diplomal/II |
| X ₁₄ | Penghasilan tiap bulan | - |

| | | |
|-----------------|--|---|
| X ₁₅ | Kepemilikan aset dalam rumah | 1. ≤Rp 500.000 2. >Rp 500.000 |
| X ₁₆ | Konsumsi daging/susu/ayam satu kali dalam seminggu | 1. Makan 0-1 kali dalam seminggu 2. Makan lebih atau sama dengan 2 kali seminggu |
| X ₁₇ | Membeli satu set pakaian baru dalam setahun | 1. Membeli hanya 1 set 1x setahun 2. Membeli lebih 1 set 1x setahun |
| X ₁₈ | Frekuensi makan dalam sehari | 1. Makan sebanyak 1x 2x dalam sehari 2. Makan lebih banyak dari 2x dalam sehari |

3.3 Teknik Analisis

Secara garis besar, langkah-langkah yang ditempuh dalam penelitian ini adalah: Pada penelitian ini, data yang digunakan adalah BAPPEDA Kabupaten Jombang tahun 2010, sebelum melakukan deskriptif dan analisis data terlebih dahulu:

1. Menentukan status kemiskinan di Kabupaten Jombang yang digunakan sebagai sampel.
2. Menentukan bantuan apa saja yang masuk dalam kategori variabel bantuan harapan berupa uang dan non uang dengan memperhatikan indikator-indikator kemiskinan penduduk Kabupaten Jombang
3. Mengumpulkan data-data semua variabel yang akan digunakan dalam pembentukan model.
4. Selanjutnya dilakukan tahap analisis deskriptif dan analisis data dengan menentukan banyaknya unit rumah tangga yang dijadikan sampel, sampel minimal yang dapat diambil sebesar 352 unit rumah tangga berstatus sangat miskin akan dipilih secara acak

IV. ANALISIS DAN PEMBAHASAN

4.1 Deskripsi Variabel yang Mempengaruhi Bantuan Yang Diharapkan Penduduk Miskin Kabupaten Jombang

Sebelum memodelkan bantuan yang diharapkan untuk rumah tangga dengan status sangat miskin dan miskin terlebih dahulu akan dideskripsikan tentang variabel yang diteliti. Variabel respon (Y) yang digunakan adalah bantuan yang diharapkan yang terdiri dari dua kategori, yaitu bantuan berupa uang dan non uang yang ditunjukkan pada Tabel 4.1 untuk rumah tangga dengan status sangat miskin. Gambaran variabel prediktor dengan skala data kontinu dengan nilai minimum, maksimum, rata-ratanya serta variabel prediktor yang datanya kategorik dengan persentase frekuensi kelompok-kelompok tertentu.

Tabel 4.1 Jumlah Dan Persentase Bantuan Yang Diharapkan Rumah Tangga Dengan Status Sangat Miskin

| Bantuan Yang Diharapkan | RT Sangat Miskin | |
|-------------------------|------------------|------------|
| | Frekuensi | Persentase |
| Berupa Uang | 180 | 51,1 |
| Non Uang | 172 | 48,9 |
| Total | 352 | 100,0 |

Dari 352 unit rumah tangga dengan status sangat miskin yang dijadikan sampel, bahwa 51,1 persen rumah tangga mengharapkan bantuan berupa uang, sedangkan 48,9 persen rumah tangga megharapkan bantuan berupa non uang

4.2 Pemodelan Bantuan yang Diharapkan Rumah Tangga dengan Status Sangat Miskin Dengan MARS

Pemodelan bantuan yang diharapkan untuk unit rumah tangga dengan status sangat miskin menggunakan pendekatan MARS sesuai langkah-langkah yang telah dijelaskan pada bab 3. Pengolahan dengan *Trial and error* yang dilakukan merujuk dari Friedman (1991) dengan mengkombinasikan banyaknya basis fungsi (BF), maksimum interaksi (MI), dan minimum observasi (MO), banyaknya fungsi basis yang digunakan dalam pengolahan ini adalah 2 sampai dengan 4 kali banyaknya variabel prediktor yang diduga berpengaruh terhadap variabel respon. Maksimum interaksi 1,2 atau 3 dengan pertimbangan jika lebih dari 3 akan menghasilkan model yang sangat kompleks, sedangkan jarak minimum antar knot atau observasi minimum (MO) antar knot sebanyak 0, 1, 2 atau 3.

