

Penerapan Metode CART pada Pengklasifikasian Bekerja dan Pengangguran di Kabupaten Subang*

Ilma Nabila¹, I Made Sumertajaya^{2‡}, and Mulianto Raharjo³

^{1,2}Department of Statistics, IPB University, Indonesia

³Kementerian Dalam Negeri, Indonesia

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

Copyright © 2022 Ilma Nabila, I Made Sumertajaya, and Mulianto Raharjo. 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

Unemployment is a complex problem faced by developing countries, including Indonesia. The high unemployment rate in Indonesia impacts poverty, so that the government seeks to carry out economic development. Subang is one of the districts that contributed 8,68 percent of the open unemployment rate in 2019 and increased by 9,48 percent in 2020. The incessant growth of industrial estates and smart city program development in Subang is one of the efforts to reduce unemployment. This study used a classification and regression tree (CART) to determine the factors that influenced unemployment status in Subang Regency. The advantage of the CART method is easy to interpret the results of the analysis. However, the accuracy of the classification tree is relatively low due to data imbalance. Therefore, this study used SMOTE method to deal with this problem. The optimal classification tree was formed from 17 terminal nodes and 6 explanatory variables. 7 terminal nodes represent work as work, and 10 terminal nodes represent unemployment as unemployment. The 6 explanatory variables consist of marital status (X3), attending job training (X5), the position in the family (X4), the education level (X2), gender (X1), and age (X6).

Keywords: CART, SMOTE, unemployment.

1. Pendahuluan

Indonesia merupakan negara berkembang yang memiliki jumlah penduduk semakin bertambah. Adanya penambahan penduduk tersebut dapat meningkatkan jumlah angkatan kerja. Angkatan kerja merupakan penduduk usia kerja yang bekerja dan pengangguran. Penduduk Indonesia yang merupakan angkatan kerja pada bulan Agustus 2020 sebanyak 138,22 juta orang (BPS 2020a). Nilai angkatan kerja tersebut cukup tinggi sehingga pemerintah berupaya melakukan pembangunan ekonomi dengan menciptakan lapangan pekerjaan dan memberikan kesempatan pada penduduk angkatan kerja agar dapat mengikuti kegiatan ekonomi di Indonesia (Muslim 2014).

* Received: Dec 2021; Reviewed: May 2022; Published: May 2022

Kesempatan dan penambahan lapangan kerja baru akan mengurangi jumlah pengangguran di Indonesia. Pengangguran merupakan masalah kompleks yang dihadapi oleh beberapa provinsi di Indonesia antara lain Jawa Barat dan khususnya di Kabupaten Subang. Kabupaten Subang merupakan kabupaten yang memiliki angka tingkat pengangguran terbuka pada urutan ke-18 di Jawa Barat. Kabupaten Subang pada akhir-akhir ini gencar membangun perindustrian dengan tujuan membuka lapangan kerja ataupun usaha. Hal tersebut dapat terlihat dengan banyaknya kawasan industri yang sudah berjalan dan adanya perencanaan pembangunan menuju smart city untuk meningkatkan sumber daya manusia berdaya saing (Sari et al. 2020). Meskipun upaya pemerintah Kabupaten Subang dilaksanakan akan tetapi pengangguran masih meningkat. Berdasarkan data survei angkatan kerja nasional (Sakernas) TPT di Kabupaten Subang sebesar 8,68 persen pada bulan Agustus 2019 dan sebesar 9,48 persen pada bulan Agustus 2020, sehingga peningkatan pengangguran pada kurun waktu satu tahun sebesar 0,8 persen (BPS 2020b).

Pemerintah melakukan salah satu upaya untuk mengurangi pengangguran dengan menerbitkan program Kartu Prakerja. Kartu Prakerja bertujuan untuk mempersiapkan dan melatih kemampuan serta keterampilan angkatan kerja agar bisa bersaing di dunia kerja (Consuello 2020). Selain itu, salah satu upaya lainnya dengan mengidentifikasi faktor-faktor yang berpengaruh terhadap pengangguran agar pemerintah dapat memberikan lapangan kerja yang tepat pada sasaran.

