

PERBANDINGAN ORDINAL FOREST DAN REGRESI LOGISTIK ORDINAL

Studi Kasus: Indeks Desa Membangun Provinsi Maluku Utara 2020

M. Yunus^{1,2}, Khairil Anwar Notodiputro¹, Bagus Sartono¹

¹IPB University, Bogor, Indonesia

²UIN Sulthan Thaha Saifuddin Jambi, Jambi, Indonesia

e-mail: muhammadyunus@apps.ipb.ac.id

Abstrak

Terdapat banyak metode yang digunakan untuk mengidentifikasi prediktor penting dan memprediksi nilai dari peubah respons ordinal. Namun, terdapat metode prediksi untuk peubah respons ordinal yang tidak menggunakan sifat ordinal tersebut. Penelitian ini menggunakan metode ordinal forest dan metode regresi logistik ordinal. Nilai akurasi dan kappa metode ordinal forest pada setiap ulangan (5 ulangan) selalu lebih besar dari pada regresi logistik ordinal. Selanjutnya, nilai akurasi dan kappa setiap kelompok berdasarkan PDRB pada metode ordinal forest selalu lebih besar dari pada regresi logistik ordinal. Sehingga metode ordinal forest lebih baik digunakan pada data peringkat status indeks desa membangun Provinsi Maluku Utara 2020.

Kata kunci: indeks desa membangun, ordinal forest, peubah respons ordinal, regresi logistik ordinal

Abstract

There are many methods used to identify the important predictors and predict the value of the ordinal response variable. However, there are prediction methods for ordinal response variables that do not use the ordinal nature. This study uses the ordinal forest method and the ordinal logistic regression method. The accuracy and kappa value of the ordinal forest method in each replication (5 replications) is always greater than the ordinal logistic regression. Furthermore, the value of accuracy and kappa of each group based on GRDP in the ordinal forest method is always greater than the ordinal logistic regression. So that the ordinal forest method is better used in data on village development index in North Maluku Province 2020.

Keywords: *ordinal forest, ordinal logistic regression, ordinal response variable, village development index*

PENDAHULUAN

Terdapat banyak metode yang digunakan untuk mengidentifikasi prediktor penting dan memprediksi nilai dari peubah respons ordinal. Namun saat ini, terdapat metode prediksi untuk peubah respons ordinal yang tidak menggunakan sifat ordinal dari peubah respons tersebut. Peubah respons ordinal sering diperlakukan sebagai peubah multinomial ataupun nominal, menerapkan metode prediksi untuk peubah respons biner ke semua kombinasi pasangan kelas dari peubah respons ordinal. Salah satu metode statistika yang sering digunakan untuk memprediksi nilai dari peubah respons ordinal adalah regresi logistik ordinal (Atinafu et al., 2023; Baek et al., 2020; Moraes et al., 2019). Pendugaan parameter regresi logistik memakai metode kemungkinan maksimum (*maximum likelihood*) (Hosmer et al., 2013). Metode kemungkinan maksimum dapat diterapkan apabila antar amatan yang satu dengan amatan yang lain diasumsikan saling bebas. Pohon keputusan ordinal terbukti efektif untuk masalah klasifikasi di mana peubah kelas menunjukkan beberapa bentuk urutan ordinal (Singer et al., 2020).

Random forest klasik yang dikenalkan oleh (Breiman, 2001) belum menerapkan sifat ordinal pada peubah respons. Pada penelitian ini, diperkenalkan metode Ordinal Forest (OF). Ordinal forest merupakan metode prediksi inovatif untuk peubah respons ordinal yang berlaku untuk kovariat dimensi rendah dan dimensi tinggi. Metode OF menggunakan sifat ordinal dari peubah ordinal tersebut. Analog dengan kelas umum model regresi untuk peubah ordinal yang dikembangkan oleh (McCullagh, 1980) yang termasuk regresi probit terurut. Metode OF didasarkan pada gagasan tentang peubah respons kontinu yang mendasari peubah respon ordinal yang diamati (Hornung, 2019).

Ordinal forest cenderung mengungguli metode lain (*multi-class random forest, regression forest, dan naive ordinal forests*) dalam hal kinerja prediksi

dan terkait mengidentifikasi prediktor penting relatif sama baiknya dengan metode lain (Hornung, 2019). Namun, menurut (Nisa et al., 2021), metode terbaik adalah metode regresi logistik ordinal karena menghasilkan nilai akurasi yang lebih tinggi serta nilai sensitivitas dan spesifisitas yang lebih seimbang di setiap kategori IPK dibandingkan metode random forest ordinal. Oleh sebab itu, penelitian ini mengidentifikasi prediktor penting dan memprediksi nilai peubah respons ordinal dengan menggunakan metode ordinal forest dan sebagai pembanding digunakan juga metode regresi logistik ordinal.

