

ANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI KEMATIAN PASIEN COVID-19 MENGGUNAKAN KLASIFIKASI BERSTRUKTUR POHON BINER DENGAN ALGORITMA QUEST

Siti Mariyam*, Yasmin Erika Faridhan, Fitria Virgantari

Program Studi Matematika, Universitas Pakuan

e-mail: sitimariyam1703@gmail.com

Diterima: 31 Januari 2022, disetujui: 1 Maret 2022, dipublikasi: 22 April 2022

Abstract: COVID-19 is an infectious disease caused by a type of coronavirus that was only discovered in December 2019. Patients with underlying medical conditions, or comorbidities, have a higher risk of developing severe illness due to COVID-19. The purpose of this study is to classify and analyze the factors which mostly affect the death in COVID-19 patients using QUEST algorithm. The main strengths of QUEST algorithm are unbiased selection of variables and high computational speed. Data used in this study are primary data of 14 variables on 460 COVID-19 patients taken from Dr. M. Goenawan Partowidigdo Lung Hospital in Cisarua, West Java, from March 2020 to January 2021. Results show that there are three significant factors that affected the death in COVID-19 patients. The first factor is the status of COVID-19 patients. The second and third factors are comorbidities, i.e. hypertension and kidney failure, respectively. The factor which mostly affected the death in COVID-19 patients is patient with probable status, with a mortality rate of 95%. The second most factor affecting the death in COVID-19 is patients with under-surveillance, suspected and confirmed status with kidney failure, where the mortality rate is 60%. The accuracy of the classification tree is 80.6%, which is quite optimal.

Keywords: COVID-19, comorbidities, death, classification, QUEST.

Abstrak: COVID-19 merupakan penyakit menular yang diakibatkan oleh jenis coronavirus yang baru ditemukan pada Desember 2019. Pasien dengan kondisi memiliki penyakit bawaan (komorbid) jauh lebih berisiko terkena penyakit yang parah akibat COVID-19. Tujuan penelitian ini yaitu mengklasifikasikan faktor-faktor yang mempengaruhi kematian pasien COVID-19 dan menganalisis faktor yang paling berpengaruh terhadap kematian pasien COVID-19 menggunakan algoritma QUEST. Kekuatan utama algoritma QUEST adalah pemilihan peubah yang tak bias dan kecepatan komputasi yang tinggi. Data yang digunakan dalam penelitian ini yaitu data primer dari 460 pasien COVID-19 yang diambil dari Rumah Sakit Paru Dr. M. Goenawan Partowidigdo di Cisarua, Jawa Barat dari bulan Maret 2020 sampai Januari 2021. Sebanyak 14 peubah digunakan dalam penelitian ini. Dari hasil pengolahan data dengan QUEST, diperoleh tiga faktor yang signifikan dalam mempengaruhi kematian pasien COVID-19. Faktor pertama yang mempengaruhi kematian pasien COVID-19 adalah status pasien COVID-19. Faktor kedua yaitu komorbid hipertensi, dan faktor ketiga adalah komorbid gagal ginjal. Faktor yang paling mempengaruhi kematian pasien COVID-19 yaitu pasien dengan status probabel dengan tingkat kematian 95%. Faktor berikutnya yang paling mempengaruhi kematian pasien dengan COVID-19 adalah pasien dengan status PDP, suspek dan terkonfirmasi yang memiliki komorbid gagal ginjal dengan tingkat kematian 60%. Ketepatan pohon klasifikasi yang terbentuk mencapai 80,6%, yang artinya ketepatan pohon klasifikasi optimal atau memiliki kesalahan yang kecil.

