

Analisis Status Indeks Desa Membangun di Kabupaten Buleleng dengan Metode CHAID

Grashella C N Ginting^{1‡}, Aam Alamudi², and Itasia Dina Sulvianti³

¹²³Department of Statistics, IPB University, Indonesia

[‡]corresponding author: grashellaclaragtginting@apps.ipb.ac.id

Copyright © 2023 Grashella C N Ginting, Aam Alamudi, and Itasia Dina Sulvianti. This is an open-access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

The integration of the SDGs into the regional development agenda is being carried out as an effort to accelerate national development. To achieve the accuracy of village development targets, the Ministry of Villages established the village development index. The province of Bali has the highest average IDM score in 2021. Of the nine districts in Bali Province, Buleleng Regency has the lowest IDM value, which is 0,7361. Therefore, the purpose of this study is to identify the variables that affect IDM in Buleleng Regency using Chi-squared Automatic Interaction Detection (CHAID) method. There is an imbalance of data from three categories of response variables. The method that is used to handle that is Synthetic Minority Oversampling Technique (SMOTE). The analysis in this study developed the CHAID model without SMOTE, with SMOTE, merging response variables categories without and with SMOTE. The results of this study indicate that the CHAID model with SMOTE without merging response variables categories performs better in classifying the status of the village development index in Buleleng Regency. The obtained accuracy is 57,3%. The variables that influenced the IDM status in Buleleng Regency are population density, distance to the regency capital, and number of recipients of direct cash assistance.

Keywords: CHAID, development, SMOTE, village development index.

*Received: Oct 2023; Reviewed: Nov 2023; Published: Des 2023

1. Pendahuluan

Sustainable Development Goals (SDGs) Desa merupakan salah satu upaya yang dapat dilakukan dalam pencapaian pembangunan keberlanjutan. Pengintegrasian SDGs ke dalam agenda pembangunan daerah diharapkan dapat mempercepat pencapaian tujuan pembangunan nasional karena desa memiliki potensi yang sangat besar dalam hal tersebut (Iskandar 2020). Pembangunan desa, sebagai upaya membangun fondasi, direncanakan secara berkelanjutan untuk menciptakan kesejahteraan bagi rakyat Indonesia dan mengurangi daerah – daerah yang miskin dan tertinggal. Dalam upaya mencapai ketepatan sasaran pembangunan desa, Kementerian Desa (2015) membentuk Indeks Desa Membangun (IDM) yang terdiri dari aspek ketahanan sosial, ketahanan ekonomi dan ketahanan ekologi. Status desa menurut nilai IDMnya dibagi menjadi desa mandiri, maju, berkembang, tertinggal, dan sangat tertinggal.

Provinsi Bali memiliki nilai rata-rata IDM tertinggi pada tahun 2021, yaitu sebesar 0,8037. Selain itu Provinsi Bali juga memiliki jumlah kabupaten dengan status IDM mandiri terbanyak, yaitu lima kabupaten. Dari sembilan kabupaten pada Provinsi Bali, Kabupaten Buleleng memiliki nilai IDM terendah yaitu sebesar 0,7361 dengan status IDM “maju” (Kementerian Desa 2021). Berdasarkan data tersebut maka akan diidentifikasi faktor-faktor yang memengaruhi status IDM di Kabupaten Buleleng.

Identifikasi peubah-peubah mengenai status IDM di Kabupaten Buleleng dapat menggunakan metode klasifikasi. Metode klasifikasi yang digunakan pada penelitian ini adalah pohon klasifikasi. Salah satu metode pohon klasifikasi yang dapat digunakan adalah *Chi-squared Automatic Interaction Detection (CHAID)*. Metode CHAID merupakan teknik iteratif yang digunakan untuk menguji peubah-peubah penjelas terhadap peubah respon dalam klasifikasi dengan menggunakan tingkat signifikansi statistik khi-kuadrat. Metode CHAID memiliki kelebihan, yaitu dapat mendeteksi setiap interaksi pada tiap-tiap pembagian pada pohon klasifikasi dan melihat faktor yang berpengaruh signifikan secara jelas (Gallagher *et al.* 2000). Suhendra *et al.* (2020) pada penelitiannya mengenai klasifikasi pemberian Kartu Keluarga Sehat (KKS) di Kota Semarang menghasilkan ketepatan klasifikasi metode CHAID sebesar 90,2% menunjukkan hasil lebih baik dibandingkan dengan metode regresi logistik dengan ketepatan klasifikasi sebesar 88%. Ketepatan klasifikasi CHAID juga lebih tinggi dibandingkan dengan metode C4.5 pada klasifikasi penderita diabetes melitus tipe 2 pada penelitian Faisal *et al.* (2017). Selain itu pada penelitian Sulviana (2018) menggunakan metode CHAID dengan peubah respon multiclass menghasilkan nilai akurasi yang cukup baik, yaitu sebesar 71,4%.