Tujuan Pembangunan Berkelanjutan (TPB)/*Sustainable Development Goals* (SDGs) adalah pembangunan yang menjaga peningkatan kesejahteraan ekonomi, sosial, dan lingkungan masyarakat secara berkesinambungan (Kementerian PPN/Bappenas, 2020). Percepatan pencapaian tujuan SDGs secara nasional dapat didukung oleh formula SDGs Desa (Kurniawan et al., 2022). SDGs Desa dapat mewujudkan desa tanpa kemiskinan dan kelaparan, ekonomi tumbuh merata, peduli kesehatan, lingkungan, pendidikan, ramah perempuan, desa berjejaring, dan desa tanggap budaya (Permendes Nomor 13, 2020). Kementerian Desa, Pembangunan Daerah Tertinggal dan Transmigrasi Republik Indonesia (Kemendes PDTT RI) telah mengeluarkan Indeks Desa Membangun (IDM) untuk memberikan informasi dan status kemajuan desa guna mendukung pembangunan desa. IDM memiliki nilai pada interval (0,1) yang meliputi aspek sosial, ekonomi dan lingkungan (Astika & Subawa, 2021). Ilustrasi dalam penelitian ini menggunakan data peringkat status IDM Provinsi Maluku Utara tahun 2020.

METODOLOGI

1. Data

Pembangunan Daerah Tertinggal (PDT) merupakan suatu proses, upaya, dan tindakan secara terencana untuk meningkatkan kualitas masyarakat dan

wilayah yang merupakan bagian integral dari pembangunan nasional (Wahyuningrum, 2020). Pembangunan desa merupakan salah satu target penting dari kebijakan pemerintah di Indonesia (Utami & Wijayanto, 2022). Data penelitian yang dipakai merupakan data Peringkat Status Indeks Desa Membangun (IDM) Provinsi Maluku Utara Tahun 2020. Nilai Indeks Pembangunan Lingkungan di Provinsi Maluku Utara secara umum memenuhi klasifikasi rendah, sedang, dan tinggi. Nilai indeks tertinggi dimiliki oleh Kota Ternate (0,43) sedangkan nilai indeks terendah dimiliki oleh Kabupaten Halmahera Selatan (0,20) (Alfana et al., 2020). Data yang digunakan terdiri dari 1.199 observasi berupa desa/kelurahan dari 117 kecamatan di 10 kabupaten/kota pada Provinsi Maluku Utara. Terdapat 163 peubah di dalam data tersebut, peubah

Status IDM (Madjid et al., 2020) menjadi peubah respons dan sisanya merupakan peubah penjelas. Penelitian ini juga menggunakan data laju pertumbuhan Produk Domestik Regional Bruto sebagai dasar pengelompokan kategori laju pertumbuhan minus, rendah, sedang dan tinggi. Data diperoleh dari Badan Pusat Statistika (BPS) Provinsi Maluku Utara.

Menurut (Kemendes PD TT RI, 2020) klasifikasi/kategori Status Desa ditetapkan dengan ambang batas sebagai berikut:

1. Desa Sangat Tertinggal: $IDM \leq 0,4907$
2. Desa Tertinggal: $0,4907 < IDM \leq 0,5989$
3. Desa Berkembang: $0,5989 < IDM \leq 0,7072$
4. Desa Maju: $0,7072 < IDM \leq 0,8155$
5. Desa Mandiri: $IDM > 0,8155$

Tabel 1. Peubah Respons dan Peubah Penjelas

No.	Peubah Respons	Kategori	Skala
1.	Status Indeks Desa Membangun (IDM)	Sangat Tertinggal Tertinggal Berkembang Maju	Ordinal

No.	Peubah Penjelas	Kategori	Skala
1.	Jumlah Keluarga Pengguna Listrik		Rasio
2.	Penerangan di Jalan Utama Desa/Kelurahan	Ada, sebagian besar Ada, sebagian kecil Tidak ada	Ordinal
3.	Tempat Buang Sampah Sementara	Ada, digunakan Ada, tidak digunakan Tidak ada	Ordinal
4.	Keberadaan Bank Sampah	Ada Tidak ada	Nominal
5.	Keberadaan Sungai di Desa/Kelurahan	Ada Tidak ada	Nominal
6.	Jumlah Warga Penderita Gizi Buruk Tahun 2019		Rasio
7.	Keberadaan Angkutan Umum	Ada, dengan trayek tetap Ada, tanpa trayek tetap Tidak ada angkutan umum	Ordinal
⋮	⋮	⋮	⋮
160.	Jumlah Industri Barang dari Kayu		Rasio
161.	Jumlah Masjid		Rasio
162.	Kejadian/Bencana Alam (Banjir)	Ada Tidak ada	Nominal

Sumber: <https://idm.kemendes.go.id/view/detil/3/publikasi>

Tabel 1. Laju Pertumbuhan Produk Domestik Regional Bruto (PDRB) (Persen) Tahun 2020

No.	Wilayah	Laju Pertumbuhan	Kategori
1.	Halmahera Barat	0,62	Rendah
2.	Halmahera Tengah	26,34	Tinggi
3.	Kepulauan Sula	0,41	Rendah
4.	Halmahera Selatan	16,22	Tinggi
5.	Halmahera Utara	-0,15	Minus
6.	Halmahera Timur	0,57	Rendah
7.	Pulau Morotai	2,68	Sedang
8.	Pulau Taliabu	1,06	Sedang
9.	Ternate	-0,85	Minus
10.	Tidore Kepulauan	1,99	Sedang
	Maluku Utara	4,92	

Sumber: <https://malut.bps.go.id>

Tabel 1. menunjukkan peubah respons dan peubah penjelas yang digunakan dalam penelitian ini. Tabel 2. menyajikan laju Pertumbuhan Produk Domestik Regional Bruto (PDRB) Menurut kabupaten/kota Provinsi Maluku Utara Atas Dasar Harga Konstan (Persen) Tahun 2020.