Kata Kunci: COVID-19, komorbid, kematian, klasifikasi, QUEST
PENDAHULUAN

COVID-19 (Coronavirus Disease-19) merupakan penyakit yang disebabkan oleh coronavirus SARS-CoV2. Virus baru ini pertama diketahui oleh WHO tanggal 31 Desember 2019, menyusul laporan adanya kluster kasus penyakit seperti pneumonia namun disebabkan oleh virus, atau disebut “viral pneumonia”, di Wuhan, Cina [1]. Penyakit yang baru ditemukan ini dengan cepat menyebar ke seluruh dunia, dan dideklarasikan sebagai pandemi oleh WHO pada 11 Maret 2020 [2]. Berdasarkan data dari Pusat Pengendalian dan Pencegahan Penyakit AS (*Centers for Disease Control and Prevention, CDC*), semua orang dengan kondisi memiliki penyakit bawaan (komorbid) jauh lebih berisiko terkena penyakit yang parah akibat COVID-19. Kasus kematian akibat COVID-19 yang terjadi pada pasien yang memiliki komorbid mencapai 94%. Komorbid yang dimaksud di antaranya adalah kanker, penyakit ginjal kronis, *chronic obstructive pulmonary disease (COPD)*/penyakit paru obstruktif kronik, kegemukan (obesitas), serta diabetes melitus tipe II [3].

Dengan beragamnya komorbid pasien, diperlukan suatu pengelompokan sebagai salah satu alat bantu dalam menganalisis komorbid yang berpengaruh terhadap pasien COVID-19. Pengelompokan atau klasifikasi dapat dilakukan dengan menggunakan pendekatan parametrik maupun pendekatan non parametrik. Pendekatan non-parametrik umumnya dilakukan pada data yang tidak memenuhi asumsi pendekatan parametrik [4]. Dalam studi ini digunakan pendekatan non parametrik, karena data pasien COVID-19 yang sebagian besar berupa peubah kategorik tidak dapat memenuhi asumsi kenormalan dalam pendekatan parametrik.

Pendekatan non parametrik yang digunakan dalam studi ini yaitu metode klasifikasi berstruktur pohon biner. Algoritma QUEST (*Quick, Unbiased and Efficient Statistical Tree*) dipilih karena sejumlah keunggulannya dibanding algoritma lainnya yang lebih banyak digunakan, seperti algoritma CART (*Classification And Regression Trees*). Proses komputasi QUEST lebih cepat daripada CART, terutama bila terdapat peubah kategorik dengan banyak kategori. Kekurangan algoritma CART lainnya di antaranya metode pemisahan (*splitting method*)-nya yang bias, terutama terhadap – sekali lagi – peubah kategorik yang memiliki lebih banyak kategori [5, 6].

Penelitian dengan menggunakan metode QUEST sebelumnya dilakukan oleh [7] mengenai faktor-faktor yang paling berpengaruh terhadap pasien penderita penyakit kencing manis di RS. Plamonia Makasar. [8] menerapkan metode QUEST dalam identifikasi faktor-faktor yang mempengaruhi penyakit diare pada balita di Indonesia.

Pada studi ini dilakukan klasifikasi faktor-faktor yang berpengaruh terhadap kematian pasien COVID-19 dengan metode klasifikasi berstruktur pohon menggunakan algoritma QUEST.

METODE KLASIFIKASI BERSTRUKTUR POHON

Metode klasifikasi berstruktur pohon merupakan metode yang digunakan untuk memprediksi suatu objek termasuk ke dalam keanggotaan kelas-kelas peubah respon kategorik, yang didasarkan atas pengukuran terhadap suatu peubah bebas atau lebih. Metode klasifikasi berstruktur pohon dibagi menjadi dua kelompok. Kelompok pertama menghasilkan pohon biner, yaitu sebuah pohon yang setiap simpulnya disekat menjadi dua simpul yang saling terpisah. Kelompok kedua adalah kelompok yang menghasilkan

pohon non-biner, yaitu sebuah pohon yang setiap simpulnya disekat menjadi dua atau lebih simpul yang terpisah [5, 9].