Penelitian sebelumnya telah dilakukan oleh Pratiwi dan Zain (2014) untuk mengklasifikasikan pengangguran terbuka di Sulawesi Utara dimana faktor-faktor yang digunakan, yaitu jenis kelamin, tingkat pendidikan, usia, status dalam keluarga, pelatihan kerja, status perkawinan, dan klasifikasi tempat tinggal. Selain itu, penelitian sebelumnya dilakukan juga oleh Yuliatin et al. (2011) menunjukkan bahwa umur, status dalam rumah tangga, status perkawinan, pendidikan, dan daerah tempat tinggal berpengaruh terhadap peluang terjadinya pengangguran di Sumatera Barat. Oleh karena itu, penelitian ini menggunakan faktor-faktor yang terdapat pada penelitian terdahulu dengan penambahan faktor, yaitu Kartu Prakerja dikarenakan Kartu Prakerja menjadi salah satu faktor yang memfasilitasi angkatan 2 kerja khususnya pengangguran untuk mendapatkan kelayakan pelatihan dan bertujuan untuk mengurangi pengangguran.

Analisis klasifikasi terdiri dari berbagai macam algoritma, diantaranya support vector machine, naive Bayes, fuzzy, decision tree, dan jaringan saraf tiruan (Wibawa et al. 2018). Menurut Sartono dan Syafitri (2010), metode yang memiliki kemampuan dalam memberikan kemudahan untuk menginterpretasikan hasil analisis dan memberikan dugaan dengan tingkat kesalahan yang kecil adalah metode classification and regression tree (CART). CART adalah analisis statistik nonparametrik yang mengidentifikasi subkelompok dari suatu populasi yang anggotanya memiliki karakteristik umum yang memengaruhi peubah respon yang diminati (Lemon et al. 2003). CART memiliki kelebihan, yaitu kekar terhadap pencilan, dapat digunakan pada ukuran data yang besar, dan peubah penjelas yang banyak (Steinberg dan Colla 1995). Selain itu, metode CART memiliki tujuan untuk mendapatkan suatu kelompok data yang homogen sebagai pencari dari suatu pengklasifikasian (Breiman et al. 1993).

Salah satu masalah yang terjadi saat menganalisis adalah adanya ketidakseimbangan data yang mengakibatkan kurangnya hasil akurasi. Oleh karena itu, perlu dilakukan penyeimbangan data antara kelas minor dan kelas mayor dengan suatu metode agar dapat meningkatkan hasil akurasi. Synthetic minority over-sampling technique (SMOTE) dapat digunakan untuk mengatasi permasalahan tersebut. Menurut Wijaya et al. (2018), SMOTE dapat meningkatkan hasil akurasi.

Tujuan pada penelitian ini adalah mengetahui faktor-faktor yang berpengaruh terhadap pengangguran di Kabupaten Subang serta karakteristiknya.

2. Metodologi

2.1 Data

Tabel 1 Peubah yang digunakan

Kode	Peubah	Skala pengukuran
Y	Status	Nominal
	0: Bekerja 1: Pengangguran	
X1	Jenis Kelamin	Nominal
	0: Perempuan 1: Laki-laki	
X2	Tingkat Pendidikan	Ordinal
	0: Tamat/Belum Tamat SD/SMP 1: SMA/SMK 2: Perguruan Tinggi	
X3	Status Perkawinan	Nominal
	0: Belum Kawin 1: Kawin 2: Cerai Mati/ Cerai Hidup	
X4	Status Dalam Keluarga	Nominal
	0: Kepala Keluarga 1: Bukan Kepala Keluarga	
X5	Mengikuti Pelatihan Kerja	Nominal
	0: Tidak Pernah 1: Pernah	
X6	Usia	Ordinal
	0: 15 – 49 1: 50 – 64 2: Usia > 64	
X7	Klasifikasi Tempat Tinggal	Nominal
	0: Pedesaan 1: Perkotaan	
X8	Kartu Pra-Kerja	Nominal
	0: Tidak Lolos 1: Lolos	

Penelitian Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari Badan Pusat Statistik (BPS). Data tersebut merupakan data mentah survei angkatan kerja nasional (Sakernas) bulan Agustus 2020. Data yang digunakan sebanyak 1598 data individu angkatan kerja. Data yang diolah menghasilkan peubah respon dengan berbentuk kategorik berupa bekerja (kategori 0) dan pengangguran (kategori 1). Peubah yang digunakan dapat dilihat pada Tabel 1.