Penelitian ini menggunakan status IDM sebagai peubah respon. Terdapat ketidakseimbangan banyaknya desa yang tergolong status mandiri, maju, dan berkembang pada Kabupaten Buleleng. Ketidakseimbangan data merupakan keadaan ketika suatu kelas lebih besar dibandingkan dengan kelas lainnya. Metode yang digunakan dalam upaya penanganan data tidak seimbang adalah *Synthetic Minority Oversampling Technique*. Dalam penelitian Salsabella (2017) data tidak seimbang untuk pemodelan risiko HIV di Papua ditangani dengan metode SMOTE. Penelitian tersebut berhasil menaikkan nilai sensitivitas sebesar 34,31%. Selain itu, pada penelitian Suantari (2020) mengenai penerapan SMOTE pada metode CHAID untuk mengklasifikasi tingkat loyalitas pelanggan menghasilkan nilai sensitivitas yang meningkat dari 67,41% menjadi 87,34%. Hal ini menunjukkan bahwa metode SMOTE efektif dalam meningkatkan ketepatan klasifikasi pada kelas minoritas.

Penelitian ini menggunakan status indeks desa membangun sebagai peubah respon. Terdapat ketidakseimbangan banyaknya desa yang tergolong status mandiri, maju, dan berkembang pada Kabupaten Buleleng. Metode yang digunakan dalam upaya penanganan data tidak seimbang adalah *Synthetic Minority Oversampling Technique* (SMOTE). Tujuan penelitian ini adalah: (1) memperoleh model terbaik dari hasil pengklasifikasian status Indeks Desa Membangun (IDM) Kabupaten Buleleng, (2) mengidentifikasi peubah yang memengaruhi status Indeks Desa Membangun (IDM) di Kabupaten Buleleng.

2. Metodologi

2.1 Bahan dan Data

Data yang digunakan pada penelitian ini adalah data sekunder yang diperoleh dari Potensi Desa (PODES) tahun 2021, peringkat Indeks Desa Membangun tahun 2021, dan publikasi setiap kecamatan di Kabupaten Buleleng dalam angka tahun 2021. Data ini terdiri dari 128 desa di Kabupaten Buleleng. Peubah yang digunakan disajikan pada Tabel 2 berikut.

Tabel 1 Peubah-peubah yang digunakan

Kode	Peubah	Keterangan	Sumber Pustaka
Y	Status Indeks Desa Membangun (IDM)	1: Mandiri 2: Maju 3: Berkembang	Ditjen PDP (2021)
X_1	Pendidikan tertinggi kepala desa	1: Tidak pernah sekolah 2: Tidak tamat SD/Sederajat 3: Tamat SD/Sederajat 4: SMP/Sederajat 5: SMA/Sederajat 6: Akademi/DIII 7: Diploma IV/S1 8: S2 9: S3	Mahayana (2013)
X_2	Banyaknya keluarga penerima Bantuan Langsung Tunai (BLT)	1: Rendah 2: Sedang 3: Tinggi	Nain (2019)
X_3	Topografi	1: Puncak/tebing 2: Lereng 3: Dataran 4: Lembah	Adinugroho et al. (2016)
X_4	Jarak ke ibukota kabupaten	1: Dekat 2: Sedang 3: Jauh	Adinugroho et al. (2016)
X_5	Kepadatan penduduk	1: Rendah 2: Sedang 3: Tinggi	Antara dan Suryana (2020)

2.2 Prosedur Analisis Data

Proses analisis pada penelitian ini menggunakan *software* SPSS version 26, MS. Excel, dan python. Tahapan- tahapan yang dilakukan dalam penelitian ini sebagai berikut:

1. Melakukan praproses data. Praproses data yang dilakukan berupa kategorisasi peubah numerik menjadi peubah kategorik. Pengkategorisasian dilakukan dengan menggunakan pendekatan sebaran dari data, pendekatan ini cukup efektif karena menunjukkan distribusi data dengan cukup jelas dan membagi data dengan baik. Berikut merupakan pengkategorian setiap peubah:
 - a. Peubah banyaknya keluarga penerima BLT (X_2) dikategorikan menjadi tiga kategori, yaitu rendah (27 keluarga $<X_2 < 119$ keluarga), sedang (119 keluarga $<X_2 < 211$ keluarga), dan tinggi ($X_2 > 211$ keluarga).
 - b. Peubah jarak ke ibukota kabupaten (X_4) dikategorikan menjadi tiga kategori, yaitu dekat (2km $<X_4 < 22$ km), sedang (22km $<X_4 < 42$ km), dan jauh ($X_4 > 42$ km)
 - c. Peubah kepadatan penduduk (X_5) dikategorikan menjadi tiga kategori, yaitu rendah (80 jiwa/km² $<X_5 < 1149,17$ jiwa/km²), sedang (1149,17 jiwa/km² $<X_5 < 2209,17$ jiwa/km²), dan tinggi ($X_5 > 2209,17$ jiwa/km²)
2. Melakukan eksplorasi data untuk melihat gambaran umum dan sebaran data.
3. Membentuk model klasifikasi CHAID

Analisis CHAID dilakukan menggunakan *software* SPSS version 26. Model CHAID dibentuk dengan batas kedalaman pohon adalah 3, minimal kasus pada *parent node* dan *child node* adalah 10 dan 5, serta level signifikansi sebesar 0,05. *Parent node* adalah simpul yang menjadi pemisah awal, sedangkan *child node* adalah simpul-simpul yang dihasilkan dari pemisahan. Validasi silang dilakukan secara otomatis pada *software* SPSS dengan 10-fold.