Metode Penelitian

Tahapan-tahapan analisis data menggunakan perangkat lunak R dengan paket ordinalForest, VGAM, dan caret sebagai berikut:

1. Eksplorasi Data
2. Analisis Ordinal Forest (OF)
Analisis dengan menggunakan ordinal forest, algoritma OF diimplementasikan pada paket R ordinalForest yang tersedia di CRAN dalam versi 2.4-2 (Hornung, 2021). Keakuratan tes diagnostik sering dinilai oleh dua probabilitas bersyarat yaitu sensitivitas dan spesifisitas (Agresti, 2019; Wulandari, 2015). Penilaian diagnostik Y memiliki skala ordinal, sebagai berikut:
spesifisitas = $P(Y \leq j | x=0)$,
sensitivitas = $P(Y > j | x=1)$
3. Analisis Regresi Logistik Ordinal
Persamaan (1) merupakan persamaan analisis dengan menggunakan regresi logistik ordinal (Agresti, 2010) diberikan:

$$P(Y \leq j | \mathbf{x}) = \pi_1(\mathbf{x}) + \dots + \pi_j(\mathbf{x}),$$

$$j = 1, \dots, J.$$

$$\text{logit} [P(Y \leq j | \mathbf{x})] = \log \frac{P(Y \leq j | \mathbf{x})}{1 - P(Y \leq j | \mathbf{x})}$$

$$\text{logit} [P(Y \leq j | \mathbf{x})]$$

$$= \log \frac{\pi_1(\mathbf{x}) + \dots + \pi_j(\mathbf{x})}{\pi_{j+1}(\mathbf{x}) + \dots + \pi_J(\mathbf{x})} \quad (1)$$

$$j = 1, \dots, J - 1$$

selanjutnya, juga dilakukan tes diagnostik yaitu sensitivitas dan spesifisitas.

4. Perbandingan Analisis Ordinal Forest dan Regresi Logistik Ordinal
Perbandingan Analisis Ordinal Forest dan Regresi Logistik Ordinal menggunakan nilai akurasi dan kappa. Berikut ini perhitungan nilai kappa (Cohen, 1968) pada persamaan (2), diberikan:

$$K_w = \frac{\sum_{i=1}^J \sum_{j=1}^J w_{ij} p_{oij} - \sum_{i=1}^J \sum_{j=1}^J w_{ij} p_{cij}}{1 - \sum_{i=1}^J \sum_{j=1}^J w_{ij} p_{cij}} \quad (2)$$

5. Ukuran Kepentingan Peubah
Ukuran yang cocok digunakan dalam (Janitza et al., 2016) untuk menilai peringkat peubah dalam hal ini diperoleh dengan VIM (Variabel Importance Measures) pada persamaan (3). Ukuran

kepentingan peubah untuk kovariat j diberikan sebagai berikut:

$$VIM_j^M = \frac{1}{ntree} \sum_{t=1}^{ntree} (MP_{tj} - M_{tj}), \quad (3)$$

$ntree$: Jumlah pohon di hutan,

M_{tj} : Kesalahan pohon t sebelum mengubah nilai variabel prediktor X_j ketika memprediksi observasi.