2.2 Prosedur Analisis Data

Analisis data yang dilakukan pada penelitian ini menggunakan bantuan software R

versi 4.0.3 serta package yang ada di software tersebut, yaitu “rpart”, “caret”, “rpart.plot”, “class” dan “DMwR”. Adapun prosedur analisis data yang dilakukan adalah sebagai berikut:

1. Mengecek missing value dan cleaning pada data agar sesuai dengan data yang ingin dianalisis.
2. Pelabelan data angkatan kerja menjadi bekerja dan pengangguran sesuai definisi BPS (2020a).
3. Melakukan eksplorasi data dan analisis statistik deskriptif.
4. Pengecekan hubungan antara peubah respon dan peubah penjelas menggunakan uji khi-kuadrat. Hipotesis dari uji khi-kuadrat adalah sebagai berikut:
 - H_0 : Tidak ada hubungan antar peubah (saling bebas)
 - H_1 : Ada hubungan antar peubah (tidak saling bebas)
5. Membagi data menggunakan 10-fold cross validation dengan persentase sembilan bagian menjadi data latih dan satu bagian menjadi data uji.
6. Membentuk pohon klasifikasi tanpa SMOTE menggunakan data latih melalui tahapan berikut:
 - a. Menentukan pemilihan pemilah berdasarkan aturan pemilahan indeks gini, penentuan simpul terminal, penandaan label kelas, pemangkasan pohon klasifikasi dan penentuan pohon klasifikasi optimal. Penelitian ini menggunakan beberapa batasan pada pembentukan pohon klasifikasi, yaitu minsplit (jumlah amatan minimum pada simpul) sebesar 20 amatan, maxdepth (maksimal kedalaman pohon) sebesar 5 kedalaman, parameter biaya kompleksitas minimum sebesar 0,001 dan jumlah amatan minimum pada simpul terminal sebesar 5.
 - b. Membuat tabel ketepatan klasifikasi agar mengetahui akurasi, sensitivitas, dan spesifisitas pada pohon klasifikasi. Pada penelitian ini spesifisitas menggambarkan akurasi amatan bekerja sebagai bekerja dan sensitivitas menggambarkan akurasi amatan pengangguran sebagai pengangguran.
7. Melakukan penanganan data tidak seimbang pada data latih dengan menggunakan SMOTE.
 - a. Menentukan k-tetangga terdekat, yaitu $k = 5$.
 - b. Menentukan persentase oversampling sebesar 350% dan undersampling 150%.
 - c. Menghitung jarak amatan pada kelas minor menggunakan rumus VDM.
 - d. Pilih satu amatan di kelas minor secara acak dan tentukan 5 tetangga terdekat dengan mengurut jarak amatan terpilih dengan semua amatan pada kelas minor. Kemudian pilih kategori kelas mayoritas pada 5 tetangga terdekat, sehingga dapat menentukan nilai per peubah penjelas untuk mendapatkan data buatan baru.
 - e. Ulangi tahap (d) hingga mendapatkan hasil yang sesuai dengan yang ditentukan.
8. Membentuk pohon klasifikasi CART dengan data tambahan SMOTE seperti tahap (6).
9. Membandingkan hasil ketepatan klasifikasi.
10. Interpretasi hasil dari pohon klasifikasi terbaik.