Terdapat empat model yang akan dibentuk, yaitu:

- a) Model tanpa penerapan SMOTE.
- b) Model dengan penerapan SMOTE.
Proses SMOTE yang diterapkan pada penelitian ini menggunakan *package imblearn.oversampling* pada *software* python, membutuhkan parameter *random_state* dan *k_neighbours*. Parameter *random_state* digunakan untuk mengatur angka acak sehingga hasil SMOTE yang dihasilkan bernilai sama untuk setiap eksekusi. Parameter *random_state* yang digunakan pada penelitian ini adalah sebesar 5. Parameter *k_neighbours* digunakan sebagai penentu jumlah *k* tetangga terdekat yang dibutuhkan untuk membuat data buatan. Pada penelitian ini menggunakan $k = 5$ yang berarti data buatan berasal dari 5 data kelas minor yang jaraknya berdekatan.
- c) Model dengan penggabungan kategori peubah respon mandiri dan maju tanpa penerapan SMOTE.
- d) Model dengan penggabungan kategori peubah respon mandiri dan maju dengan penerapan SMOTE.
Setelah dilakukan penggabungan peubah respon kategori mandiri dan maju, selanjutnya dilakukan penerapan SMOTE untuk mengatasi

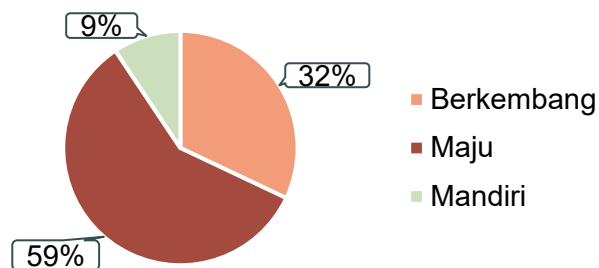
ketidakseimbangan data dengan proses penerapan SMOTE sama seperti langkah b.

4. Mengevaluasi ketepatan klasifikasi model yang diperoleh dan memilih model klasifikasi terbaik.
5. Menentukan peubah-peubah yang memengaruhi status IDM Kabupaten Buleleng dan menginterpretasi model.
6. Kesimpulan.

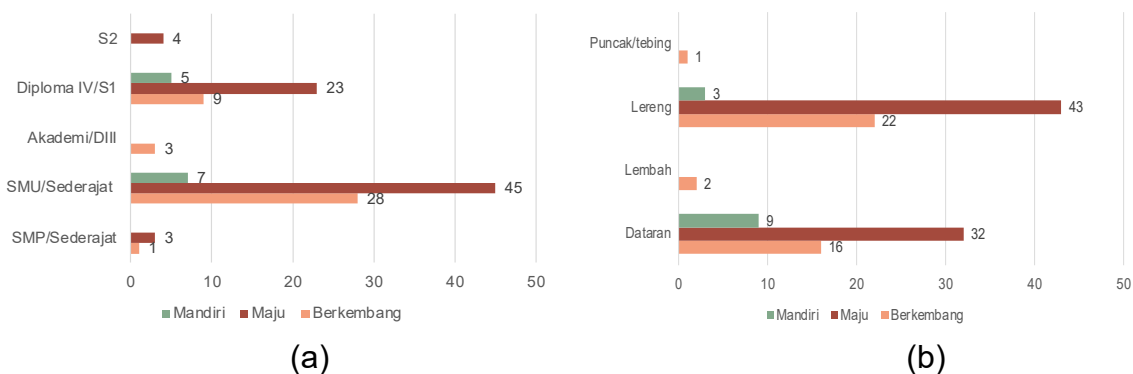
3. Hasil dan Pembahasan

3.1 Deskripsi Data

Data yang digunakan pada penelitian ini berjumlah 128 desa di Kabupaten Buleleng. Peubah respon yaitu status indeks desa membangun (IDM) terdiri atas 12 desa berstatus mandiri, 75 desa berstatus maju, dan 41 desa berstatus berkembang. Gambar 1 menunjukkan persentase status mandiri sebesar 9%, status maju sebesar 59%, dan status berkembang sebesar 32%. Perbedaan persentase status IDM tersebut menunjukkan bahwa data tidak seimbang, dengan status mandiri merupakan kelas minoritas.