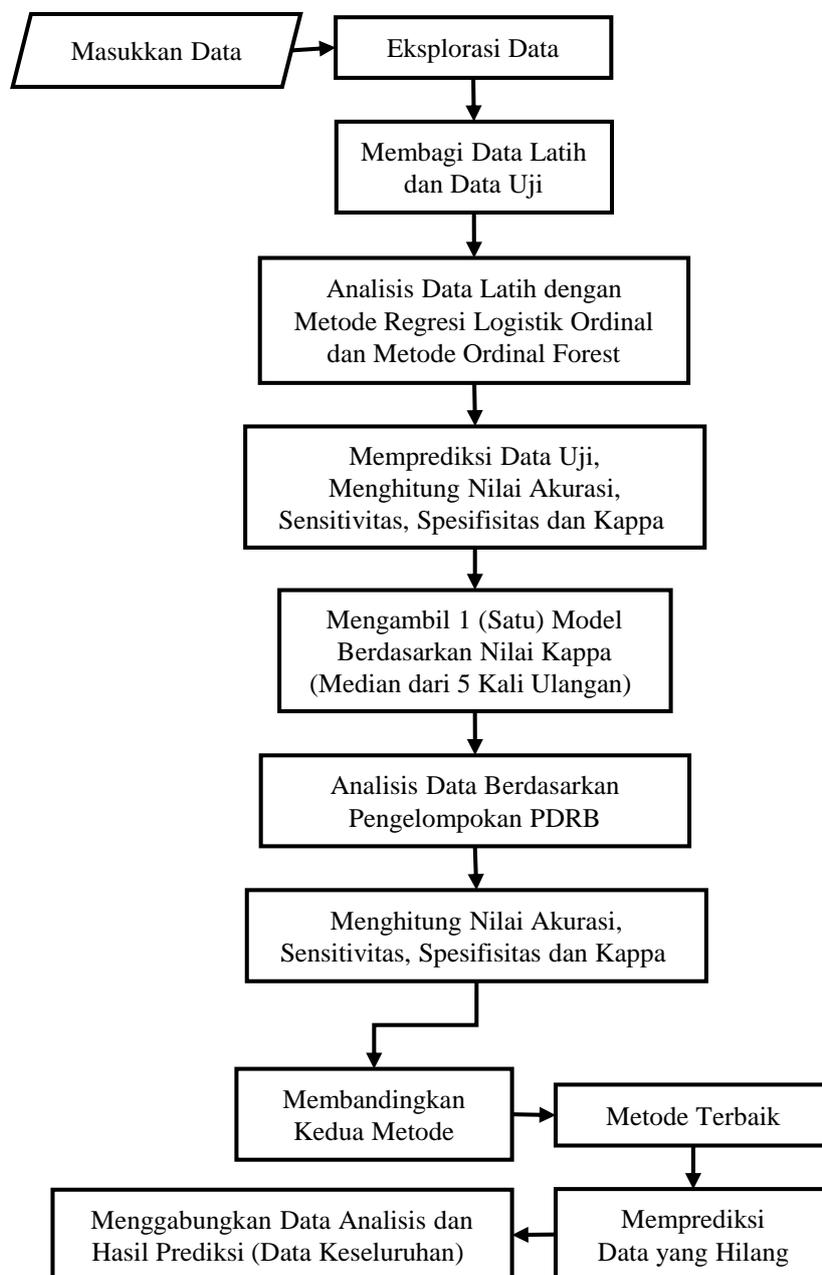
MP_{tj} : Kesalahan pohon t sesudah mengubah nilai variabel prediktor X_j ketika memprediksi observasi.

6. Prediksi Status IDM yang Mengalami Missing Value (Nilai yang Hilang)

Prediksi terhadap status IDM yang mengalami Missing Value (Nilai yang Hilang) pada data Peringkat Status Indeks Desa Membangun (IDM) Provinsi Maluku Utara Tahun 2020.

2. Prosedur Analisis Data

Prosedur analisis data dalam penelitian ini, disajikan pada Gambar 1. Diagram Alir Prosedur Analisis Data, sebagai berikut:



Gambar 1. Diagram Alir Prosedur Analisis Data

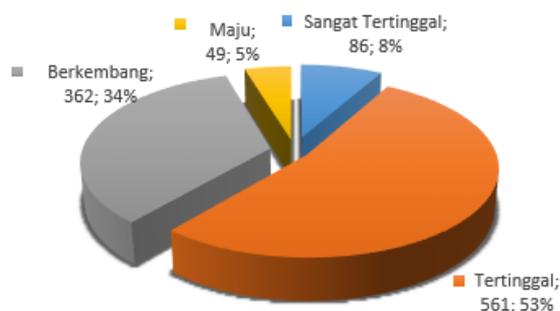
Tabel 3. Kategori Sebenarnya Dan Hasil Prediksi Analisis OF

Prediksi	Kategori Sebenarnya			
	Sangat Tertinggal	Tertinggal	Berkembang	Maju
Sangat Tertinggal	4	1	0	0
Tertinggal	12	96	24	0
Berkembang	0	14	50	3
Maju	0	0	1	7
Total	16	111	75	10

HASIL DAN PEMBAHASAN

1. Eksplorasi Data

Data peringkat status Indeks Desa Membangun (IDM) Provinsi Maluku Utara Tahun 2020. Data terdiri dari 1.199 observasi (desa/kelurahan), namun terdapat 141 desa/kelurahan mengalami *missing value* (nilai yang hilang) terkait dengan peubah respons Status IDM, sehingga tidak diikutsertakan ke dalam analisis. Sisa desa/kelurahan yang digunakan dalam penelitian ini adalah 1.058 desa/kelurahan. Selanjutnya, nilai yang hilang dari peubah Status IDM akan diprediksi menggunakan metode terbaik yang diperoleh pada penelitian ini. Gambar 2 menunjukkan bahwa sebagian besar status IDM di Provinsi Maluku Utara tahun 2020 dikategorikan sebagai desa tertinggal yaitu sebesar 53% (561 observasi).



Gambar 2. Sebaran Status IDM di 1.058 Desa/Kelurahan

2. Analisis Ordinal Forest

Analisis ordinal forest (OF) menggunakan data latih sebesar 80% (846 desa/kelurahan) dan data uji sebesar 20% (212 desa/kelurahan). Proses analisis tersebut dilakukan sebanyak 5 kali ulangan

dengan data latih dan data uji yang berbeda dan dengan persentase yang sama. Diperoleh nilai kappa dari setiap ulangan dan diambil median dari 5 nilai kappa tersebut. Tabel 3. menunjukkan kategori sebenarnya dan hasil prediksi menggunakan analisis ordinal forest (OF) dengan nilai kappa sebesar 0,5356.

