

PENERAPAN METODE POHON REGRESI *STEPWISE LINEAR* DENGAN ALGORITMA *GUIDE* DALAM MENGANALISIS PENGARUH KINERJA PROGRAM GIZI TERHADAP PREVALENSI *UNDERWEIGHT* DI INDONESIA

ISNANI, IZZATI RAHMI HG, YUDIANTRI ASDI

*Program Studi S1 Matematika,
Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Andalas,
Kampus UNAND Limau Manis Padang, Indonesia.
email : isnani.math14@gmail.com*

Diterima 9 Maret 2019 Direvisi 7 April 2019 Dipublikasikan 7 Mei 2019

Abstrak. *Underweight* merupakan keadaan gizi kurang yang merupakan akibat dari kekurangan asupan zat gizi yang masuk ke dalam tubuh. Berdasarkan Pemantauan Status Gizi (PSG) tahun 2017, wilayah Indonesia mengalami prevalensi *underweight* sebesar 17,8% yang mana prevalensi ini melewati batas aman kejadian *underweight* menurut WHO (*World Health Organization*). Oleh karena itu, perlu dianalisis pengaruh kinerja program gizi terhadap prevalensi *underweight* di Indonesia agar dapat dijadikan acuan untuk menyelesaikan permasalahan *underweight* di Indonesia. Salah satu metode yang dapat digunakan yaitu metode pohon regresi *stepwise linear* dengan algoritma *GUIDE* (*Generalized, Unbiased Interaction Detection and Estimation*). Hasil analisis data dengan menggunakan metode ini menunjukkan bahwa kejadian *underweight* dapat dikelompokkan menjadi 17 kelompok berdasarkan karakteristiknya oleh 8 kinerja program gizi, sedangkan model yang terbentuk disetiap simpul dipengaruhi oleh 13 kinerja program gizi, dimana ketepatan model yang dihasilkan meningkat dari 0,2796 menjadi 0,6227. Dengan kata lain, dugaan yang diperoleh dari model yang terbentuk mampu menerangkan pengaruh kinerja program gizi terhadap prevalensi *underweight* di Indonesia.

Kata Kunci: Pohon Regresi *Stepwise Linear*, Algoritma *GUIDE*, Prevalensi *Underweight*

1. Pendahuluan

Underweight merupakan keadaan gizi kurang yang merupakan akibat dari kekurangan asupan zat gizi yang masuk ke dalam tubuh. Gizi buruk dan gizi kurang adalah status gizi yang didasarkan pada indeks Berat Badan menurut Umur (BB/U). Berdasarkan Pemantauan Status Gizi (PSG) tahun 2017, prevalensi *underweight* secara nasional pada tahun 2017 adalah 17,8%, terdiri dari 3,8% status gizi buruk dan 14% status gizi kurang [4]. Hasil ini menunjukkan angka yang sama dengan prevalensi *underweight* secara nasional di tahun 2016 yaitu sebesar 17,8% terdiri dari 3,4% status gizi buruk dan 14,4% status gizi kurang [3].

Berdasarkan data tersebut, perlu adanya upaya untuk mengatasi masalah gizi, yaitu dengan memperhatikan masalah dan kinerja program gizi di Indonesia. Salah satu metode yang dapat diterapkan pada masalah dan kinerja program gizi di Indonesia yaitu metode pohon regresi *stepwise linear* dengan algoritma GUIDE. Metode *stepwise linear* merupakan salah satu prosedur terbaik dalam menyeleksi peubah karena dapat mencegah masuknya lebih banyak peubah penjelas daripada yang diperlukan sambil memperbaiki persamaannya pada setiap tahap, sehingga penggunaannya direkomendasikan [2]. Sedangkan keunggulan dari pemakaian algoritma *GUIDE* yaitu dapat mengatasi masalah penyimpangan (bias) [6].