Maka pada Tabel 4.2 adalah model-model hasil kombinasi pemilihan model terbaik yang dilihat dari GCV terkecil, namun bila GCV bernilai sama maka pertimbangan selanjutnya dengan melihat R^2 terbesar, namun bila beberapa model masih memiliki nilai R^2 yang sama maka dilihat pada model yang memiliki ketepatan klasifikasi terbesar, jika masih memiliki nilai yang sama maka pertimbangan selanjutnya adalah dengan melihat kombinasi BF, MI dan MO terkecil.

Tabel.4.2 Trial And Error Penentuan Model Terbaik MARS Rumah Tangga Dengan Status Sangat Miskin

| Kombinasi | | | GCV | MSE | R ² | Keakuratan Klasifikasi |
|------------|-----------|-----------|--------------|--------------|----------------|------------------------|
| BF | MI | MO | | | | |
| 36 | 1 | 0 | 0,236 | 0,223 | 0,061 | 63,92 |
| 36 | 1 | 1 | 0,237 | 0,226 | 0,058 | 65,63 |
| 36 | 1 | 2 | 0,237 | 0,226 | 0,058 | 65,63 |
| 36 | 1 | 3 | 0,237 | 0,226 | 0,058 | 65,63 |
| 36 | 2 | 0 | 0,238 | 0,216 | 0,051 | 67,33 |
| 36 | 2 | 1 | 0,239 | 0,216 | 0,050 | 68,75 |
| 36 | 2 | 2 | 0,239 | 0,216 | 0,048 | 67,33 |
| 36 | 2 | 3 | 0,238 | 0,216 | 0,052 | 67,33 |
| 36 | 3 | 0 | 0,232 | 0,212 | 0,078 | 66,48 |
| 36 | 3 | 1 | 0,227 | 0,197 | 0,098 | 71,31 |
| 36 | 3 | 2 | 0,228 | 0,195 | 0,092 | 73,01 |
| 36 | 3 | 3 | 0,225 | 0,200 | 0,105 | 64,77 |
| 54 | 1 | 1 | 0,237 | 0,230 | 0,057 | 62,50 |
| 54 | 1 | 2 | 0,237 | 0,230 | 0,057 | 62,50 |
| 54 | 1 | 3 | 0,237 | 0,226 | 0,058 | 65,63 |
| 54 | 2 | 0 | 0,237 | 0,216 | 0,056 | 67,33 |
| 54 | 2 | 1 | 0,239 | 0,216 | 0,048 | 68,75 |
| 54 | 2 | 2 | 0,240 | 0,216 | 0,047 | 67,33 |
| 54 | 2 | 3 | 0,239 | 0,213 | 0,050 | 68,18 |
| 54 | 3 | 0 | 0,233 | 0,212 | 0,074 | 66,48 |
| 54 | 3 | 1 | 0,226 | 0,193 | 0,102 | 72,73 |
| 54 | 3 | 2 | 0,223 | 0,196 | 0,114 | 73,30 |
| 54 | 3 | 3 | 0,228 | 0,194 | 0,093 | 72,44 |
| 72 | 1 | 0 | 0,237 | 0,230 | 0,056 | 62,50 |
| 72 | 1 | 1 | 0,237 | 0,230 | 0,057 | 62,50 |
| 72 | 1 | 2 | 0,237 | 0,230 | 0,057 | 62,50 |
| 72 | 1 | 3 | 0,237 | 0,230 | 0,056 | 62,50 |
| 72 | 2 | 0 | 0,238 | 0,216 | 0,055 | 67,33 |
| 72 | 2 | 1 | 0,238 | 0,215 | 0,054 | 55,11 |
| 72 | 2 | 2 | 0,235 | 0,225 | 0,067 | 63,64 |
| 72 | 2 | 3 | 0,239 | 0,216 | 0,050 | 67,33 |
| 72* | 3* | 0* | 0,222 | 0,203 | 0,115 | 70,45 |
| 72 | 3 | 1 | 0,225 | 0,187 | 0,105 | 74,15 |
| 72 | 3 | 2 | 0,224 | 0,190 | 0,108 | 70,45 |
| 72 | 3 | 3 | 0,229 | 0,209 | 0,087 | 69,32 |