Berdasarkan Tabel 3. menunjukkan bahwa sebanyak 157 desa/kelurahan diprediksi benar oleh metode tersebut. Terdapat 55 desa/kelurahan yang diprediksi tidak sesuai dengan kategori sebenarnya. Selanjutnya, dari 55 desa/kelurahan tersebut, terdapat 27 desa/kelurahan yang diprediksi lebih tinggi dari kategori yang sebenarnya dan 28 desa/kelurahan yang diprediksi lebih rendah dari kategori sebenarnya. Tabel 4. menyajikan nilai evaluasi metode ordinal forest berupa nilai akurasi, nilai sensitivitas dan nilai spesifisitas setiap kategori status IDM.

Berdasarkan Tabel 4. dapat dilihat bahwa nilai akurasi menggunakan metode logistik ordinal sebesar 74,06% artinya sebanyak 157 desa/kelurahan dari 212 desa/kelurahan diprediksi sesuai dengan kategorinya, sisanya 55 desa/kelurahan diprediksi tidak sesuai dengan kategorinya. Nilai sensitivitas untuk kategori sangat tertinggal sebesar 25,00% artinya sebanyak 4 desa/kelurahan dari 16 desa/kelurahan diprediksi sesuai dengan kategorinya, sisanya 12 desa/kelurahan diprediksi tidak sesuai dengan kategorinya dan nilai spesifisitas sebesar 99,49%. Nilai sensitivitas untuk kategori tertinggal sebesar 86,49% artinya sebanyak 96 desa/kelurahan dari 111 desa/kelurahan diprediksi sesuai dengan kategorinya, sisanya 15 desa/kelurahan diprediksi tidak

Tabel 4. Nilai Evaluasi Metode Ordinal Forest (%)

Nilai	Kategori Status IDM	Status IDM
Akurasi		74,06
Sensitivitas	Sangat Tertinggal	25,00
	Tertinggal	86,49
	Berkembang	66,67
	Maju	70,00
Spesifisitas	Sangat Tertinggal	99,49
	Tertinggal	64,36
	Berkembang	87,59
	Maju	99,51

sesuai dengan kategorinya dan nilai spesifisitas sebesar 64,36%. Nilai sensitivitas untuk kategori berkembang sebesar 66,67% artinya sebanyak 50 desa/kelurahan dari 75 desa/kelurahan diprediksi sesuai dengan kategorinya, sisanya 25 desa/kelurahan diprediksi tidak sesuai dengan kategorinya dan nilai spesifisitas sebesar 87,59%. Nilai sensitivitas untuk kategori maju sebesar 70,00% artinya sebanyak 7 desa/kelurahan dari 10 desa/kelurahan diprediksi sesuai dengan kategorinya, sisanya 3 desa/kelurahan diprediksi tidak sesuai dengan kategorinya dan nilai spesifisitas sebesar 99,51%.

3. Analisis Regresi Logistik Ordinal

Analisis regresi logistik ordinal menggunakan data latih sebesar 80% (846 desa/kelurahan) dan data uji sebesar 20% (212 desa/kelurahan). Proses analisis tersebut dilakukan sebanyak 5 kali ulangan dengan data latih dan data uji yang berbeda dan dengan persentase yang sama. Diperoleh nilai kappa dari setiap ulangan dan diambil median dari 5 nilai kappa tersebut. Tabel 5. menunjukkan kategori sebenarnya dan hasil prediksi

menggunakan analisis regresi logistik ordinal dengan nilai kappa sebesar 0,3922.

Berdasarkan Tabel 5. menunjukkan bahwa sebanyak 137 desa/kelurahan diprediksi benar oleh metode tersebut. Terdapat 75 desa/kelurahan yang diprediksi tidak sesuai dengan kategori sebenarnya. Selanjutnya, dari 75 desa/kelurahan tersebut, terdapat 37 desa/kelurahan yang diprediksi lebih tinggi dari kategori yang sebenarnya dan 38 desa/kelurahan yang diprediksi lebih rendah dari kategori sebenarnya. Tabel 6. menyajikan nilai evaluasi metode regresi logistik ordinal berupa nilai akurasi, nilai sensitivitas dan nilai spesifisitas setiap kategori status IDM.

Berdasarkan Tabel 6. dapat dilihat bahwa nilai akurasi menggunakan metode logistik ordinal sebesar 64,62% artinya sebanyak 137 desa/kelurahan dari 212 desa/kelurahan diprediksi sesuai dengan kategorinya, sisanya 75 desa/kelurahan diprediksi tidak sesuai dengan kategorinya. Nilai sensitivitas untuk kategori sangat tertinggal sebesar 31,25% artinya sebanyak

Tabel 5. Kategori Sebenarnya Dan Hasil Prediksi Analisis Regresi Logistik Ordinal

Prediksi	Kategori Sebenarnya			
	Sangat Tertinggal	Tertinggal	Berkembang	Maju
Sangat Tertinggal	5	5	2	0
Tertinggal	11	85	29	0
Berkembang	0	20	39	2
Maju	0	1	5	8
Total	16	111	75	10

Tabel 6. Nilai Evaluasi Metode Regresi Logistik Ordinal (%)

Nilai	Kategori Status IDM	Status IDM
Akurasi		64,62
Sensitivitas	Sangat Tertinggal	31,25
	Tertinggal	76,58
	Berkembang	52,00
	Maju	80,00
Spesifisitas	Sangat Tertinggal	96,43
	Tertinggal	60,40
	Berkembang	83,94
	Maju	97,03