2. Landasan Teori

2.1. Analisis Regresi Linier Berganda

Analisis regresi linier berganda yaitu analisis regresi yang digunakan untuk memodelkan hubungan antara satu peubah respon (Y) dengan beberapa peubah penjelas (X). Model dari analisis regresi linier berganda yaitu [7]:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \cdots + \beta_k X_{ki} + \varepsilon_i \quad (2.1)$$

dimana,

- Y_i : nilai peubah respon pada pengamatan ke- i ($i = 1, 2, \dots, n$)
- X_{ji} : nilai peubah penjelas ke- j pada pengamatan ke- i ($j = 1, 2, \dots, k$)
- β_m : koefisien regresi ke- m ($m = 0, 1, \dots, k$)
- ε_i : galat pada pengamatan ke- i ($i = 1, 2, \dots, n$)

Ketepatan model regresi dapat dilihat dari koefisien determinasi. Nilai koefisien determinasi adalah ukuran kebaikan model regresi. Koefisien determinasi dapat dirumuskan sebagai berikut [7]:

$$R^2 = \frac{JKR}{JKT} \times 100\% = \frac{\mathbf{b}^T \mathbf{X}^T \mathbf{Y} - n\bar{Y}^2}{\mathbf{Y}^T \mathbf{Y} - n\bar{Y}^2} \times 100\%. \quad (2.2)$$

2.2. Analisis Regresi Bertatar (Stepwise)

Regresi bertatar (*stepwise*) merupakan salah satu prosedur pemilihan persamaan regresi terbaik dengan cara menyisipkan peubah penjelas satu per satu sampai diperoleh persamaan regresi yang memuaskan. Prosedur dasarnya adalah sebagai berikut [2]:

- (1) Hitung korelasi semua peubah penjelas (X) dengan peubah respon (Y). Pilih peubah penjelas yang paling berkorelasi dengan peubah respon sebagai peubah yang pertama kali dimasukkan ke dalam regresi, misalkan X_1 .
- (2) Hitung persamaan regresi linier ordo pertama $\hat{Y} = f(X_1)$.
- (3) Uji koefisien regresi yang terbentuk untuk mengetahui apakah peubah X_1 nyata atau tidak. Jika hasilnya nyata, cari peubah penjelas kedua untuk dimasukkan

ke dalam persamaan regresi dengan memeriksa koefisien korelasi parsial Y dengan semua peubah penjelas yang berada di luar persamaan regresi, yaitu X_j , dimana $j \neq 1$. Dengan kata lain, koreksi hubungan garis lurus antara Y dan X_j dengan X_1 . Misalkan peubah yang terpilih adalah X_2 . Selanjutnya cari persamaan regresi kedua $\hat{Y} = f(X_1, X_2)$. Namun, jika hasilnya tidak nyata maka proses berhenti dan ambil $Y = \bar{Y}$ sebagai model terbaik.

- (4) Lakukan lagi uji koefisien regresi pada persamaan regresi kedua, perhatikan peningkatan nilai R^2 . Kemudian nilai- F parsial untuk kedua peubah yang ada di dalam persamaan diuji dengan cara membandingkan nilai- F parsial terendah dengan nilai- F tabel. Jika uji ini nyata maka peubah dipertahankan di dalam model. Sebaliknya, jika uji ini tidak nyata maka peubah dikeluarkan dari model.
- (5) Proses terus berlanjut sampai akhirnya tidak ada lagi peubah yang akan dimasukkan atau dikeluarkan dari model.

2.3. Metode Pohon Regresi Stepwise Linear

Metode pohon regresi secara teknis dikenal sebagai metode penyekatan rekursif biner. Prosesnya adalah biner karena kumpulan data yang disebut simpul selalu disekat menjadi dua sekatan yang disebut simpul anak. Pada setiap simpul dibentuk suatu model regresi yang diperoleh dari prosedur *stepwise linear*. Langkah-langkah pembentukan pohon regresi yaitu [1]:

- (1) Pemilihan Peubah Penyekat.
Pembentukan pohon regresi dilakukan dengan cara penyekatan data pengamatan pada tiap simpul yang dibagi ke dalam dua simpul anak. Aturan penyekatannya adalah sebagai berikut [1]:
 - (a) Tiap penyekatan hanya bergantung pada nilai yang berasal dari satu peubah penjelas.
 - (b) Jika peubah penjelas numerik X_j , penyekatan yang diperbolehkan adalah $X_j \leq c$ untuk $c \in \mathbb{R}$, dimana c adalah nilai tengah antara dua nilai amatan peubah X_j yang berurutan dan berbeda nilai. Misal X_j mempunyai n nilai yang berbeda maka terdapat sebanyak $n - 1$ penyekatan.
 - (c) Jika peubah penjelas kategorik X_j , penyekatan yang terjadi berasal dari semua kemungkinan penyekatan berdasarkan terbentuknya dua anak gugus yang saling lepas (*disjoint*). Misal peubah X_j merupakan peubah kategorik nominal dengan L kategori, maka akan ada $2^{L-1} - 1$ penyekatan, sedangkan jika peubah X_j merupakan peubah kategorik ordinal, maka akan ada $L - 1$ penyekatan.