Istilah-istilah yang digunakan dalam pohon klasifikasi di antaranya [10, 11]:

- a. Simpul induk/simpul akar (*root node*), yaitu himpunan data di awal
- b. Simpul daun/simpul akhir (*leaf node/terminal node*), yaitu sebuah simpul yang sudah tidak bisa disekat kembali.
- c. Simpul dalam/simpul antara (*intermediate node*), yaitu simpul yang tidak termasuk simpul induk dan simpul akhir
- d. Peubah penyekat, yaitu peubah bebas yang digunakan untuk menyekat sebuah simpul. Peubah penyekat bisa berupa peubah kategorik atau peubah kontinu.

QUEST

QUEST merupakan algoritma pohon klasifikasi yang dikembangkan dari algoritma FACT (*Fast and Accurate Classification Tree*) [12]. Kekuatan utama algoritma QUEST adalah pemilihan peubah yang tak bias dan kecepatan komputasi yang tinggi [6].

Dalam algoritma QUEST, pemilihan peubah penyekat dan simpul penyekat dilakukan secara terpisah. Statistik uji F digunakan untuk peubah numerik, sedangkan statistik uji kai kuadrat (χ^2) digunakan untuk peubah kategorik. Peubah penyekat dipilih berdasarkan peubah yang memiliki tingkat homogenitas yang paling besar. Apabila peubah penyekat yang terpilih berupa peubah kategorik, maka perlu dilakukan transformasi ke peubah numerik, kemudian dilakukan analisis diskriminan kuadrat.

Algoritma QUEST dibagi menjadi tiga bagian yaitu, algoritma pemilihan peubah prediktor penyekat, algoritma penentuan simpul penyekat, dan algoritma penghentian pembentukan pohon [5].

METODOLOGI PENELITIAN

Data

Data yang digunakan dalam penelitian ini berupa informasi mengenai pasien COVID-19 yang diambil dari Rumah Sakit Paru Dr. M. Goenawan Partowidigdo, Cisarua, Jawa Barat. Data diambil dari bulan Maret 2020 sampai Januari 2021, sebanyak 672 pasien. Peubah yang digunakan dalam penelitian ini sebanyak 14 peubah, yang terdiri dari 13 peubah kategorik dan satu peubah numerik. Peubah tak bebas dalam penelitian ini yaitu kondisi pasien COVID-19 (sembuh dan meninggal). Peubah bebasnya yaitu status pasien COVID-19, umur, kehamilan, ruangan perawatan, komorbid diabetes melitus, tuberkulosis, penyakit jantung, kekurangan nutrisi, asma, penyakit hati, gagal ginjal, pneumonia, dan hipertensi.

Peubah status terdiri dari empat kategori, yaitu [13]:

1. PDP (pasien dalam pengawasan), artinya pasien dengan demam lebih dari 38 derajat Celcius disertai salah satu gejala pernapasan seperti batuk/sesak/sakit tenggorokan/pneumonia ringan hingga berat, serta 14 hari sebelum timbul gejala memiliki riwayat kontak dengan pasien konfirmasi atau probabel COVID-19.
2. Probabel, artinya pasien suspek dengan gejala ISPA berat, atau meninggal dengan

- gejala COVID-19 tanpa adanya hasil pemeriksaan laboratorium PCR.
3. Suspek, artinya seseorang dengan penyakit pernapasan akut yang parah (memerlukan rawat inap), serta tanpa adanya diagnosis lainnya yang sepenuhnya menjelaskan manifestasi klinis.
 4. Terkonfirmasi, artinya seseorang dengan adanya hasil tes laboratorium terinfeksi COVID-19 terlepas dari ada tidaknya tanda gejala klinis.

Tahapan Analisis

1. Pengolahan data

Setelah melalui proses penyaringan, data yang digunakan dalam studi ini sebanyak 460 pasien COVID-19; terdiri dari 108 pasien dengan kondisi meninggal dan 352 pasien dengan kondisi sembuh.

Data diolah dengan bantuan program QUEST [14]. Perhitungan manual juga ditampilkan dalam tulisan ini. Langkah-langkah algoritma QUEST ditampilkan pada tahapan 2 – 7.