3. Hasil dan Pembahasan

3.1 Eksplorasi Data

Data amatan individu dari survei angkatan kerja nasional (Sakernas) yang dilaksanakan pada bulan Agustus 2020 di Kabupaten Subang yang digunakan pada

penelitian ini sebanyak 1598 individu berkategori angkatan kerja. Hasil pelabelan yang telah dilakukan pada data angkatan kerja yang masuk ke dalam kategori bekerja sebanyak 1223 individu dan kategori pengangguran sebanyak 375 individu. Sebagian besar individu di Kabupaten Subang termasuk ke dalam kategori bekerja dengan presentase sebesar 76,53% dan pengangguran sebesar 23,47%.

Tabel 2 Pengangguran dan bekerja berdasarkan peubah penjelas

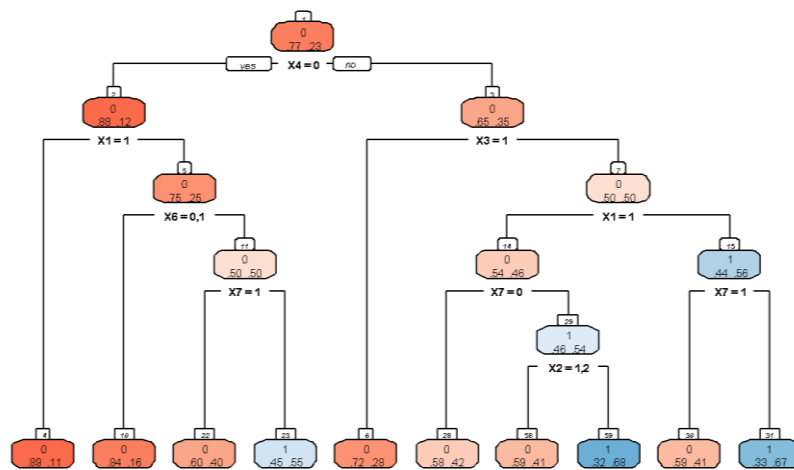
Peubah X	Kategori	Jumlah Bekerja	Jumlah Pengangguran	Jumlah Amatan	Pengangguran (%)
Jenis Kelamin	0: Perempuan	439	210	649	56,00
	1: Laki-laki	784	165	949	44,00
Tingkat Pendidikan	0: Tamat/ Belum Tamat SD/SD/SMP	888	242	1130	64,53
	1: SMA/SMK	247	112	359	29,87
	2: Perguruan Tinggi	88	21	109	5,60
Status Perkawinan	0: Belum kawin	112	113	225	30,13
	1: Kawin	983	218	1201	58,13
	2: Cerai mati/hidup	128	44	172	11,73
Status dalam Keluarga	0: Kepala keluarga	498	272	770	72,53
	1: Bukan Kepala keluarga	725	103	828	27,47
Mengikuti Pelatihan kerja	0: Tidak pernah	1091	348	1439	92,80
	1: Pernah	132	27	159	7,20
Usia	0: 15 – 49	763	268	1031	71,47
	1: 50 – 64	349	71	420	18,93
	2: Usia >64	111	36	147	9,60
Klasifikasi Tempat Tinggal	0: Pedesaan	880	225	1105	60,00
	1: Perkotaan	343	120	463	32,00
Kartu Prakerja	0: Tidak Lolos	1221	375	1596	100,00
	1: Lolos	2	0	2	0,00

Tabel 2 menjelaskan mengenai pengangguran dan bekerja berdasarkan peubah penjelas. Pada peubah jenis kelamin, perempuan yang pengangguran lebih mendominasi daripada laki-laki, yaitu sebesar 44,00% sedangkan perempuan sebesar 56,00%. Pengangguran pada peubah tingkat pendidikan didominasi oleh tingkat pendidikan tamat/belum tamat SD/SD/SMP. Pada peubah status perkawinan didominasi oleh pengangguran yang status kawin dengan persentase sebesar 58,13%. Pengangguran dengan status kepala keluarga sebesar 72,53% lebih besar daripada pengangguran dengan status bukan kepala keluarga sebesar 27,47%. Individu pengangguran yang tidak mengikuti pelatihan kerja sebesar 92,80% lebih besar daripada individu pengangguran yang pernah mengikuti pelatihan kerja.