Berdasarkan aturan tersebut, tentukan semua kemungkinan penyekatan untuk setiap peubah penjelas. Kemudian pilih sekatan terbaik dari semua kemungkinan sekatan tersebut dan sekat simpul menjadi simpul kiri dan simpul kanan.

Jumlah kuadrat sisaan pada simpul ke- t dijadikan sebagai kriteria kehomogenan di dalam masing-masing simpul. Misalkan simpul t berisi anak con-
toh (X_n, Y_n) , n adalah banyaknya amatan, dan $\hat{y}(t)$ diduga dari substitusi nilai

peubah ke dalam model di simpul t tersebut, maka jumlah kuadrat sisaan pada simpul t dinyatakan sebagai:

$$JKS(t) = \sum_{X_n \in t} [y_n - \hat{y}(t)]^2. \quad (2.3)$$

Misalkan ada penyekatan s yang menyekat t menjadi simpul anak kiri t_L dan simpul anak kanan t_R yang dinyatakan sebagai berikut:

$$JKS(t_L) = \sum_{X_n \in t_L} [y_n - \hat{y}(t_L)]^2, \quad (2.4)$$

dan

$$JKS(t_R) = \sum_{X_n \in t_R} [y_n - \hat{y}(t_R)]^2, \quad (2.5)$$

maka total jumlah kuadrat sisaan dapat ditulis sebagai berikut:

$$JKS_{total} = JKS(t_L) + JKS(t_R), \quad (2.6)$$

sehingga fungsi penyekat yang digunakan adalah:

$$\emptyset(s, t) = JKS(t) - JKS_{total},$$

atau

$$\emptyset(s, t) = JKS(t) - JKS(t_L) - JKS(t_R). \quad (2.7)$$

Sedangkan untuk sekatan terbaik s^* adalah:

$$\emptyset(s^*, t) = \max_{s \in S} \emptyset(s, t). \quad (2.8)$$

(2) Penghentian Sekatan.

Suatu simpul t tidak dapat disekat lagi jika memenuhi salah satu kondisi dari aturan penghentian sekatan sebagai berikut [1]:

- (a) Jika pohon regresi mencapai batas nilai maksimum kedalaman pohon dari spesifikasi.
- (b) Jika ukuran amatan pada simpul induk kurang dari ukuran minimum spesifikasi, atau berisi pengamatan dengan jumlah yang terlalu sedikit.

(3) Penentuan Nilai Dugaan Respon.

Pada metode pohon regresi, jika suatu simpul merupakan simpul akhir maka dugaan respon bagi pengamatan dalam simpul akhir tersebut adalah rata-rata dugaan responnya.

2.4. Metode GUIDE

Metode GUIDE (*Generalized, Unbiased Interaction Detection and Estimation*) merupakan salah satu prosedur dalam pembentukan pohon regresi dengan menggunakan algoritma untuk menyusun model sesuai dengan partisi data yang disajikan dalam pohon bilangan berpasangan (*binary decision tree*).

Berikut adalah beberapa algoritma GUIDE yang digunakan untuk dugaan linier dan memilih antara sepasang uji interaksi [5].

Algoritma 1 : Uji khi-kuadrat untuk dugaan linier.