Keterangan : *: kombinasi model terbaik

Berdasarkan kriteria pemilihan model terbaik MARS maka yang terpilih adalah dengan model BF: 72, MI: 3 dan MO: 0 dengan bentuk model :

$$\hat{f}(x) = 0.291 + 0.677 \cdot \text{BF11} + 0.364 \cdot \text{BF13} - 0.383 \cdot \text{BF19} + 0.078 \cdot \text{BF61} + 0.373 \cdot \text{BF68} \quad (4.1)$$

Interpretasi variabel-variabel dalam setiap basis fungsi adalah sebagai berikut :

1. BF11 : memuat interaksi antara variabel status kepemilikan bangunan tempat tinggal (X1), bahan bakar memasak (X11) dan jenis dinding tempat tinggal (X5)
2. BF13 : memuat interaksi antara variabel ijazah terakhir kepala keluarga (X13), tempat buang air besar (X7) dan jenis dinding tempat tinggal (X5)
3. BF19 : memuat interaksi antara variabel tempat buang air besar (X7), bahan bakar memasak (X11) dan jenis dinding tempat tinggal (X5)
4. BF61 : memuat interaksi antara variabel luas lantai (X3), bahan bakar memasak (X11) dan jenis atap tempat tinggal (X4)
5. BF68 : memuat interaksi antara variabel status kepemilikan bangunan tempat tinggal (X1), jenis dinding tempat tinggal (X5) dan jenis atap tempat tinggal (X4)

Pada Tabel 4.3 menampilkan variabel-variabel yang mempengaruhi nilai GCV

Tabel 4.3 Variabel-Variabel Yang Mempengaruhi Pengurangan Nilai GCV Pada RT Dengan Status Sangat Miskin

| No | Variabel | Pengurangan Nilai GCV |
|----|---|-----------------------|
| 1 | Status Kepemilikan Bangunan Tempat Tinggal (X1) | 0,257 |
| 2 | Jenis Dinding Tempat Tinggal (X5) | 0,253 |
| 3 | Jenis Atap Tempat Tinggal (X4) | 0,249 |
| 4 | Bahan Bakar Memasak (X11) | 0,240 |
| 5 | Tempat Buang Air Besar (X7) | 0,232 |
| 6 | Ijazah Terakhir Kepala Keluarga (X13) | 0,228 |
| 7 | Luas Lantai (M ²)(X3) | 0,223 |
| 8 | Luas Kavling M2 (X2) | 0,222 |
| 9 | Jenis Lantai (X6) | 0,222 |
| 10 | Tempat Pembuangan Tinja (X8) | 0,222 |
| 11 | Sumber Penerangan Utama (X9) | 0,222 |
| 12 | Sumber Air Minum (X10) | 0,222 |
| 13 | Tempat Pengobatan (X12) | 0,222 |
| 14 | Penghasilan Tiap Bulan (X14) | 0,222 |
| 15 | Kepemilikan Aset (X15) | 0,222 |
| 16 | Konsumsi Daging atau Susu atau Ayam (X16) | 0,222 |
| 17 | Membeli 1 set pakaian dalam setahun (X17) | 0,222 |
| 18 | Frekuensi makan dalam sehari (X18) | 0,222 |

Pada Tabel di atas dapat terlihat hanya 7 variabel yang memiliki tingkat kepentingan dalam model dari 18 variabel yang digunakan dalam penelitian ini, maka dapat dijelaskan : Variabel status kepemilikan bangunan tempat tinggal memiliki pengaruh paling besar dalam mengurangi nilai GCV sebesar 0,257.