Berdasarkan hasil eksplorasi pengangguran di Kabupaten Subang lebih besar di pedesaan dibanding di perkotaan. Pengangguran di Kabupaten Subang lebih besar yang tidak lolos daripada yang lolos kartu prakerja. Menurut hasil survei angkatan kerja nasional BPS (2020a), hal ini disebabkan oleh individu yang tidak mengetahui adanya program Kartu Prakerja dan tidak mendaftar saat adanya pendaftaran Kartu Prakerja

3.2 Analisis CART tanpa SMOTE

Pada penerapan metode CART, data pengamatan di penelitian ini dibagi menjadi data latih dan data uji. Pembagian data ini menggunakan *10-fold cross validation*. Tahap pertama, yaitu menentukan pemilah yang terbaik untuk menjadi pemilah utama. Pemilah utama yang didapatkan berdasarkan perhitungan indeks Gini adalah status dalam keluarga. Pohon klasifikasi menghasilkan 10 simpul terminal dengan melibatkan 6 peubah penjelas, yaitu status dalam keluarga (X4), jenis kelamin (X1), status perkawinan (X3), usia (X6), klasifikasi tempat tinggal (X7), tingkat pendidikan (X2) dengan kedalaman pohon sebesar 5.



Gambar 1 Pohon klasifikasi CART tanpa SMOTE

Tahap selanjutnya adalah proses pemangkasan pohon klasifikasi untuk mendapatkan pohon optimal atau pohon yang lebih sederhana. Nilai CP yang menghasilkan nilai kesalahan relatif paling kecil, yaitu untuk memangkaskan pohon klasifikasi maksimal sehingga mendapatkan pohon optimal. Pada kasus ini nilai kesalahan relatif paling kecil yang dihasilkan sebesar 0,97867 dengan CP sebesar 0,000 sehingga tidak perlu melakukan pemangkasan karena pohon sudah optimal. Pohon klasifikasi optimal dapat dilihat pada Gambar 1 terbentuk dari 7 simpul terminal yang menjelaskan kategori bekerja dan 3 simpul terminal yang menjelaskan kategori pengangguran. Hasil pohon klasifikasi bekerja dan pengangguran di Kabupaten Subang dievaluasi dengan menghitung nilai akurasi, sensitivitas, dan spesifisitas dengan menggunakan *10-fold cross validation*.

Tabel 3 Ketepatan klasifikasi CART tanpa SMOTE menggunakan data latih

Pengulangan dalam <i>Cross Validation</i>			
<i>Fold</i>	Akurasi	Spesifisitas	Sensitivitas
1	78,44%	97,19%	16,72%
2	78,46%	96,82%	11,43%

3	77,76%	92,35%	30,76%
4	78,58%	95,99%	22,35%
5	78,16%	97,55%	14,07%
6	78,72%	97,10%	17,96%
7	78,16%	96,06%	21,68%
8	78,58%	97,37%	17,21%
9	78,30%	95,72%	21,83%
10	78,58%	97,66%	14,29%
<i>Mean</i>	78,37%	96,38%	18,83%

Tabel 3 menjelaskan nilai akurasi, sensitivitas, dan spesifisitas pada masing-masing *fold*. Pada masing-masing *fold* untuk nilai spesifisitas cukup tinggi dibandingkan dengan nilai sensitivitas. Hal ini menunjukkan bahwa tingkat ketepatan klasifikasi pada pengangguran sebagai pengangguran atau sensitivitas cukup kecil. Nilai rata-rata akurasi sebesar 78,37%, spesifisitas sebesar 96,38%, dan sensitivitas sebesar 18,83%.