- (a) Tentukan sisaan dari dugaan model linier untuk peubah- n dan peubah- f , keluarkan peubah- s dan peubah- c .
- (b) Untuk setiap peubah- n , bagi data menjadi empat kelompok pada kuartil contoh, bentuk tabel kontingensi 2×4 dengan tanda dari sisaan (positif versus negatif) sebagai baris dan kelompok data sebagai kolom. Hitung jumlah observasi untuk setiap baris dan kolom serta hitung statistik χ^2 dan p -value tabel dari distribusi $\chi^2_{(3)}$.
- (c) Lakukan hal yang sama untuk setiap peubah- s dan peubah- c . Selanjutnya, kategori dari peubah digunakan untuk membentuk kolom dari tabel kontingensi. Hilangkan kolom dengan jumlah observasi pada kolom sama dengan nol.
- (d) Untuk mendeteksi interaksi antara sepasang peubah- n (X_i, X_j), bagi ruang contoh (X_i, X_j) menjadi empat kuadran dengan menyekat *range* dari setiap peubah menjadi dua bagian pada median contoh. Bentuk tabel kontingensi 2×4 dengan menggunakan tanda sisaan sebagai baris dan kuadran sebagai kolom. Hitung statistik χ^2 dan p -value. Selanjutnya, hilangkan kolom dengan jumlah observasi pada kolom sama dengan nol.
- (e) Lakukan hal yang sama untuk setiap pasangan peubah- s .
- (f) Lakukan juga hal yang sama untuk setiap pasangan peubah- c , gunakan pasangan nilai tersebut untuk membagi ruang contoh. Sebagai contoh, jika X_i dan X_j masing-masing mempunyai c_i dan c_j kategori, statistik χ^2 dan p -value dihitung dari tabel dengan 2 baris jumlah kolom $c_i c_j$ dikurangi dengan jumlah kolom yang jumlah observasi pada kolom sama dengan nol.
- (g) Hitung statistik χ^2 dan p -value untuk setiap pasang (X_i, X_j) dimana X_i adalah peubah- n dan X_j adalah peubah- c . Jika X_j mempunyai c kategori, bentuk tabel kontingensi berukuran $2 \times (2c - \text{jumlah kolom dengan jumlah observasi pada kolom sama dengan nol})$.
- (h) Dengan cara yang sama, hitung statistik χ^2 dan p -value untuk setiap pasang (X_i, X_j) dimana X_i adalah peubah- s dan X_j adalah peubah- c .
- (i) Terakhir, lakukan hal yang sama untuk setiap pasang (X_i, X_j) dimana X_i adalah peubah- s dan X_j adalah peubah- n seperti pada langkah d.

Jika nilai p -value terkecil berasal dari uji *curvature*, maka peubah dengan p -value terkecil tersebut dipilih untuk menyekat simpul. Jika p -value terkecil berada pada uji interaksi, maka gunakan algoritma berikut untuk memilih salah satu peubah sebagai penyekat simpul.

Algoritma 2 : Memilih peubah di antara sepasang uji interaksi.

Misalkan sepasang peubah dipilih karena uji interaksinya paling signifikan di antara semua uji *curvature* dan uji interaksi.

- (a) Jika diantaranya tidak ada peubah- n , pilih salah satu peubah dengan p -value *curvature* terkecil.
- (b) Jika keduanya merupakan peubah- n , sekat simpul sepanjang nilai tengah dari tiap peubah. Pilih peubah yang menghasilkan jumlah kuadrat sisaan minimum.
- (c) Jika salah satu peubah pasti merupakan peubah- n , maka pilih peubah lainnya untuk menyekat simpul.

2.5. Masalah dan Kinerja Program Gizi

Data yang digunakan dalam penelitian ini yaitu data masalah dan kinerja program gizi pada 430 kabupaten/kota di Indonesia yang diperoleh dari Pemantauan Status Gizi (PSG) tahun 2017 [4].

Masalah gizi yang dibahas pada penelitian ini yaitu kejadian *underweight*. *Underweight* merupakan keadaan gizi kurang yang merupakan akibat dari kekurangan asupan zat gizi yang masuk ke dalam tubuh. Prevalensi *underweight* dijadikan sebagai peubah respon pada penelitian ini.

Kinerja program gizi digunakan sebagai peubah penjelas dalam penelitian ini yang terdiri dari 15 program, yaitu : kabupaten/kota di Indonesia (X_1), inisiasi menyusui dini ≥ 1 jam (X_2), inisiasi menyusui dini < 1 jam (X_3), ASI dalam 24 jam terakhir (X_4), ASI eksklusif (X_5), balita yang memiliki kartu menuju sehat (X_6), vitamin A balita usia 6-59 bulan (X_7), balita kurus dapat pemberian makanan tambahan (X_8), timbang ≥ 4 kali (X_9), ibu hamil risiko kekurangan energi kronik (X_{10}), wanita usia subur risiko kekurangan energi kronik (X_{11}), ibu hamil kekurangan energi kronik dapat pemberian makanan tambahan (X_{12}), ibu hamil dapat tablet tambah darah ≥ 90 tablet (X_{13}), ibu hamil dapat tablet tambah darah < 90 tablet (X_{14}), dan konsumsi garam beryodium (X_{15}).