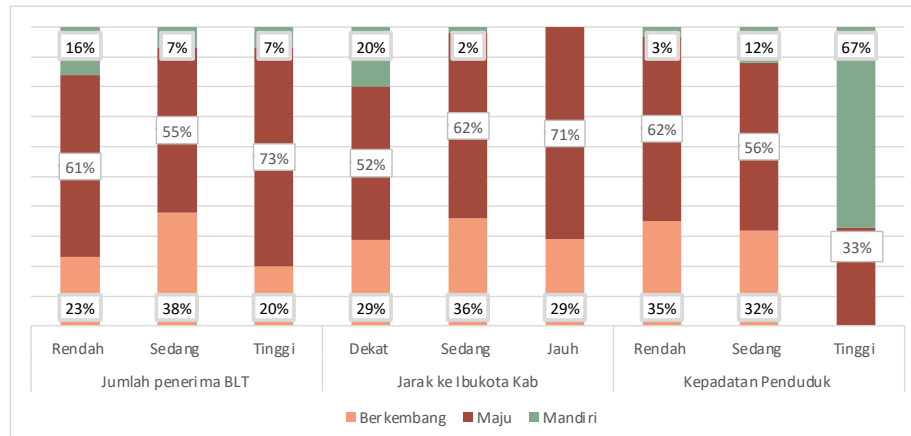
Gambar 1 Persentase status indeks desa membangun di Kabupaten Buleleng



Gambar 2 Persentase status indeks desa membangun di Kabupaten Buleleng berdasarkan (a) pendidikan tertinggi kepala desa dan (b) topografi

Gambar 2 menunjukkan pendidikan tertinggi kepala desa dan topografi desa di Kabupaten Buleleng menurut status desa. Pada bagian (a) dapat dilihat bahwa mayoritas kepala desa di Kabupaten Buleleng telah menyelesaikan pendidikan hingga tingkat SMU/Sederajat dan Diploma IV/S1. Pada bagian (b) menunjukkan bahwa topografi desa di Kabupaten Buleleng sebagian besar terdiri dari lereng dan dataran. Berdasarkan Gambar 3 dapat diketahui bahwa desa dengan status mandiri lebih

banyak dikategorikan ke dalam banyaknya keluarga penerima BLT yang rendah. Peubah jarak ke ibukota kabupaten dengan kategori jauh paling banyak bertatus maju dengan persentase sebesar 71% dan status berkembang sebesar 29%. Desa yang memiliki status indeks desa membangun mandiri tidak ada yang termasuk ke dalam kategori jarak yang jauh dari ibukota kabupaten. Selain itu, dapat diketahui bahwa mayoritas desa dengan kepadatan penduduk yang tinggi memiliki status mandiri, namun tidak ada desa yang memiliki status berkembang.



Gambar 3 Persentase status indeks desa membangun di Kabupaten Buleleng berdasarkan banyaknya keluarga penerima BLT, jarak ke ibukota Kabupaten, dan kepadatan penduduk

3.2 Uji Khi-Kuadrat

Uji khi-kuadrat digunakan untuk menguji apakah terdapat hubungan yang signifikan antara peubah respon dengan peubah penjelas. Peubah penjelas yang paling signifikan terhadap peubah respon akan menjadi peubah pemisah pertama pada pohon keputusan CHAID (Nasution dan Widyaningsih 2022). Tabel 1 menyajikan informasi hasil uji kebebasan khi-kuadrat antara setiap peubah penjelas dan peubah respon. Kepadatan penduduk dan jarak ke ibukota Kabupaten merupakan peubah yang memiliki hubungan signifikan dengan status IDM pada taraf nyata 5%. Kepadatan penduduk adalah peubah penjelas yang paling signifikan serta diduga akan menjadi pemisah simpul pertama pada pohon klasifikasi yang terbentuk.

Tabel 2 Hasil khi-kuadrat setiap peubah penjelas terhadap peubah status IDM

Peubah Penjelas	Khi-kuadrat	Nilai-p
Pendidikan tertinggi kepala desa	11,379	0,181
Jumlah keluarga yang menerima BLT	5,179	0,269
Topografi	11,244	0,081
Jarak ke ibukota Kabupaten	12,768	0,012*
Kepadatan Penduduk	39,852	0,000*

*signifikan pada taraf nyata 5%

3.3 Analisis CHAID Tanpa Menggunakan SMOTE

Analisis CHAID dibentuk menggunakan tiga kategori peubah respon yang tidak seimbang tanpa penanganan data tidak seimbang. Model klasifikasi yang dibentuk menggunakan *10-fold cross validation*. Nilai rata-rata dari hasil *10-fold* diperkirakan

sebagai nilai risiko kesalahan klasifikasi (Wong et al. 2021). Hasil analisis CHAID menunjukkan bahwa terdapat satu peubah penjelas yang berpengaruh terhadap status indeks desa membangun, yaitu kepadatan penduduk. Evaluasi model yang dihasilkan dapat dilihat melalui Tabel 2. Nilai akurasi yang dihasilkan sebesar 60,9%, sensitivitas 60,9%, spesifisitas 80,5%, dan rata-rata nilai risiko kesalahan klasifikasi 10-*fold* sebesar 0,391. Nilai tersebut dapat dikatakan cukup baik dalam melakukan pemodelan, akan tetapi model tidak dapat memprediksi status berkembang.