5 desa/kelurahan dari 16 desa/kelurahan diprediksi sesuai dengan kategorinya, sisanya 11 desa/kelurahan diprediksi tidak sesuai dengan kategorinya dan nilai spesifisitas sebesar 96,43%. Nilai sensitivitas untuk kategori tertinggal sebesar 76,58% artinya sebanyak 85 desa/kelurahan dari 111 desa/kelurahan diprediksi sesuai dengan kategorinya, sisanya 26 desa/kelurahan diprediksi tidak sesuai dengan kategorinya dan nilai spesifisitas sebesar 60,40%. Nilai sensitivitas untuk kategori berkembang sebesar 52,00% artinya sebanyak 39 desa/kelurahan dari 75 desa/kelurahan diprediksi sesuai dengan kategorinya, sisanya 36 desa/kelurahan diprediksi tidak sesuai dengan kategorinya dan nilai spesifisitas sebesar 83,94%. Nilai sensitivitas untuk kategori maju sebesar 80,00% artinya sebanyak 8 desa/kelurahan dari 10 desa/kelurahan diprediksi sesuai dengan kategorinya, sisanya 2 desa/kelurahan diprediksi tidak sesuai dengan kategorinya dan nilai spesifisitas sebesar 97,03%.

4. Perbandingan Ordinal Forest dan Regresi Logistik Ordinal

Pemodelan terbaik dipilih berdasarkan nilai akurasi dan kappa.

Metode yang mempunyai nilai akurasi dan kappa terbesar merupakan metode yang terbaik. Tabel 7. menyajikan perbandingan ordinal forest dan regresi logistik ordinal. Berdasarkan Tabel 7. bahwa nilai akurasi dan kappa setiap ulangan dari metode ordinal forest lebih besar dari pada regresi logistik ordinal. Didapatkan metode ordinal forest lebih baik digunakan pada data peringkat status indeks desa membangun Provinsi Maluku Utara tahun 2020. Model yang diambil adalah 1 (satu) model dengan set.seed (137) berdasarkan nilai kappa (median dari 5 kali ulangan). Selanjutnya, dilakukan analisis data Status Indeks Desa Membangun Provinsi Maluku Utara Tahun 2020 yang telah dikelompokkan berdasarkan laju pertumbuhan PDRB. Tabel 8. menyajikan perbandingan regresi logistik ordinal dan ordinal forest pada data pengelompokan berdasarkan PDRB tahun 2020.

Berdasarkan Tabel 8. bahwa nilai akurasi dan kappa setiap kelompok pada metode ordinal forest selalu lebih besar dari pada regresi logistik ordinal. Sehingga didapatkan metode ordinal forest lebih baik digunakan pada data pengelompokan berdasarkan PDRB di status indeks desa membangun Provinsi Maluku Utara tahun 2020.

Tabel 7. Perbandingan Ordinal Forest (OF) dan Regresi Logistik Ordinal (RLO)

Metode	Nilai	set.seed (17)	set.seed (27)	set.seed (67)	set.seed (137)	set.seed (257)	Media
OF	Akurasi	0,7500	0,7547	0,7028	0,7406	0,7311	0,7406
	Kappa	0,5415	0,5414	0,4565	0,5356	0,5128	0,5356
RLO	Akurasi	0,6981	0,6557	0,6462	0,6462	0,6934	0,6557
	Kappa	0,4642	0,3833	0,3695	0,3922	0,4614	0,3922

Tabel 8. Perbandingan Ordinal Forest (OF) dan Regresi Logistik Ordinal (RLO) pada Data Pengelompokan Berdasarkan PDRB Tahun 2020

Metode	Nilai	Laju Pertumbuhan PDRB			
		Minus	Rendah	Sedang	Tinggi
OF	Akurasi	0,9490	0,9568	0,9569	0,9314
	Kappa	0,9143	0,9252	0,9233	0,8621
RLO	Akurasi	0,7398	0,7666	0,7703	0,7582
	Kappa	0,5397	0,5856	0,5863	0,4849

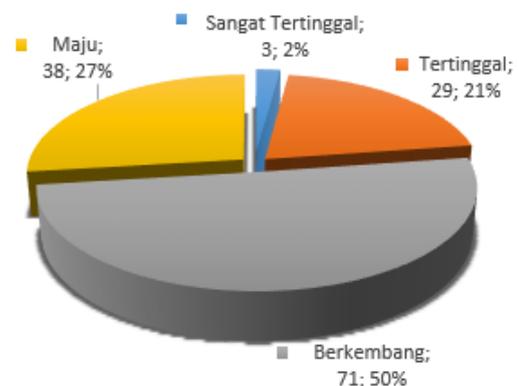
5. Ukuran Kepentingan Peubah

Ukuran kepentingan peubah penjelas yang digunakan pada penelitian ini adalah VIM (*Variable Importance Measures*). Tingkat kepentingan 7 peubah penjelas numerik yang lebih tinggi dibandingkan dengan peubah penjelas lainnya adalah jarak restoran/rumah makan, jarak warung/kedai makanan minuman, jarak tempat praktik dokter, jarak kantor bupati/walikota, jarak minimarket/swalayan, jumlah keluarga pengguna listrik dan jarak apotek. Selanjutnya, tingkat kepentingan 7 peubah penjelas kategorik yang lebih tinggi dibandingkan dengan peubah penjelas lainnya adalah tempat buang sampah keluarga, sumber air untuk mandi/cuci, lalu lintas (transport) dari/ke desa/kelurahan, sumber air untuk minum, keberadaan angkutan umum, bahan bakar untuk memasak dan sinyal telepon seluler/handphone. Hal ini sesuai dengan penelitian yang dilakukan (Prasetya et al., 2019) bahwa nilai Indeks Jaminan Ekonomi dan Sosial cenderung tinggi.