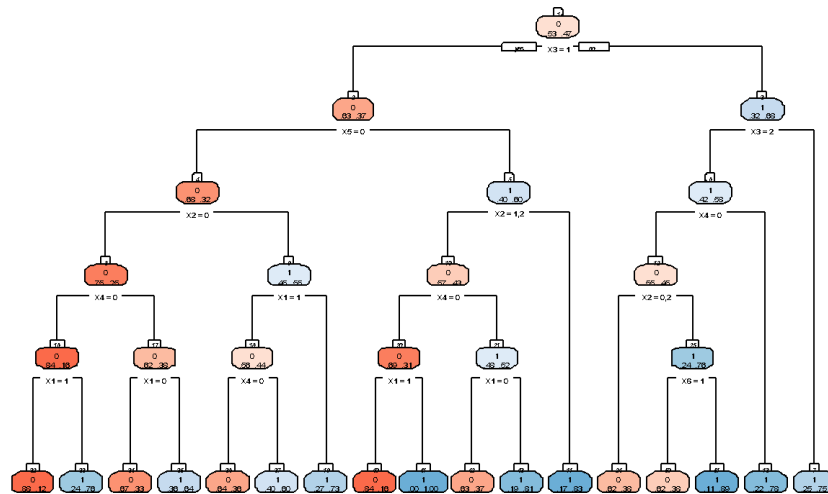
Tabel 4 Ketepatan klasifikasi CART tanpa SMOTE menggunakan data uji

Observasi	Prediksi		Ketepatan
	Bekerja a	Pengangguran n	
Bekerja	1188	35	97,14%
Pengangguran n	310	65	17,33%
Akurasi Total			78.41%

Berdasarkan Tabel 4 ketepatan klasifikasi CART tanpa SMOTE dengan menggunakan data uji menghasilkan nilai akurasi sebesar 78,41%, spesifisitas sebesar 97,14%, dan sensitivitas sebesar 17,33%. Ketepatan klasifikasi pada bekerja sebagai bekerja (spesifisitas) cukup tinggi yaitu sebanyak 1188 individu sedangkan ketepatan klasifikasi pengangguran sebagai pengangguran (sensitivitas) cukup kecil yaitu sebanyak 65 individu. Adapun kesalahan klasifikasi bekerja sebagai pengangguran sebanyak 35 individu dan pengangguran sebagai bekerja sebanyak 310 individu.

3.3 Analisis CART dengan SMOTE

Penerapan CART dengan SMOTE pada penelitian ini menggunakan data asli yang digabungkan dengan data buatan hasil SMOTE. Tahap pertama melakukan pemilihan pemilah. Pemilah terbaik pada analisis CART dengan SMOTE adalah peubah status perkawinan (X3) yang dapat dilihat pada Gambar 2. Peubah status perkawinan memiliki nilai impuritas paling tinggi daripada peubah penjelas lainnya. Pohon klasifikasi yang terbentuk menghasilkan 17 simpul terminal dan melibatkan 6 peubah penjelas, yaitu status perkawinan (X3), mengikuti pelatihan kerja (X5), status dalam keluarga (X4), tingkat pendidikan (X2), jenis kelamin (X1), dan usia (X6).



Gambar 2 Pohon klasifikasi CART dengan SMOTE

Proses pemangkasan dilakukan setelah mendapatkan pohon maksimal. Nilai kesalahan relatif minimum yang didapatkan pada pohon klasifikasi CART dengan SMOTE adalah sebesar 0,5300 dengan CP 0,0000. Hal tersebut menunjukkan bahwa tidak perlu dilakukan pemangkasan karena pohon dapat dikatakan sudah optimal. Pohon klasifikasi optimal menghasilkan 7 simpul terminal yang mengklasifikasikan angkatan kerja berkategori bekerja dan 10 simpul terminal yang mengklasifikasikan angkatan kerja berkategori pengangguran.