Tabel 3 Hasil klasifikasi analisis CHAID tanpa SMOTE

Aktual	Prediksi			Persentase Benar (%)
	Mandiri	Maju	Berkembang	
Mandir	6	6	0	50%
Maju	3	72	0	96%
Berkembang	0	41	0	0%
Persentase Keseluruhan (%)	7%	93%	0%	60,9%

3.4 Analisis CHAID dengan Menggunakan SMOTE

Penerapan SMOTE dalam penanganan data tidak seimbang menghasilkan proporsi yang sama untuk ketiga kategori peubah respon dengan jumlah data keseluruhan sebanyak 225 amatan. Hasil evaluasi CHAID dengan data yang sudah seimbang disajikan pada Tabel 3. Nilai ketepatan klasifikasi secara keseluruhan adalah sebesar 57,3%, dengan sensitivitas sebesar 43,4%, spesifisitas sebesar 74,6%, serta rata-rata nilai risiko kesalahan klasifikasi 10-*fold* sebesar 0,484.

Tabel 4 Hasil klasifikasi analisis CHAID dengan SMOTE

Aktual	Prediksi			Persentase Benar (%)
	Mandiri	Maju	Berkembang	
Mandiri	63	1	11	84,0%
Maju	17	44	14	58,7%
Berkembang	11	42	22	29,3%
Persentase Keseluruhan (%)	40,4%	38,7%	20,9%	57,3%

Model CHAID dengan penerapan SMOTE menghasilkan tiga peubah penjelas yang berpengaruh terhadap status indeks desa membangun dengan sepuluh simpul dan enam simpul terminal yang terbentuk. Peubah penjelas yang berpengaruh terhadap peubah respon adalah kepadatan penduduk, jarak ke ibukota kabupaten, dan banyaknya keluarga penerima BLT.

3.5 Analisis CHAID dengan Penggabungan Kategori Peubah Respon

Penggabungan peubah respon kategori mandiri dan maju merupakan salah satu upaya yang dilakukan untuk mengatasi ketidakseimbangan data. Kedua kategori tersebut digabungkan karena memiliki karakteristik yang lebih mirip serta nilai IDM yang dekat. Persentase status mandiri dan maju adalah sebesar 68%, sedangkan status berkembang adalah sebesar 32%. Pada analisis CHAID dengan penggabungan kedua kategori peubah respon menghasilkan hanya kepadatan penduduk yang menjadi penyusun pohon klasifikasi dan tidak ada desa yang diduga memiliki status IDM berkembang. Hal ini dapat disebabkan masih terdapat ketidakseimbangan pada

data setelah penggabungan kategori peubah respon. Ketidakseimbangan dapat memberikan implikasi pada analisis yaitu memengaruhi kinerja dan akurasi model (He dan Garcia 2009). Penerapan SMOTE dilakukan untuk mengatasi hal tersebut. Jumlah data keseluruhan yang terbentuk setelah dilakukan penerapan SMOTE adalah sebanyak 174. Tabel 4 menunjukkan evaluasi hasil analisis CHAID dengan akurasi sebesar 60,3%, sensitivitas sebesar 24,1%, spesifisitas sebesar 96,6% dan rata-rata nilai risiko kesalahan klasifikasi 10-fold model sebesar 0,448. Hasil analisis CHAID menunjukkan bahwa terdapat dua peubah penjelas yang menyusun pohon klasifikasi yaitu kepadatan penduduk dan banyaknya keluarga penerima BLT.

Tabel 5 Hasil klasifikasi analisis CHAID dengan penggabungan kategori peubah respon serta penerapan SMOTE

Aktual	Prediksi		Persentase Benar (%)
	Mandiri dan Maju	Berkembang	
Mandiri dan Maju	21	66	24,1%
Berkembang	3	84	96,6%
Persentase Keseluruhan (%)	13,8%	86,2%	60,3%