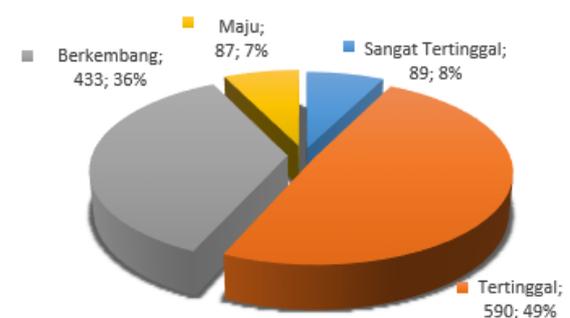
6. Prediksi Status IDM yang Mengalami Missing Value (Nilai yang Hilang)

Terdapat 141 desa/kelurahan mengalami *missing value* (nilai yang hilang) terkait dengan peubah respons Status IDM, sehingga dilakukan prediksi dengan metode terbaik yaitu ordinal forest. Gambar 3. menunjukkan sebaran hasil prediksi status IDM di 141 desa/kelurahan. Sebagian besar desa/kelurahan yang mengalami nilai yang hilang diprediksi sebagai desa yang berstatus berkembang yaitu sebesar 71 desa/kelurahan (50%). Hanya ada 3 desa/kelurahan (2%) yang

diprediksi sebagai desa yang berstatus sangat tertinggal. Gambar 4. menunjukkan sebaran status IDM di 1.199 desa/kelurahan pada Provinsi Maluku Utara tahun 2020 (secara keseluruhan termasuk hasil prediksi).



Gambar 3. Sebaran Status IDM di 141 Desa/Kelurahan



Gambar 4. Sebaran Status IDM di 1.199 Desa/Kelurahan

SIMPULAN

Nilai akurasi dan kappa metode ordinal forest pada setiap ulangan (5 ulangan) selalu lebih besar dari pada regresi logistik ordinal. Selanjutnya, nilai akurasi dan kappa setiap kelompok berdasarkan PDRB pada metode ordinal forest selalu

lebih besar dari pada regresi logistik ordinal. Sehingga disimpulkan bahwa metode ordinal forest lebih baik digunakan pada data peringkat status indeks desa membangun Provinsi Maluku Utara tahun 2020. Tingkat kepentingan peubah penjelas numerik dan kategorik yang tertinggi dibandingkan dengan peubah penjelas lainnya adalah jarak restoran/rumah makan dan tempat buang sampah keluarga. Hasil prediksi nilai yang hilang sebanyak 141 desa/kelurahan dengan metode ordinal forest, sebagian besar desa/kelurahan yang mengalami nilai yang hilang diprediksi sebagai desa yang berstatus berkembang yaitu sebesar 71 desa/kelurahan (50%). Hanya ada 3 desa/kelurahan (2%) yang diprediksi sebagai desa yang berstatus sangat tertinggal.