Tabel 5 Ketepatan klasifikasi CART dengan SMOTE pada data latih

Pengulangan dalam <i>Cross Validation</i>			
<i>Fold</i>	Akurasi	Spesifisitas	Sensitivitas
1	78,74%	75,00%	61,48%
2	72,82%	79,60%	62,96%
3	72,12%	83,55%	59,55%
4	71,87%	80,26%	62,96%
5	75,00%	78,28%	68,88%
6	72,47%	75,81	72,59%
7	72,56%	80,92%	63,70%
8	72,12%	77,63%	65,92%
9	69,33%	82,23%	65,92%
10	74,21%	82,89%	67,64%
<i>Mean</i>	73,12%	79,62%	65,16%

Nilai akurasi, sensitivitas, dan spesifisitas dari masing-masing *fold* mengalami kenaikan dan penurunan yang dapat dilihat pada Tabel 5. Jika dibanding dengan hasil akurasi pohon klasifikasi CART tanpa SMOTE, nilai rata-rata akurasi dan spesifisitas pada data latih pohon klasifikasi CART dengan SMOTE mengalami penurunan sebesar 5,25% dan 16,76%. Sehingga nilai akurasi dan spesifisitas menjadi 73,12% dan 79,62%. Beda halnya dengan nilai sensitivitas yang mengalami kenaikan sehingga menjadi 65,16%.

Tabel 6 Ketepatan klasifikasi CART dengan SMOTE pada data uji

Observasi	Prediksi		Ketepatan n
	Bekerj a	Penganggura n	
Bekerja	998	225	81,60%
Penganggura n	220	155	41,33%
Akurasi Total			72,15%

Setelah dilakukan penerapan model pohon klasifikasi dengan SMOTE pada data uji menghasilkan nilai akurasi dan spesifisitas yang menurun, namun nilai sensitivitas meningkat. Tabel 6 menjelaskan tentang ketepatan klasifikasi dengan nilai akurasi sebesar 72,15%, spesifisitas (ketepatan klasifikasi bekerja sebagai bekerja) sebesar 81,60%, dan sensitivitas (ketepatan klasifikasi pengangguran sebagai pengangguran) sebesar 41,33%.

4. Simpulan dan Saran

4.1 Simpulan

Hasil pohon klasifikasi terbaik, yaitu pohon klasifikasi CART dengan SMOTE yang menghasilkan sebanyak 17 simpul terminal terdiri dari 7 simpul terminal untuk kategori bekerja dan 10 simpul terminal untuk kategori pengangguran. Peubah penjas yang membentuk pohon klasifikasi terdiri dari 6 peubah penjas, yaitu status perkawinan (X3), pelatihan kerja (X5), status dalam keluarga (X4), tingkat pendidikan (X2), jenis kelamin (X1), dan usia (X6). Karakteristik pada pengangguran di Kabupaten Subang dijelaskan pada simpul 33, 35, 37, 19, 41, 43, 11, 51, 13, dan 7.

Pengangguran pada simpul terminal menghasilkan beberapa karakteristik, yaitu jika individu memiliki status kawin, tidak pernah mengikuti pelatihan kerja, tingkat pendidikan tamat/belum tamat SD/SD/SMP, status dalam keluarga, yaitu kepala keluarga, jenis kelamin perempuan. Jika individu memiliki status kawin, tidak pernah mengikuti pelatihan kerja, tingkat pendidikan tamat/belum tamat SD/SD/SMP, status dalam keluarga, yaitu bukan kepala keluarga, jenis kelamin laki-laki. Jika status kawin, tidak pernah mengikuti pelatihan kerja, tingkat pendidikan SMA/SMK dan perguruan tinggi, jenis kelamin laki-laki, dan status dalam keluarga, yaitu bukan kepala keluarga. Jika status kawin, tidak pernah mengikuti pelatihan kerja, tingkat pendidikan SMA/SMK dan perguruan tinggi, dan jenis kelamin perempuan. Jika status kawin, pernah mengikuti pelatihan kerja, tingkat pendidikan SMA/SMK dan perguruan tinggi, status dalam keluarga, yaitu kepala keluarga, dan jenis kelamin perempuan. Jika status kawin, pernah mengikuti pelatihan kerja, tingkat pendidikan SMA/SMK dan perguruan tinggi, status dalam keluarga, yaitu bukan kepala keluarga, dan jenis kelamin laki-laki. Jika status kawin, pernah mengikuti pelatihan kerja, dan tingkat pendidikan tamat/belum tamat SD/SD/SMP. Jika status belums kawin dan cerai mati/hidup, status dalam keluarga, yaitu kepala keluarga, dan tingkat pendidikan SMA/SMK, dan usia 15 – 49 dan 50 – 64 tahun. Jika status belum kawin dan cerai mati/hidup, status dalam keluarga, yaitu bukan kepala keluarga. Jika status belum kawin dan cerai mati/hidup, kawin.