3. Pembahasan

3.1. Pembentukan Pohon Regresi

Pembentukan pohon regresi dapat dilakukan dengan prosedur berikut:

(1) Pemilihan Peubah Penyekat.

Tahap pertama dalam membentuk pohon regresi adalah pemilihan peubah penyeikat dengan menggunakan algoritma *GUIDE*. Penyekat yang terpilih adalah penyeikat dengan nilai *p-value* terkecil. Awal mulanya bentuk suatu model dari 430 data dengan menggunakan prosedur *stepwise linear* sehingga diperoleh model $Y = 27,037 - 0,080X_5 - 0,069X_8 + 0,1513X_{10} + 0,2718X_{11} - 0,0703X_{15}$ pada simpul induk (t_1). Kemudian lakukan prosedur *GUIDE* untuk memilih peubah penyeikat pada t_1 . Ternyata peubah pertama yang menyeikat simpul induk adalah peubah balita kurus dapat pemberian makanan tambahan, sehingga peubah ini menjadi faktor dominan yang mampu mengelompokkan prevalensi *underweight* dengan baik. Peubah balita kurus dapat pemberian makanan tambahan ini kemudian menyeikat simpul induk (t_1) menjadi simpul kiri (t_2) dan simpul kanan (t_3), dimana untuk $X_8 \leq 40,95$ sebanyak 173 data masuk ke (t_2) sedangkan untuk $X_8 > 40,95$ sebanyak 257 data masuk ke (t_3).

Lakukan hal yang sama untuk membentuk suatu model pada setiap simpul. Kemudian simpul-simpul tersebut akan diseikat lagi oleh peubah penjelas lainnya. Prosedur ini terus berlangsung sampai dipenuhinya kriteria penghentian sekatan.

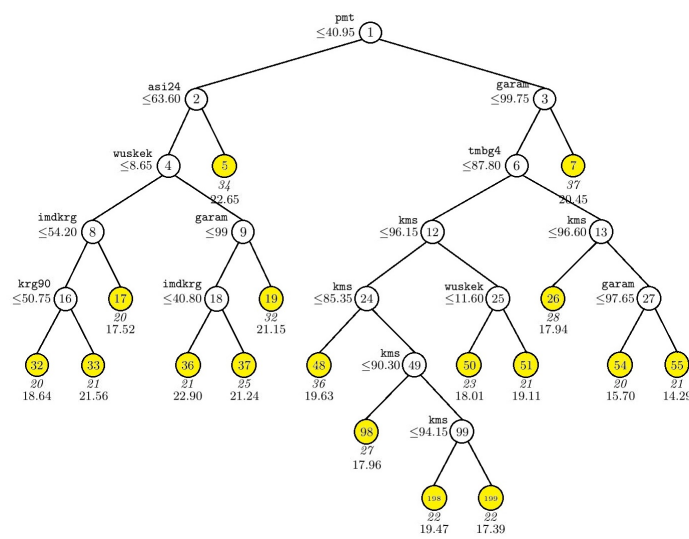
(2) Penghentian Sekatan.

Pembentukan pohon regresi akan dihentikan ketika struktur pohon sudah men-

capai kedalaman 10 atau data pengamatan pada setiap simpul kurang dari 40 amatan namun tidak kurang dari 20 amatan.

(3) Penentuan Nilai Dugaan Respon.

Nilai dugaan respon pada setiap simpul akhir diperoleh dengan cara mensubstitusikan nilai peubah penjelas ke dalam model regresi yang terbentuk di setiap simpul akhir.



Gambar 1. Struktur Pohon Regresi

3.2. Analisis Pohon Regresi

Dari Gambar 1 dapat dilihat bahwa pohon regresi yang terbentuk dibangun oleh 33 simpul, dimana 1 simpul induk (*root node*), 16 simpul dalam (*internal node*), dan 17 simpul akhir (*terminal node*). Berdasarkan pohon regresi yang terbentuk, kejadian *underweight* dapat dibagi menjadi 17 kelompok. Kelompok karakteristik pohon regresi didasarkan pada peubah penyekat yang mampu mengelompokkan data dengan baik. Dapat dilihat bahwa pohon regresi disekat oleh 8 peubah dimana peubah balita kurus dapat pemberian makanan tambahan merupakan peubah penyekat pertama yang paling mampu mengelompokkan data dengan baik. Kemudian diikuti oleh peubah-peubah pemberian ASI dalam 24 jam terakhir, wanita usia subur risiko kekurangan energi kronik, inisiasi menyusui dini < 1 jam, konsumsi garam beryodium, timbang ≥ 4 kali, balita yang memiliki kartu menuju sehat, dan ibu hamil dapat tablet tambah darah < 90 tablet.