3.6 Penentuan Model Terbaik

Model klasifikasi terbaik dipilih berdasarkan kemampuan model dalam memprediksi kelas minoritas maupun mayoritas. Oleh karena itu, pemilihan model terbaik dilakukan dengan membandingkan model analisis CHAID dengan penerapan SMOTE tanpa penggabungan kategori peubah respon dan dengan penggabungan kategori peubah respon. Tabel 5 menunjukkan bahwa model CHAID tanpa penggabungan kategori peubah respon menghasilkan akurasi yang lebih tinggi dibandingkan model dengan penggabungan kategori peubah respon, meskipun perbedaannya tidak begitu signifikan. Nilai sensitivitas yang dihasilkan penggabungan kategori peubah respon lebih kecil dibandingkan dengan model tanpa penggabungan kategori peubah respon. Namun spesifisitas yang dihasilkan penggabungan kategori peubah respon lebih tinggi. Nilai sensitivitas yang tinggi berarti model cenderung tidak melewati desa yang sebenarnya masuk kelas positif, sedangkan nilai spesifisitas yang tinggi berarti model cenderung tidak salah mengklasifikasikan desa yang sebenarnya masuk kelas negatif. Pada penelitian ini diperlukan keseimbangan antara sensitivitas dan spesifisitas. Oleh karena itu, model analisis CHAID dengan penerapan SMOTE tanpa penggabungan kategori peubah respon akan digunakan untuk melihat karakteristik dari status IDM di Kabupaten Buleleng.

Tabel 6 Evaluasi model analisis CHAID dengan penerapan SMOTE

Analisis CHAID	Akurasi	Sensitivitas	Spesifisitas	Rata-rata nilai risiko kesalahan klasifikasi 10-fold
Tanpa penggabungan kategori peubah respon	57,3%	43,4%	74,6%	0,484
Dengan penggabungan	60,3%	24,1%	96,6%	0,448

kategori peubah
respon

3.7 Interpretasi Model

Hasil analisis CHAID menunjukkan bahwa kepadatan penduduk merupakan peubah yang paling signifikan terhadap status indeks desa membangun di Kabupaten Buleleng. Kepadatan penduduk menjadi pemecah simpul pertama, kemudian diikuti oleh jarak ke ibukota kabupaten, dan banyaknya keluarga penerima BLT. Pohon keputusan yang dihasilkan dapat dilihat pada Gambar 4. Pohon analisis klasifikasi CHAID menghasilkan enam simpul terminal. Karakteristik desa yang memiliki status mandiri dengan peluang terbesar dijelaskan oleh tiga simpul, yaitu simpul terminal 1, simpul terminal 4, dan simpul terminal 6. Simpul terminal 6 dengan karakteristik kepadatan penduduk yang lebih tinggi memiliki nilai n yang lebih besar dibandingkan simpul terminal lainnya. Simpul terminal 3 memiliki peluang yang besar untuk menjelaskan karakteristik status maju dan berkembang dengan n yang besar. Simpul terminal 5 memiliki persentase yang sama dalam menjelaskan karakteristik status maju dan berkembang. Klasifikasi status indeks desa membangun beserta karakteristiknya menurut analisis CHAID dapat dilihat pada Tabel 9.

Tabel 7 Karakteristik status IDM hasil analisis CHAID

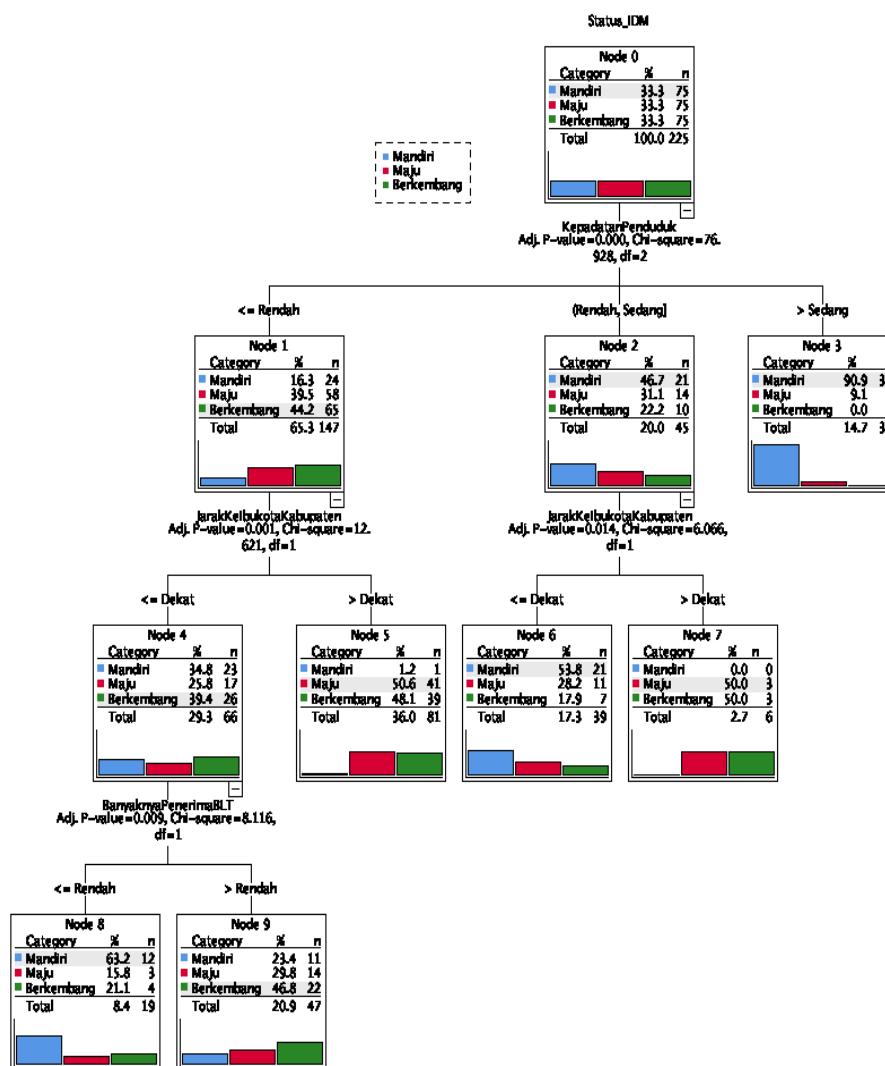
Simpul terminal	Karakteristik	Persentase			n
		Mandiri	Maju	Berkembang	
1	Kepadatan penduduk sangat rendah; jarak ke ibukota kabupaten sangat dekat; banyaknya penerima BLT sangat rendah	63,2%	15,8%	21,1%	19
2	Kepadatan penduduk sangat rendah; jarak ke ibukota kabupaten sangat dekat; banyaknya penerima BLT cukup tinggi	23,4%	29,8%	46,8%	47
3	Kepadatan penduduk sangat rendah; jarak ke ibukota kabupaten cukup jauh	1,2%	50,6%	48,1%	81
4	Kepadatan penduduk berada diantara rendah dan sedang atau cukup tinggi; jarak ke ibukota kabupaten sangat dekat	53,8%	28,2%	17,9%	39
5	Kepadatan penduduk berada diantara rendah dan sedang atau cukup tinggi; jarak ke ibukota kabupaten cukup jauh	0%	50%	50%	6
6	Kepadatan penduduk tinggi	90,9%	9,1%	0%	33