2. Pemilihan peubah penyekat

Pemilihan peubah penyekat dilakukan dengan uji statistik untuk menghitung nilai signifikan. Uji statistik untuk peubah numerik dilakukan dengan uji F, sedangkan uji *chi-square*/kai kuadrat (χ^2) dilakukan untuk peubah kategorik.

3. Pemilihan dan perbandingan nilai-p

Pemilihan nilai-p dilakukan berdasarkan peubah yang memiliki nilai-p terkecil.

4. Penentuan simpul penyekat

Hasil peubah nilai-p numerik dan nilai-p peubah kategorik dibandingkan. Apabila peubah yang terpilih peubah kategorik, maka peubah tersebut diubah menjadi peubah numerik dengan cara transformasi ke dalam sebuah vektor dummy I dimensi.

$$v = (v_1, v_2, \dots, v_I) \quad (1)$$

$$v_i = \begin{cases} 1 & x=b_i \\ 0 & x \neq b_i \end{cases}, i = 1, 2, \dots, I \quad (2)$$

5. Analisis diskriminan kuadratik

Langkah-langkah analisis diskriminan kuadratik adalah sebagai berikut :

- a. Penentuan rataan dan ragam dari pengamatan peubah respon.
- b. Perhitungan nilai ragam minimal dari beberapa kategori peubah respon.
- c. Apabila hasil minimal ragam tersebut tidak sama dengan nol, maka sumbu X dibagi menjadi tiga interval yaitu $(-\infty, d_1)$, (d_1, d_2) dan (d_2, ∞) dengan d merupakan akar-akar dari persamaan:

$$P(G|t)s_G^{-1}\phi\left(\frac{x-\bar{x}_G}{s_G}\right) = P(B|t)s_H^{-1}\phi\left(\frac{x-\bar{x}_H}{s_H}\right) \quad (3)$$

Kedua ruas diberi log untuk memperoleh persamaan kuadrat $ax^2 + bx + c = 0$

dengan

$$a = s_G^2 - s_H^2 \quad (4)$$

$$b = 2(\bar{x}_G s_H^2 - \bar{x}_H s_G^2) \quad (5)$$

$$c = (\bar{x}_H s_G)^2 - (x_G s_H)^2 + 2s_G^2 s_H^2 \log \left(\frac{P(G|t) s_H}{P(H|t) s_G} \right) \quad (6)$$

- d. Akar-akar dari persamaan tersebut dipilih berdasarkan akar yang memiliki nilai yang paling mendekati nilai rata-rata dari peubah tiap kategori.
6. Penghentian pembentukan pohon
Penghentian pembentukan pohon dilakukan apabila tidak ada simpul yang tidak bisa disekat kembali.
 7. Pembentukan diagram pohon
Pohon klasifikasi terbentuk berdasarkan peubah yang menjadi simpul induk, simpul dalam dan simpul akhir yang didapatkan dari tahapan kedua sampai keenam, dan digambarkan dalam sebuah diagram pohon.
 8. Interpretasi
Pada tahap ini, interpretasi dilakukan berdasarkan diagram pohon yang terbentuk pada tahap ketujuh.

HASIL DAN PEMBAHASAN

Penentuan Simpul Awal (t_0)

Penentuan simpul awal (t_0) diawali dengan mencari peubah nilai-p terkecil. Nilai-p dihasilkan dengan bantuan *software* SPSS. Hasil uji statistik dapat dilihat pada Tabel 1.