4.3 Evaluasi Keakuratan Klasifikasi pada Model Rumah Tangga dengan Status Sangat Miskin dengan Metode MARS

Evaluasi klasifikasi unit rumah tangga yang mengharapkan bantuan berupa uang (-1) dan berupa non uang (1) berdasarkan model MARS yang diperoleh pada persamaan (4.1), dihitung dengan menggunakan nilai keakuratan klasifikasi dapat dilihat pada Tabel 4.4.

Tabel 4.4 Hasil Klasifikasi dengan Metode MARS RT Sangat Miskin

| Kelas Aktual | Prediksi Kelas | | Keakuratan Klasifikasi (%) |
|----------------------------------|------------------|----------------------|----------------------------|
| | Bantuan Uang(-1) | Bantuan Non_Uang (1) | |
| Bantuan Uang(-1) | 132 | 48 | 73,30 |
| Bantuan Non_Uang (1) | 56 | 116 | 67,44 |
| Keakuratan Klasifikasi Total (%) | | | 70,50 |
| Aper (100% 70,50 %) | | | 29,50 |

Pada Tabel di atas dapat dijelaskan bahwa dari 180 unit rumah tangga sangat miskin yang mengharapkan bantuan uang yang diteliti, sebanyak 132 atau 73,30% unit rumah tangga terklasifikasi dengan tepat, untuk 172 unit rumah tangga sangat miskin yang , mengharapkan bantuan berupa non uang sebanyak 116 atau 67,44% unit rumah tangga terklasifikasi dengan benar, maka secara keseluruhan keakuratan klasifikasi untuk unit rumah tangga sangat miskin dengan metode MARS sebesar 70,50% dengan tingkat kesalahan klasifikasi sebesar 29,50% dan sebesar 73,30 % dari 180 unit rumah tangga yang mengharapkan bantuan uang masuk dalam klasifikasi yang benar (*Sensitivity*) kemudian sebesar 67,44% dari 152 unit rumah tangga yang mengharapkan bantuan Non uang masuk dalam klasifikasi yang benar (*Specificity*).

4.4 Keakuratan Klasifikasi pada Metode *Boosting* MARS RT Sangat Miskin

Pada tahap ini dilakukan *boosting* dari model terbaik yang dapat dilihat pada persamaan (4.1) dengan 5, 10,25 dan 50 replikasi dan untuk tiap replikasi tersebut akan dihitung keakuratan klasifikasinya yang dapat dilihat pada Tabel (4.5)

Tabel 4.5 Keakuratan Klasifikasi Metode *Boosting* MARS RT Sangat Miskin

| Replikasi | KeakuratanKlasifikasi Total (%) |
|-----------|---------------------------------|
| 5 | 61,93 |
| 10 | 60,79 |
| 25 | 60,79 |
| 50 | 60,79 |

Pada Tabel di atas terlihat bahwa keakuratan klasifikasi pada 5 replikasi sebesar 61,93% kemudian untuk 10, 25 sampai 50 replikasi keakuratan klasifikasinya sama sebesar 60,79 % hal ini dapat diartikan bahwa pada saat replikasi 10 keakuratan klasifikasi sudah konvergen dan stabil sehingga penambahan replikasi tidak mempengaruhi tingkat keakuratan klasifikasi dan akan menghasilkan keakuratan klasifikasi yang sama maka keakuratan klasifikasi yang paling baik terjadi 5 replikasi.

Untuk tingkat kesalahan klasifikasi dengan metode *boosting* MARS adalah sebesar 38,07%, sehingga jika dibandingkan dengan tingkat kesalahan pada klasifikasi dengan model MARS 29,50% maka kesalahan klasifikasi pada metode MARS *boosting* lebih besar sehingga dapat disimpulkan pada penelitian bantuan yang diharapkan untuk Kabupaten Jombang metode yang lebih baik digunakan adalah metode MARS klasifikasi terlihat dari tingkat keakuratan klasifikasi keseluruhan sebesar 70,50% .

V. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan pembahasan yang telah dibahas, maka kesimpulan dalam penelitian ini adalah:

Dilihat dari tingkat keakuratan klasifikasi maupun tingkat kesalahan klasifikasi pada metode MARS klasifikasi dan *boosting* MARS pada rumah tangga dengan status sangat

miskin dapat disimpulkan bahwa pada penelitian bantuan yang diharapkan pada Kabupaten Jombang maka metode MARS klasifikasi lebih baik dibandingkan metode boosting MARS, hal ini ditunjukkan dengan tingkat keakuratan klasifikasi keseluruhan sebesar 70,50% dan tingkat kesalahan pada klasifikasi dengan model MARS sebesar 29,50%.

5.2 Saran

Beberapa saran untuk penelitian selanjutnya, dalam pengambilan sampel diharapkan lebih banyak apabila variabel prediktornya berskala kategorik, dan untuk lebih memahami pemilihan sampel yang digunakan, perlunya keseriusan dari pihak BAPPEDA dalam pengambilan data sehingga data yang digunakan lengkap agar hasil klasifikasi yang diperoleh dengan metode MARS klasifikasi maupun *boosting* MARS lebih akurat sehingga mampu mendapatkan tingkat keakuratan klasifikasi lebih tinggi.

Beberapa saran untuk pemerintah adalah dalam pemberian bantuan terhadap rumah tangga miskin sebaiknya lebih memperhatikan faktor-faktor yang mempengaruhi tingkat kebutuhan bantuan rumah tangga tersebut agar pemberian bantuan sesuai dengan kebutuhan dan diharapkan pemerintah lebih terfokus memberikan bantuan untuk perbaikan infrastruktur kategori-kategori dalam variabel terkait.

DAFTAR PUSTAKA

- Badan Pusat Statistik, (2010), *Survei Sosial Ekonomi Juli 2010, Buku 3, Pedoman Pencacahan Kor (Untuk Pencacah dan Kortim)*, BPS, Jakarta.
- Bappenas, (2010,) *Laporan Pencapaian Tujuan Pembangunan Millenium Indonesia 2010*, Kementerian Perencanaan Pembangunan Nasional/Badan Perencanaan Pembangunan Nasional (Bappenas).
- Fernanda W.J, (2013), *Boosting Neural Network Dan Boosting Cart Pada Klasifikasi Diabetes Militus Tipe II* .Thesis .Surabaya : Program Pasca Sarjana, Institute Teknologi Sepuluh Nopember.
- Friedman, J.H., (1991), *Multivariate Adaptive Regression Spline (With Discussion)*, *The Annals of Statistics*, Vol. 19, hal. 1-141.
- Friedman, J.H., (2001). *Greedy function approximation: a gradient boosting machine*. *Ann. Statist.* 29, 1189–1232.
- Jalaluddin, M,(2009), *Pemodelan Partisipasi Anak Dalam Kegiatan Ekonomi di Sumatera Barat Menggunakan Regresi Logistik dan MARS.*: Tesis Institut Teknologi Sepuluh Nopember, Serabaya.
- Okun, O.(2011). *Feature Selection and Ensemble in Machine Learning Bioinformatics: Algorithmic Classification and Implementations*. United States of America: IGI Global
- Otok, B.W. (2010), *Pendekatan Multivariate Daptive Regression Spline (MARS) Pada Pengelompokan Zona Musim Suatu Wilayah*. *Jurnal statistik* Vol 10.Nopember 2010. FMIPA ITS Surabaya
- Otok, B.W. (2008), *Pendekatan Bootstrap Pada Model Multivariate Adaptive Regression Splines (MARS)*, Desertasi, Universitas Gadjah Mada, Yogyakarta.
- Otok, B.W. Guritno, S., Subanar, Haryatmi, S. (2006), *Bootstrap dalam MARS untuk Klasifikasi Perbankan*. *Inferensi Jurnal Statistik*, Volume 2, NO. 1, Januari 2006. FMIPA ITS Surabaya.