Cara memilih peubah penyekat pada tiap simpul yaitu dengan mengikuti algoritma *GUIDE*. Sedangkan cara untuk menetapkan nilai peubah yaitu sekat simpul sepanjang nilai tengah dari data berurutan. Kemudian cari satu per satu jum-

lah kuadrat sisaan di setiap simpul kiri dan simpul kanan. Kemudian cari jumlah kuadrat sisaan di simpul tersebut dengan rumus pada Persamaan 2.7. Kemudian pilih sekatan yang mampu menghasilkan penurunan jumlah kuadrat sisaan terbesar. Pada simpul induk, sekatan terbaik yang terpilih yaitu $X_8 \leq 40,95$. Begitu juga untuk simpul kiri (t_2) dan simpul kanan (t_3). Algoritma *GUIDE* akan menganggap t_2 dan t_3 sebagai simpul induk yang baru. Kemudian langkah pada algoritma *GUIDE* akan diulang dari langkah pertama. Selengkapnya, karakteristik pohon regresi tersebut dapat dilihat pada Tabel 1.

Tabel 1. Karakteristik Pohon Regresi pada Setiap Simpul

Simpul	n	Dugaan y	Peubah Penciri
32T	20	18,63	$X_8 \leq 40,98$; $X_4 \leq 63,6$; $X_{11} \leq 8,65$; $X_3 \leq 54,2$; $X_{14} \leq 50,75$
33T	21	21,56	$X_8 \leq 40,98$; $X_4 \leq 63,6$; $X_{11} \leq 8,65$; $X_3 \leq 54,2$; $X_{14} > 50,75$
17T	20	17,52	$X_8 \leq 40,98$; $X_4 \leq 63,6$; $X_{11} \leq 8,65$; $X_3 > 54,2$
36T	21	22,90	$X_8 \leq 40,98$; $X_4 \leq 63,6$; $X_{11} > 8,65$; $X_{15} \leq 99$; $X_3 \leq 40,8$
37T	25	21,24	$X_8 \leq 40,98$; $X_4 \leq 63,6$; $X_{11} > 8,65$; $X_{15} \leq 99$; $X_3 > 40,8$
19T	32	21,15	$X_8 \leq 40,98$; $X_4 \leq 63,6$; $X_{11} > 8,65$; $X_{15} > 99$
5T	34	22,65	$X_8 \leq 40,98$; $X_4 > 63,6$
48T	36	19,63	$X_8 > 40,98$; $X_{15} \leq 99,75$; $X_9 \leq 87,8$; $X_6 \leq 96,15$; $X_6 \leq 85,35$
98T	27	17,96	$X_8 > 40,98$; $X_{15} \leq 99,75$; $X_9 \leq 87,8$; $X_6 \leq 96,15$; $X_6 > 85,35$; $X_6 \leq 90,3$
198T	22	19,47	$X_8 > 40,98$; $X_{15} \leq 99,75$; $X_9 \leq 87,8$; $X_6 \leq 96,15$; $X_6 > 85,35$; $X_6 > 90,3$; $X_6 \leq 94,15$
199T	22	17,39	$X_8 > 40,98$; $X_{15} \leq 99,75$; $X_9 \leq 87,8$; $X_6 \leq 96,15$; $X_6 > 85,35$; $X_6 > 90,3$; $X_6 > 94,15$
50T	23	18,01	$X_8 > 40,98$; $X_{15} \leq 99,75$; $X_9 \leq 87,8$; $X_6 > 96,15$; $X_{11} \leq 11,6$
51T	21	19,11	$X_8 > 40,98$; $X_{15} \leq 99,75$; $X_9 \leq 87,8$; $X_6 > 96,15$; $X_{11} > 11,6$
26T	28	17,94	$X_8 > 40,98$; $X_{15} \leq 99,75$; $X_9 > 87,8$; $X_6 \leq 96,6$
54T	20	15,69	$X_8 > 40,98$; $X_{15} \leq 99,75$; $X_9 > 87,8$; $X_6 > 96,6$; $X_{15} \leq 97,65$
55T	21	14,29	$X_8 > 40,98$; $X_{15} \leq 99,75$; $X_9 > 87,8$; $X_6 > 96,6$; $X_{15} > 97,65$
7T	37	20,45	$X_8 > 40,98$; $X_{15} > 99,75$