Tabel 1. Hasil uji statistik simpul awal

Peubah	Nilai statistik uji	Nilai-p
Kai kuadrat		
Status (x_1)	64,904	0,000
Kehamilan (x_3)	1,238	0,266
Ruangan perawatan (x_4)	3,309	0,124
Diabetes melitus (x_5)	4,388	0,036
Tuberkulosis (x_6)	1,009	0,315
Jantung (x_7)	0,545	0,450
Kekurangan nutrisi (x_8)	0,891	0,345
Asma (x_9)	0,616	0,432
Penyakit hati (x_{10})	2,204	0,138
Gagal ginjal (x_{11})	11,455	0,001
Pneumonia (x_{12})	1,393	0,238
Hipertensi (x_{13})	19,237	0,001
Uji F		
Umur (x_2)	0,015	0,904

Dari Tabel 1 dapat dilihat bahwa peubah yang memiliki nilai p terkecil yaitu 0,0000 adalah peubah status (x_1). Statistik hitung untuk peubah x_1 dengan $\alpha = 0,05$ adalah $\frac{0,05}{13} = 0,0038$. Berdasarkan perbandingan nilai $0,0000 < 0,0038$ maka peubah x_1 terpilih sebagai simpul awal. Peubah x_1 terdiri dari 4 kategori yaitu PDP, probabel, suspek atau terkonfirmasi. Sehingga untuk mendapatkan titik penyekat, peubah ditransformasikan menjadi peubah numerik. Hasil transformasi koordinat diskriminan terbesar dapat dilihat pada Tabel 2.

Tabel 2. Tabulasi silang x_1 hasil transformasi dengan y

y	x_1				Total
	6,051 $x_1 = 0$	-5,487 $x_1 = 1$	6,052 $x_1 = 2$	5,843 $x_1 = 3$	
1	132	1	80	139	352
0	43	19	23	23	108
Total	175	20	103	162	460

Dari Tabel 2 dapat ditentukan nilai rata-rata kategori meninggal (\bar{x}_0) dan kategori sembuh (\bar{x}_1); nilai ragam kategori meninggal (S_0) dan kategori sembuh (S_1); serta nilai peluang masing-masing kategori, sehingga diperoleh hasil sebagai berikut:

$$\bar{x}_0 = 3,977$$

$$\bar{x}_1 = 5,930$$

$$S_0 = 4,511$$

$$S_1 = 0,637$$

$$p(0|t) = 0,235$$

$$p(1|t) = 0,765$$

Analisis diskriminan kuadratik dilakukan terhadap peubah x_1 agar diperoleh titik penyekat. Titik penyekat yaitu akar dari persamaan $ax^2 + bx + c = 0$, dengan

$$a = S_0^2 - S_1^2 = 19,943$$

$$b = 2(\bar{x}_0 S_1^2 - \bar{x}_1 S_0^2) = -238,113$$

$$c = (\bar{x}_0 S_0^2)^2 - (\bar{x}_1 S_1^2)^2 + 2S_0^2 S_1^2 \ln \left| \frac{p(0|t) S_1^2}{p(1|t) S_0^2} \right| = 6459,48$$

Maka analisis diskriminan kuadratiknya yaitu sebagai berikut :

$$-238,113^2 - 4 \cdot 19,943 \cdot 6459,48 = -458587,83, \text{ dikarenakan hasil kurang dari nol}$$

$$\text{maka } d = \frac{3,977+5,930}{2} = 4,9535$$

Dengan demikian, simpul t_0 disekat menjadi dua simpul, yaitu simpul $t_1 \leq 4,9535$ dan $t_2 > 4,9535$, dan dapat disimpulkan bahwa x_1 memotong di nilai $\xi = -5,487$. Sehingga, simpul t_1 terdiri dari pasien dengan status probabel yang meninggal sebanyak 19 pengamatan serta pasien dengan status probabel yang sembuh sebanyak 1 pasien. Simpul t_2 terdiri dari pasien dengan status PDP, suspek, atau terkonfirmasi yang meninggal sebanyak 89, dan pasien dengan status PDP, suspek, atau terkonfirmasi yang sembuh sebanyak 351 orang.

Proses selanjutnya yaitu penentuan simpul t_2 dan simpul-simpul berikutnya sampai terdapat salah satu syarat penghentian pembentukan pohon terpenuhi.