4.2 Saran

Saran untuk penelitian selanjutnya, peneliti dapat menambahkan peubah lain untuk pembentukan pohon klasifikasi. Peneliti selanjutnya dapat meneliti pengaruh faktor pandemi Covid-19 terhadap pengangguran dikarenakan survei angkatan kerja nasional (Sakernas) pada bulan Agustus 2020 dilaksanakan saat pandemi Covid-19.

Daftar Pustaka

- [BPS] Badan Pusat Statistik. 2020a. Berita Resmi Statistik Keadaan Ketenagakerjaan Indonesia Agustus 2020. Jakarta (ID): Badan Pusat Statistik.
- [BPS] Badan Pusat Statistik. 2020b. Berita Resmi Statistik Keadaan ketenagakerjaan Provinsi Jawa Barat Agustus 2020. Jakarta (ID): Badan Pusat Statistik.
- Breiman L, Friedman JH, Olshen RA, Stone CJ. 1993. *Classification and Regression Trees*. New York (US): Chapman and Hall.
- Consuello Y. 2020. Analisis efektivitas Kartu Pra-Kerja di tengah pandemi covid-19. *'Adalah: Buletin Hukum dan Keadilan*. 4(1): 93-100.
- Lemon SC, Roy J, Clark MA, Friedmann PD, Rakowski W. 2003. Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. *Annals of Behavioral Medicine*. 26(3): 172–181.
- Muslim MR. 2014. Pengangguran terbuka dan determinasinya. *Jurnal Ekonomi dan Studi Pembangunan*. 15(2): 171–181.
- Pratiwi FE, Zain I. 2014. Klasifikasi pengangguran terbuka menggunakan CART (*Classification and Regression Tree*) di Provinsi Sulawesi Utara. *Jurnal Sains dan Seno POMITS*. 3(1): 54–59.
- Sari D, Takariani CSD, Praditya D, Puspitasari L. 2020. Analisis Strategis Kabupaten Subang Menuju Smart City. *Jurnal Penelitian Komunikasi*. 23(2):181–198. doi:10.20422/jpkv2i23.726.
- Sartono B, Syafitri UD. 2010. Metode pohon gabungan: solusi pilihan untuk mengatasi kelemahan pohon regresi dan klasifikasi tunggal. *Forum Statistika dan Komputasi*. 15(1): 1–7.
- Steinberg D, Colla P. 1995. *CART: Tree-Structured Non- Parametric Data Analysis*. San Diego, CA: Salford Systems.
- Wibawa AP, Purnama MGA, Akhbar MF, Dwiyanto FA. 2018. Metode-metode klasifikasi. Di dalam: *Seminar Ilmu Komputer dan Teknologi Informasi*, editor. ISSN: 2540-7902 dan p-ISSN 2541-366X. Malang: Universitas Negeri Malang. hlm 134-138.
- Wijaya J, Soleh AM, Rizki A. 2018. Penanganan Data Tidak Seimbang pada Pemodelan Rotation Forest Keberhasilan Studi Mahasiswa Program Magister IPB. *Xplore J Stat*. 2(2):32–40. doi:10.29244/xplore.v2i2.99.
- Yuliatin, Huseno T, Febriani. 2011. Pengaruh karakteristik kependudukan terhadap pengangguran di Sumatera Utara. *Jurnal Manajemen dan Kewirausahaan*. 2(2): 15–43.
- Yuliatin, Huseno T, Febriani. 2011. Pengaruh karakteristik kependudukan terhadap pengangguran di Sumatera Utara. *Jurnal Manajemen dan Kewirausahaan*. 2(2): 15–43.