Dapat dilihat bahwa peubah penyekat yang paling dominan pada setiap simpul adalah peubah balita yang memiliki kartu menuju sehat dan diikuti oleh peubah balita kurus dapat pemberian makanan tambahan. Namun peubah penyekat terbaik pada pohon regresi tersebut adalah peubah balita kurus dapat pemberian makanan tambahan yang kemudian disusul oleh peubah ASI dalam 24 jam terakhir sebagai peubah penyekat terbaik kedua. Dengan adanya pengelompokkan data berdasarkan karakteristik tersebut, maka diharapkan model mampu menghasilkan prevalensi *underweight* dengan ketepatan yang baik. Dari proses pengolahan *software GUIDE* dapat dilihat model dari setiap simpul seperti pada Tabel 2.

Pada Tabel 2 dapat dilihat bahwa peubah yang berperan dalam pembentukan model ada sebanyak 13 peubah yaitu peubah inisiasi menyusu dini ≥ 1 jam, inisiasi menyusu dini < 1 jam, ASI eksklusif, balita yang memiliki kartu menuju sehat, vitamin A balita usia 6-59 bulan, balita kurus dapat pemberian makanan tambahan, timbang ≥ 4 kali, ibu hamil risiko kekurangan energi kronik, wanita usia subur risiko kekurangan energi kronik, ibu hamil kekurangan energi kronik dapat pemberian makanan tambahan, pemberian tablet tambah darah ≥ 90 tablet, pemberian tablet tambah darah < 90 tablet, dan konsumsi garam beryodium. Dapat dilihat bahwa ada dua peubah yang paling sering muncul dalam model, yaitu peubah

Tabel 2. Model Regresi pada Setiap Simpul

Simpul	Model Regresi	R^2
32T	$Y = 17,286 - 0,31073X_8 + 0,73516X_{10}$	0,6288
33T	$Y = 423,06 + 1,2174X_2 - 0,14260X_5 + 0,78886X_{10} - 0,29108X_{13} - 4,0699X_{15}$	0,7991
17T	$Y = -4,3751 - 0,39582X_5 - 0,87278X_6 + 1,2335X_7$	0,7917
36T	$Y = 75,805 + 0,29577X_6 - 0,88589X_7$	0,3868
37T	$Y = 27,852 - 0,32712X_2 + 0,36943X_3 - 0,26058X_{15}$	0,7387
19T	$Y = 6,9172 - 0,21460X_3 + 0,23451X_9 + 0,52068X_{10}$	0,4101
5T	$Y = 2,7110 + 0,73604X_2 + 0,11944X_5 + 0,36140X_{10} + 0,06881X_{13}$	0,7488
48T	$Y = 10,221 + 0,11617X_3 + 0,48413X_{11} + 0,093055X_{12} - 0,17252X_{13}$	0,6723
98T	$Y = 21,184 - 0,11664X_5 + 0,54589X_{10} - 0,086004X_{15}$	0,6562
198T	$Y = 35,260 - 0,20480X_5 - 0,22668X_8 + 0,54181X_{11} + 0,16840X_{12} - 0,13496X_{14}$	0,8422
199T	$Y = 59,199 - 0,44305X_7$	0,2115
50T	$Y = 14,034 - 0,25989X_5 + 0,28602X_9 - 0,74106X_{10}$	0,6834
51T	$Y = 19,110$	0,3799
26T	$Y = -1,5571 + 0,12254X_3 - 0,72967X_6 + 0,83502X_7 + 0,60768X_{10} - 0,12848X_{15}$	0,7771
54T	$Y = 26,271 - 0,097207X_5 - 0,14477X_8 + 0,51420X_{11}$	0,5737
55T	$Y = 94,562 - 0,22396X_2 - 0,92896X_9 + 0,41912X_{10}$	0,6468
7T	$Y = 20,449$	0,3588

ibu hamil risiko kekurangan energi kronik dan peubah ASI eksklusif. Maka dapat dikatakan kedua peubah ini merupakan peubah yang paling berpengaruh dalam pembentukan model. Artinya peubah ini dapat menjelaskan keragaman prevalensi *underweight* dengan baik pada setiap kelompok pengamatan.