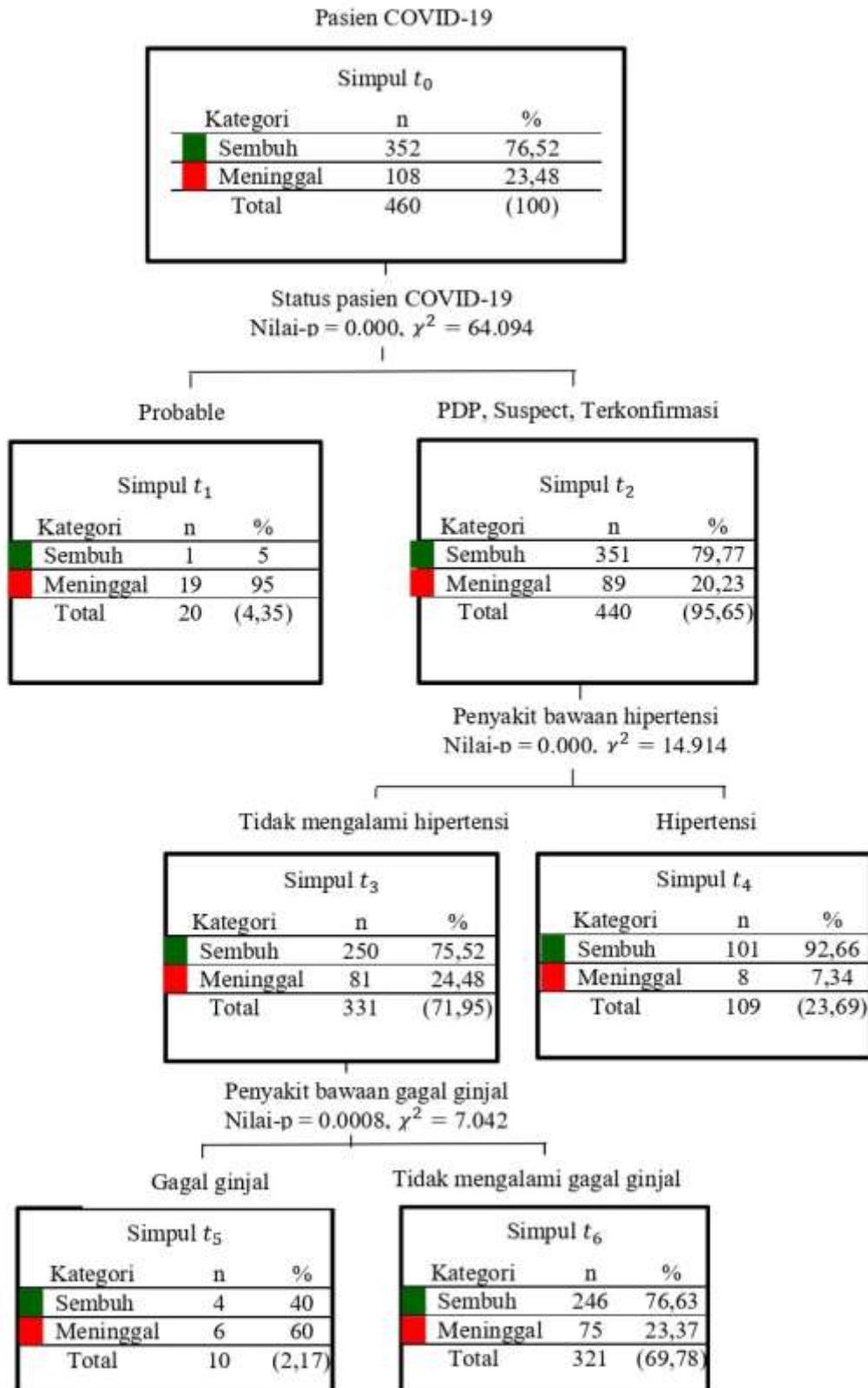
Pohon Klasifikasi yang Terbentuk

Faktor pertama yang mempengaruhi kematian pasien COVID-19 adalah statusnya. Pasien dengan status probabel diklasifikasikan sebagai meninggal, dan dalam pohon ditempatkan ke sebelah kiri sebagai simpul akhir. Pada simpul akhir pertama ini terdapat 20 pengamatan, dengan pasien meninggal sebanyak 19 orang. Tiga status lainnya yaitu PDP, suspect, dan terkonfirmasi ditempatkan ke sebelah kanan sebagai simpul dalam. Pada simpul dalam ini terdapat 440 pengamatan, dengan pasien meninggal sebanyak 89.

Faktor kedua yang mempengaruhi kematian pasien COVID-19 adalah komorbid hipertensi. Pasien yang tidak memiliki komorbid hipertensi berada di sebelah kiri sebagai simpul dalam dengan total 331 pengamatan, terdiri dari 81 pasien yang meninggal. Pasien yang memiliki komorbid hipertensi berada di sebelah kanan sebagai simpul akhir dengan 109 pengamatan, yang terdiri dari 8 pasien yang meninggal.

Faktor ketiga yang mempengaruhi kematian pasien COVID-19 adalah komorbid gagal ginjal. Pasien yang memiliki komorbid gagal ginjal ditempatkan di sebelah kiri dan yang tidak memiliki komorbid gagal ginjal di sebelah kanan. Kedua kategori ini merupakan simpul akhir dengan masing-masing jumlah pasien 6 dan 75 pasien yang meninggal.

Faktor yang paling mempengaruhi kematian pasien COVID-19 yaitu pasien dengan status probabel dengan tingkat kematian 95%, dan pasien dengan status PDP, suspect, atau terkonfirmasi yang memiliki komorbid gagal ginjal dengan tingkat kematian 60%. Berdasarkan detail tersebut, terbentuklah pohon pada Gambar 1.



Gambar 1. Pohon klasifikasi yang terbentuk

Ketepatan Klasifikasi Pohon