Pohon klasifikasi yang terbentuk mendefinisikan karakteristik status indeks desa membangun di Kabupaten Buleleng. Berdasarkan kepadatan penduduk yang tinggi, desa memiliki kecenderungan yang besar untuk dikategorikan dalam status mandiri. Selain itu, kelompok desa dengan karakteristik jarak yang sangat dekat ke ibukota kabupaten dan banyaknya penerima BLT yang sangat rendah juga cenderung memiliki status mandiri. Desa dengan status maju dan berkembang cenderung memiliki kepadatan penduduk yang sangat rendah dan jarak ke ibukota kabupaten cukup jauh. Kepadatan penduduk yang tinggi dapat mendorong pembangunan di desa tersebut, seperti pembangunan infrastruktur dan pertumbuhan ekonomi. Wahyuni *et al.* (2021) menjelaskan bahwa kepadatan penduduk yang tinggi berpengaruh negatif terhadap kemiskinan disebabkan oleh kepadatan penduduk yang tinggi dapat mendorong aktivitas ekonomi di daerah tersebut serta jangkauan infrastruktur pada daerah tersebut lebih efisien dan biasanya daerah tersebut menjadi pusat ibukota atau berada di sekitar ibukota. Banyaknya keluarga penerima BLT mendeskripsikan penduduk miskin di desa tersebut, banyaknya penerima BLT yang sangat rendah cenderung memiliki kategori mandiri. Hal ini menunjukkan bahwa semakin rendah penduduk miskin di desa tersebut maka pembangunan di desa tersebut semakin baik karena semakin tinggi partisipasi masyarakat dalam pembangunan desa. Hal ini sejalan dengan penelitian Amalia dan Syawie (2015) yang menyatakan bahwa dengan perpaduan antara berbagai kegiatan pemerintahan dan kegiatan partisipasi

masyarakat dalam suatu mekanisme yang baik maka pembangunan desa akan berhasil dalam rangka mempercepat proses pencapaian kemandirian desa.

4. Simpulan

Hasil analisis CHAID dengan tiga kategori peubah respon dengan menggunakan SMOTE tanpa penggabungan kategori peubah respon dapat menghasilkan kinerja yang lebih baik dalam mengklasifikasikan status indeks desa membangun (IDM) di Kabupaten Buleleng. Model CHAID dengan menggunakan SMOTE tanpa penggabungan kategori peubah respon menghasilkan nilai akurasi sebesar 57,3%. Evaluasi kinerja klasifikasi juga memperhatikan keseimbangan antara sensitivitas dan spesifisitas. Hal ini untuk menghindari kesalahan dalam mengklasifikasikan desa agar program lebih tepat sasaran. Penanganan ketidakseimbangan data menggunakan SMOTE pada penelitian ini memberikan hasil yang cukup baik ditunjukkan dengan hasil klasifikasi tidak hanya pada kelas mayoritas. Peubah yang memengaruhi status



Gambar 4 Model Klasifikasi dengan SMOTE tanpa penggabungan kategori peubah respon

indeks desa membangun di Kabupaten Buleleng adalah kepadatan penduduk, jarak ke ibukota kabupaten, dan banyaknya keluarga yang menerima Bantuan Langsung Tunai (BLT). Kepadatan penduduk menjadi karakteristik utama dalam pengklasifikasian status indeks desa membangun di Kabupaten Buleleng.