Pada pembentukan model regresi untuk semua data pengamatan diperoleh R^2 sebesar 0,2796. Setelah dilakukan penyekatan lebih lanjut, ternyata model regresi pada 17 kelompok pengamatan mengalami peningkatan pada 16 kelompok pengamatan dimana terdapat 1 kelompok yang mengalami penurunan yang tidak terlalu signifikan. Artinya dengan melakukan penyekatan, ketepatan model menjadi lebih baik.

Dengan adanya pembentukan pohon regresi, data dapat dikelompokkan dengan baik. Hal ini dapat dilihat dari pembentukan pohon regresi yang menghasilkan R^2 sebesar 0,6227. Artinya terjadi peningkatan R^2 dari 0,2796 menjadi 0,6227. Dapat dikatakan bahwa ketepatan model menjadi lebih baik dengan adanya proses pembentukan pohon regresi.

4. Kesimpulan

Dari uraian sebelumnya dapat disimpulkan bahwa:

- (1) Peubah-peubah yang berperan dalam pembentukan pohon regresi terdiri dari 8 peubah, dimana peubah balita kurus dapat pemberian makanan tambahan merupakan peubah penyekat pertama yang paling mampu mengelompokkan data dengan baik, diikuti oleh peubah ASI dalam 24 jam terakhir, wanita usia subur risiko kekurangan energi kronik, inisiasi menyusui dini kurang dari 1 jam, konsumsi garam beryodium, timbang lebih dari 4 kali, balita yang memiliki kartu menuju sehat, dan ibu hamil dapat tablet tambah darah kurang dari 90 tablet.

- (2) Ketepatan model pada pembentukan model regresi untuk semua data pengamatan yaitu sebesar 0,2796. Setelah dilakukan penyekatan, ternyata ketepatan model regresi mengalami peningkatan pada 16 kelompok dan mengalami penurunan yang tidak terlalu signifikan pada 1 kelompok. Artinya dengan melakukan penyekatan, ketepatan model menjadi lebih baik. Selain itu, terdapat dua peubah yang paling berpengaruh dalam pembentukan model yaitu peubah ibu hamil risiko kekurangan energi kronik dan peubah ASI eksklusif.
- (3) Pohon regresi yang terbentuk menghasilkan R^2 sebesar 0,6227. Artinya terjadi peningkatan R^2 dari 0,2796 menjadi 0,6227. Dengan kata lain, dugaan yang diperoleh dari model yang terbentuk mampu menerangkan pengaruh kinerja program gizi terhadap prevalensi *underweight* di Indonesia.

5. Ucapan Terima kasih

Penulis mengucapkan terima kasih kepada Dr. Ferra Yanuar, Dr. Dodi Devianto, Monika Rianti Helmi, M.Si, dan Radhiatul Husna, M.Si yang telah memberikan kritikan dan masukan sehingga makalah ini dapat diselesaikan dengan baik.

Daftar Pustaka

- [1] Breiman, et al. 1984. *Classification And Regression Tree*. Chapman and Hall, New York
- [2] Draper, N.R. dan H. Smith. 1981. *Analisis Regresi Terapan Edisi Kedua*. Diterjemahkan oleh : Bambang Sumantri. Gramedia Pustaka Utama, Jakarta
- [3] Kementerian Kesehatan Republik Indonesia. 2017. Hasil Pemantauan Status Gizi (PSG) dan Penjelasannya Tahun 2016. http://www.kesmas.kemkes.go.id/assets/upload/dir_519d41d8cd98f00/files/Buku-Saku-Hasil-PSG-2016_842.pdf , diakses pada Minggu, 4 November 2018 pukul 20.53 WIB
- [4] Kementerian Kesehatan Republik Indonesia. 2018. Buku Saku Pemantauan Status Gizi Tahun 2017. http://www.kesmas.kemkes.go.id/assets/upload/dir_519d41d8cd98f00/files/Buku-Saku-Nasional-PSG-2017_975.pdf, diakses pada Minggu, 4 November 2018 pukul 21.10 WIB
- [5] Loh, W.Y. 2002. *GUIDE User Manual*. Department of Statistics University of Wisconsin, Madison
- [6] Loh, W.Y. 2002. Regression Tree with Unbiased Variable Selection and Interaction Detection. *Statistica Sinica*. **12** (2002) : 361 – 386
- [7] Montgomery, D.C and E.A. Peck. 1992. *Introduction to Linear Regression Analysis 2nd Edition*. Wiley, New York