Ketepatan klasifikasi (*correct classification*) terdiri atas APER, *sensitivity*, dan *specificity* [15]. Ketepatan klasifikasi dapat dihitung berdasarkan Tabel 3.

Tabel 3. Ukuran ketepatan klasifikasi

Amatan	Dugaan		Total
	Meninggal	Sembuh	
Meninggal	28	80	108
Sembuh	9	343	352
Total	37	423	460

Persentase ketepatan klasifikasi kategori meninggal dapat dihitung sebagai nilai *Sensitivity*, sebagai berikut :

$$Sensitivity = \frac{28}{37} \times 100\% = 75,6\%$$

Persentase ketepatan klasifikasi kategori sembuh dapat dihitung sebagai nilai *Specificity*, sebagai berikut:

$$Specificity = \frac{343}{423} \times 100\% = 81\%$$

Akurasi ketepatan klasifikasi keseluruhan dapat dihitung sebagai berikut:

$$APER = \frac{28+343}{460} \times 100\% = 80,6\%$$

Secara ringkas, nilai ketepatan klasifikasi dalam kategori pasien COVID-19 meninggal yaitu sebesar 75,6 %. Nilai ketepatan klasifikasi dalam kategori pasien COVID-19 sembuh yaitu sebesar 81%. Secara keseluruhan ketepatan klasifikasi pohon yang terbentuk yaitu 80,6%. Dengan kata lain, peluang kesalahan klasifikasi pohon yaitu 0,194, yang artinya pohon klasifikasi ini optimal.

KESIMPULAN

Dari hasil klasifikasi terhadap data pasien COVID-19 di RS Paru Dr. M. Goenawan Partowidigdo, Cisarua, Kab. Bogor, terdapat faktor-faktor yang mempengaruhi kematian pasien COVID-19, yaitu peubah status pasien COVID-19, komorbid hipertensi, dan gagal ginjal. Faktor yang paling mempengaruhi kematian pasien COVID-19 yaitu pasien COVID-19 dengan status probabel, serta pasien dengan status PDP, suspect, dan terkonfirmasi yang memiliki komorbid gagal ginjal. Ketepatan pohon klasifikasi ini mencapai 80,6 %, yang artinya pohon klasifikasi ini optimal, dengan peluang kesalahan klasifikasi yang kecil sebesar 19,4%.