DAFTAR PUSTAKA

- Agresti, Alan. (2010). *Analysis of Ordinal Categorical Data* (Second Edition). John Wiley and Sons.
- Agresti, Alan. (2019). An Introduction to Categorical Data Analysis. In *Wiley Series in Probability and Statistics* (Third Edition, Vol. 170, Issue 4). https://doi.org/10.1111/j.1467-985x.2007.00506_2.x
- Alfana, A. M. F., Kusmanto, Y., & Pitoyo, A. J. (2020). *Analisis Indeks Pembangunan Lingkungan di Provinsi Maluku Utara Analisis Sumberdaya Manusia View project do an exercise View project*. <https://www.researchgate.net/publication/344778695>
- Astika, A. N., & Subawa, N. S. (2021). Evaluasi Pembangunan Desa Berdasarkan Indeks Desa Membangun. *Jurnal Ilmiah Muqoddimah : Jurnal Ilmu Sosial, Politik Dan Humaniora*, 5(2), 223–232. <http://jurnal.um-tapsel.ac.id/index.php/muqoddimah>
- Atinafu, B. T., Tarekegn, F. N., & Kebede, W. M. (2023). Assessment of the Level of Social Support and Associated Factors Among Cancer Patients in the Comprehensive Cancer Center at Ethiopia: Ordinal Logistic Regression Analysis Level of Social Support and Associated Factors Among Cancer Patients. *Heliyon*, 9(5). <https://doi.org/10.1016/j.heliyon.2023.e15688>
- Baek, W., Jang, Y., Park, C. G., & Moon, M. (2020). Factors Influencing Satisfaction with Patient-Controlled Analgesia Among Postoperative Patients Using a Generalized Ordinal Logistic Regression Model. *Asian Nursing Research*, 14(2), 73–81. <https://doi.org/10.1016/j.anr.2020.03.001>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5–32.
- Cohen, J. (1968). Weighed Kappa: Nominal Scale Agreement with Provision for Scaled Disagreement or Partial Credit. *Psychological Bulletin*, 70(4), 213–220.
- Hornung, R. (2019). Ordinal Forests. *Journal of Classification*, 37(1), 4–17. <https://doi.org/10.1007/s00357-018-9302-x>
- Hornung, R. (2021). Ordinal Forests: Prediction and Variable Ranking with Ordinal Target Variables. In *R package version 2.4-2*. Springer. <https://doi.org/10.1007/s00357-018-9302-x>
- Hosmer, D. W., Lemeshow, Stanley., & Sturdivant, R. X. (2013). *Applied Logistic Regression* (Third Edition). New Jersey (US): John Wiley and Sons.
- Janitza, S., Tutz, G., & Boulesteix, A. L. (2016). Random Forest for Ordinal Responses: Prediction and Variable Selection. *Computational Statistics and Data Analysis*, 96, 57–73. <https://doi.org/10.1016/j.csda.2015.10.005>
- Kemendes PDTT RI. (2020). *Standar Operasional Prosedur (SOP) Update Data Indeks Desa Membangun Tahun 2020*. Direktorat Jenderal Pembangunan dan Pemberdayaan Masyarakat Desa: Kementerian Desa, Pembangunan Daerah Tertinggal dan Transmigrasi, Republik Indonesia.
- Kementerian PPN/Bappenas. (2020). *Pedoman Teknis Penyusunan Rencana Aksi Tujuan Pembangunan Berkelanjutan (TPB)/ Sustainable Development Goals (SDGs)*. Kementerian Perencanaan Pembangunan

- Nasional / Badan Perencanaan dan Pembangunan Nasional (BAPPENAS).
- Kurniawan, E., Amidi, Gunawan, Susilowati, N., Paranti, L., & Santi W, D. G. (2022). *Panduan UNNES GIAT Penguatan Generasi Milenial Mendukung SDGs Desa*.
- Madjid, T., Rachmawaty, R., Fachri, M., Wikantosa, B., Uguy, N. S., Muklis, Agusta, I., Hariyanti, Mustakim, Dwiatmojo, R., Azizah, K., & Sukoyo. (2020). *Peringkat Status Indeks Desa Membangun (IDM)* (2020th ed.). Direktorat Jenderal Pembangunan dan Pemberdayaan Masyarakat Desa: Kementerian Desa, Pembangunan Daerah Tertinggal dan Transmigrasi, Republik Indonesia.
- McCullagh, P. (1980). Regression Models for Ordinal Data. *Journal of the Royal Statistical Society. Series B (Methodological)*, 42(2), 109–142.
- Moraes, J. O., Cruz, E. A., Pinheiro, Í., Oliveira, T. C. M., Alvarenga, V., Sant'Ana, A. S., & Magnani, M. (2019). An Ordinal Logistic Regression Approach to Predict the Variability on Biofilm Formation Stages by Five Salmonella Enterica Strains on Polypropylene and Glass Surfaces as Affected by pH, Temperature and NaCl. *Food Microbiology*, 83, 95–103. <https://doi.org/10.1016/j.fm.2019.04.012>
- Nisa, Z. I., Soleh, A. M., & Wijayanto, H. (2021). Identifikasi Faktor-Faktor yang Memengaruhi Prestasi Mahasiswa Menggunakan Regresi Logistik Ordinal dan Random Forest Ordinal. *Xplore: Journal of Statistics*, 10(1), 88–101. <https://doi.org/10.29244/xplore.v10i1.465>
- Permendesa Nomor 13. (2020). *Peraturan Menteri Desa, Pembangunan Daerah Tertinggal dan Transmigrasi Republik Indonesia*.
- Prasetya, Y. E., Hidayat, A. R. T., & Dinanti, D. (2019). Village Development Index of Probolinggo Coastal Villages Case study: Bhinor Village, Paiton District. *IOP Conference Series: Earth and Environmental Science*, 328(1), 1–8. <https://doi.org/10.1088/1755-1315/328/1/012056>
- Singer, G., Anuar, R., & Ben-Gal, I. (2020). A Weighted Information-Gain Measure for Ordinal Classification Trees. *Expert Systems with Applications*, 152. <https://doi.org/10.1016/j.eswa.2020.113375>
- Utami, N. A., & Wijayanto, A. W. (2022). Classification of Village Development Index at Regency/Municipality Level Using Bayesian Network Approach with K-Means Discretization. *Jurnal Aplikasi Statistika & Komputasi Statistik, Khusus*, 95–206.
- Wahyuningrum, V. (2020). Penerapan Radial Basis Function Neural Network dalam Pengklasifikasian Daerah Tertinggal di Indonesia. *Jurnal Aplikasi Statistika & Komputasi Statistik*, 12(1), 37–54.
- Wulandari, I. (2015). Performansi Piecewise Polynomial Smooth Support Vector Machine untuk Klasifikasi Desa Tertinggal di Provinsi Kalimantan Timur Tahun 2011. *Jurnal Aplikasi Statistika & Komputasi Statistik*, 7(1), 21–44.