Daftar Pustaka

- Adinugroho G, Arifa'ilah I, Elvina S, Inriyatni S, Aisyah T. 2016. Pola spasial indeks kesulitan geografis dan pengaruhnya terhadap pembangunan Kabupaten Gunungkidul. *Jurnal Plano Madani*. 5(2): 158-170.
- Amalia AD, Syawie M. 2015. Pembangunan kemandirian desa melalui konsep pemberdayaan: suatu kajian dalam perspektif sosiologi. *Sosio Informa*. 2(1): 175-188.
- Antara IGMY, Suryana IGPE. 2020. Pengaruh tingkat kepadatan penduduk terhadap indeks pembangunan manusia di Provinsi Bali. *Media Komunikasi Geografi*. 21(1): 63-73.
- [Ditjen PDP] Direktorat Jenderal Pembangunan Desa dan Perdesaan. 2021. Peringkat Indeks Desa Membangun Tahun 2021. Jakarta (ID): Direktorat Jenderal Pembangunan Desa dan Perdesaan.
- Faisal M, Nasution YN, Deny F. 2017. Perbandingan kinerja metode klasifikasi chi-square automatic interaction detection (CHAID) dengan metode klasifikasi algoritma C4.5 pada studi kasus: penderita diabetes melitus tipe 2 di Samarinda tahun 2015. *Jurnal Eksponensial*. 8(2):119-124.
- Gallagher AC, Monroe HM, Fish J. 2000. An iterative approach to classification analysis. *Journal of Applied Statistics*. 29 : 256-266.
- He H, Garcia EA. 2009. Learning from imbalanced data. *IEE Transactions on Knowledge and Data Engineering*. 21(9): 1263-1284 doi:10.1109/tkde.2008.239
- Iskandar AH. 2020. SDGs Desa Percepatan Pencapaian Tujuan Pembangunan Nasional Berkelanjutan. Jakarta (ID): Yayasan Pustaka Obor Indonesia.
- [KDPDTT] Kementerian Desa, Pembangunan Daerah Tertinggal dan Transmigrasi. 2015. Indeks Desa Membangun. Jakarta (ID): Kementerian Desa, Pembangunan Daerah Tertinggal dan Transmigrasi.
- Mahayana W. 2013. Peran kepala desa dalam meningkatkan pembangunan desa di desa bumi rapak kecamatan kaubun kabupaten kutai timur. *Jurnal Ilmu Pemerintahan*. 1(1): 400-414.
- Nain U. 2019. Pembangunan Desa dalam Perspektif Sosiohistoris. Makassar (ID): Garis Khatulistiwa.
- Nasution CN, Widyaningsih Y. 2022. Klasifikasi pemilih dalam pemilu 2019 di Indonesia menggunakan regresi logistik multinomial dan Chi-Square Automatic Decision Tree (CHAID). *Jurnal Statistika dan Aplikasinya*. 6(2):148-159.
- Salsabella ET. 2017. Penanganan Masalah Kelas Data Tidak Seimbang pada Pemodelan Risiko HIV di Papua [skripsi]. Bogor (ID): Institut Pertanian Bogor.

- Suantari NGA. 2020. Penerapan SMOTE pada metode CHAID untuk mengklasifikasi tingkat loyalitas pelanggan [skripsi]. Bogor (ID): Institut Pertanian Bogor.
- Suhendra MA, Ispriyanti D, Sudarno. 2020. Ketepatan klasifikasi pemberian kartu keluarga sejahtera di Kota Semarang menggunakan metode regresi logistik biner dan metode CHAID. *Jurnal Gaussian*. 9(1):64-74.
- Sulviana V. 2018. Penggunaan metode CHAID (Chi-Squared Automatic Interaction Detection) pada segmentasi tren penjualan berbagai jenis minuman ringan di Indonesia [skripsi]. Bogor (ID): Institut Pertanian Bogor.
- Wahyuni AY, Juanda B, Purnamadewi YL. 2021. Analisis pengaruh alokasi DAK masing-masing bidang terhadap tingkat kemiskinan di Kabupaten Pandeglang dan Kabupaten Lebak. *Jurnal Wilayah dan Lingkungan*. 9(1): 1-17 doi:10.14710/jwl.9.1.1-17.
- Wong M, Nikolaos T, Federica N, Tejal A, Davor J. 2021. Risk of Pre-Malignancy or Malignancy in Postmenopausal Endometrial Polyps: A CHAID Decision Tree Analysis. *Diagnostics*. <https://doi.org/10.3390/diagnostics11061094>