DAFTAR PUSTAKA

- [1] [WHO] World Health Organization. (2021). Coronavirus disease (COVID-19). <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/coronavirus-disease-covid-19> [Diakses 5 April 2022].
- [2] [WHO] World Health Organization. (2020). WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. <https://www.who.int/director-general/speeches/detail/who-director-general-s->

- [opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020](#) [Diakses 13 April 2022].
- [3] [CDC] Centers for Disease Control and Prevention. (2022). Underlying medical conditions associated with higher risk for severe COVID-19: Information for healthcare professionals. <https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-care/underlyingconditions.html> [Diakses 13 April 2022].
- [4] Vayssières, M.P., Plant, R.E., Allen-Diaz, B.H. (2000). Classification trees: An alternative non-parametric approach for predicting species distributions. *Journal of Vegetation Science*. **11**: 679-694. <https://doi.org/10.2307/3236575>
- [5] Loh, W.Y. (2014a). Classification and regression tree methods. *Wiley StatsRef: Statistics Reference Online*. <https://doi.org/10.1002/9781118445112.stat03886>
- [6] Loh, W.Y. (2014b). Fifty years of classification and regression trees. *International Statistical Review*. **82** (3): 329-348. <https://doi.org/10.1111/insr.12016>
- [7] Alwi, W., Tahir, N.A. (2015). Metode klasifikasi berstruktur pohon dengan algoritma QUEST. *Jurnal MSA*. **3** (2): 14-24.
- [8] Rizki, V. D., Setyawan, Y. (2018). Penerapan metode *Quick, Unbiased, Efficient Statistical Trees (QUEST)* untuk menentukan faktor-faktor yang mempengaruhi penyakit diare pada balita di Indonesia. *Jurnal Statistika Industri dan Komputasi*. **3** (1):1-10. <https://doi.org/10.34151/statistika.v3i01.1074>
- [9] Han, J., Kamber, M., Pei, J. (2011). *Data Mining Concepts and Techniques*. San Francisco: Morgan Kaufmann Publishers.
- [10] Sartono, B. (2015). Pohon klasifikasi - Bagian 1: Gambaran umum dan algoritma dasar yang perlu diketahui. Seri Tulisan Data Mining. <http://bagusco.staff.ipb.ac.id/files/2016/05/Pohon-Klasifikasi.pdf> [Diakses 13 April 2022].
- [11] Misshuari, I.W., Chairunnisah. (2020). Penerapan metode Classification And Regression Trees (CART) untuk menentukan faktor-faktor yang mempengaruhi pembayaran kredit oleh nasabah di PT BPRS Gebu Prima Medan. *Karismatika*. **6**(3): 10-19. <https://doi.org/10.24114/jmk.v6i3.22202>
- [12] Loh, W.Y., Vanichsetakul, N. (1998). Tree-Structured Classification via Generalized Discriminant Analysis. *Journal of the American Statistical Association*. **83**(403): 715-725. <https://doi.org/10.1080/01621459.1988.10478652>
- [13] Rahayu, S., Nugraha, S.M., Manik, C., Handayani, N., Jovina, T., Budiati, N. (2020). *COVID-19: The Nightmare or Rainbow*. Mata Aksara. Thesis doi: [10.31237/osf.io/tvyuw](https://doi.org/10.31237/osf.io/tvyuw)
- [14] Shih, Y. S. 2020. *QUEST user manual*. Department of Mathematics, National Chung Cheng University, Taiwan.

- [15] Johnson, R.A., Wichern, D.W. (2007). *Applied Multivariate Statistical Analysis*. 6th edition. London: Pearson.