

KLASIFIKASI SEKOLAH PADA PENERIMAAN MAHASISWA BARU JALUR SNMPTN DI IPB UNIVERSITY MENGGUNAKAN METODE REGRESI LOGISTIK MULTINOMIAL

FARAH AULIA



**DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2022**



@Hak cipta milik IPB University

IPB University



IPB University
— Bogor Indonesia —

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Perpustakaan IPB University

PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Klasifikasi Sekolah pada Penerimaan Mahasiswa Baru Jalur SNMPTN di IPB University Menggunakan Metode Regresi Logistik Multinomial” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Agustus 2022

Farah Aulia
G14180063



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

ABSTRAK

FARAH AULIA. Klasifikasi Sekolah pada Penerimaan Mahasiswa Baru Jalur SNMPTN di IPB University Menggunakan Metode Regresi Logistik Multinomial. Dibimbing oleh UTAMI DYAH SYAFITRI dan RAHMA ANISA.

IPB University merupakan perguruan tinggi yang memiliki kualitas pendidikan baik, begitu pula dengan penerimaan mahasiswa barunya. Salah satu jalur penerimaan mahasiswa baru yang ada di IPB University yaitu SNMPTN. Hal yang ingin diketahui IPB University adalah membuat fungsi klasifikasi sekolah pendaftar SNMPTN menggunakan metode regresi logistik multinomial. Sebelum dilakukan klasifikasi, dilakukan analisis gerombol dengan metode *two-step cluster* menggunakan data tahun 2016-2020. Metode *two-step cluster* menghasilkan empat gerombol optimal. Gerombol 1 adalah sekolah dengan peringkat terendah (komitmen, kualitas, dan konsistensi rendah). Gerombol 2 dan 3 adalah sekolah yang memiliki kriteria khusus pada kategori tertentu. Gerombol 4 adalah sekolah yang memiliki peringkat tertinggi (komitmen, kualitas, dan konsistensi tinggi). Hasil dari analisis gerombol akan dipakai untuk klasifikasi sebagai peubah respon. Model regresi logistik multinomial memberikan ketepatan klasifikasi sebesar 98,86%. Peluang sekolah setiap kategori dapat diperoleh dari model logit yang terbentuk. Peubah signifikan dari model yang dihasilkan mampu memprediksi peluang sekolah untuk masuk dalam gerombol tertentu. Model yang didapatkan dari hasil klasifikasi akan divalidasi menggunakan data tahun 2021. Validasi yang dihasilkan masih belum cukup bagus memprediksi karena nilai akurasi sebesar 61,8%. Nilai akurasi dikatakan bagus jika berada pada selang 80-110%. Hal ini dikarenakan setiap tahun terjadi perubahan sekolah yang mendaftar SNMPTN di IPB University.

Kata kunci: data kategorik, Regresi Logistik Multinomial, SNMPTN

ABSTRACT

FARAH AULIA. School Classification on New Student Admissions on the SNMPTN Path at IPB University Using the Multinomial Logistics Regression Method. Supervised by UTAMI DYAH SYAFITRI and RAHMA ANISA.

IPB University is considered as one of the best universities, with quite qualified student enrollment. One of the new student admissions paths at IPB University is SNMPTN. This study aimed to generate a classification function for the school of college student enrollment, in this case, SNMPTN on IPB University, with a multinomial logistic regression analysis method. The researcher conducted the cluster analysis by two-step cluster using 2016 until 2020 data and produced four optimal clusters. The clustering was proceed and resulted four optimal clusters. The first cluster ranked lowest based on commitment, quality, and high school student enrollment consistency. The second and third clusters were schools with particular criteria for specific categories. The last cluster was the other three clusters highest commitment, quality, and consistency. In this case, the multinomial logistic regression model had 98,86% accuracy. The probability of all clusters in a high

school could be derived from the logit model. Logistic regression's coefficient was used to predict probability value which falls into four clusters model. The prediction model of the actual data on years interval 2016 until 2020 for data testing 2021 was still below expectation because low accuracy data was 61,8%.

Keywords: categorical data, Multinomial Regression Logistic, SNMPTN

@Tak cipta milik IPB University

IPB University



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2022
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.



KLASIFIKASI SEKOLAH PADA PENERIMAAN MAHASISWA BARU JALUR SNMPTN DI IPB UNIVERSITY MENGGUNAKAN METODE REGRESI LOGISTIK MULTINOMIAL

FARAH AULIA

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana pada
Program Studi Statistika dan Sains Data

**DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2022**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Tim Penguji pada Ujian Skripsi:
1 Dr. Ir. Erfiani, M.Si

Judul Skripsi : Klasifikasi Sekolah pada Penerimaan Mahasiswa Baru Jalur
SNMPTN di IPB University Menggunakan Metode Regresi
Logistik Multinomial

Nama : Farah Aulia
NIM : G14180063

@Hak cipta milik IPB University

Disetujui oleh

Pembimbing 1:
Dr. Utami Dyah Syafitri, S.Si., M.Si



Pembimbing 2:
Rahma Anisa, M.Si



Diketahui oleh

Ketua Departemen:
Dr. Anang Kurnia, M.Si
NIP 197308241997021001

Tanggal Ujian: 15 Juli 2022

Tanggal Lulus:

PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanaahu wa ta'ala atas segala karunia-Nya sehingga skripsi ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan Agustus 2021 sampai bulan Juli 2022 ini ialah klasifikasi, dengan judul “Klasifikasi Sekolah pada Penerimaan Mahasiswa Baru Jalur SNMPTN di IPB University Menggunakan Metode Regresi Logistik Multinomial”.

Terima kasih penulis ucapkan kepada para pembimbing, Ibu Dr. Utami Dyah Syafitri, S.Si., M.Si dan Ibu Rahma Anisa, M.Si yang telah membimbing dan banyak memberi saran, serta Ibu Dr. Ir. Erfiani sebagai dosen penguji pada ujian sidang penulis. Di samping itu, penghargaan penulis sampaikan kepada staf Tata Usaha Departemen Statistika atas bantuannya dalam kelancaran administrasi. Penulis juga mengucapkan terima kasih kepada teman-teman Statistika Angkatan 55 dan sahabat-sahabat penulis atas segala bantuan dan dukungannya selama ini. Ungkapan terima kasih juga disampaikan kepada mama, papa, adik-adik penulis, serta seluruh keluarga yang telah memberikan dukungan, doa, dan kasih sayangnya.

Semoga skripsi ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Agustus 2022

Farah Aulia



DAFTAR ISI

DAFTAR TABEL	xi
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xi
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Tujuan	2
TINJAUAN PUSTAKA	3
2.1 Pemeringkatan dan Kategori Sekolah	3
2.2 <i>Two-Step Cluster</i>	4
2.3 Regresi Logistik	5
2.4 Regresi Logistik Multinomial	6
2.5 Pendugaan Parameter	7
2.6 Pengujian Simultan	8
2.7 Pengujian Parsial	8
2.8 Interpretasi Koefisien	8
2.9 Ketepatan Klasifikasi Model	9
III METODE	10
3.1 Data	10
3.2 Prosedur Analisis Data	10
IV HASIL DAN PEMBAHASAN	13
4.1 Deskripsi Data	13
4.2 Analisis Gerombol	15
4.3 Analisis Regresi Logistik Multinomial	17
4.4 Validasi Model	21
V SIMPULAN DAN SARAN	22
5.1 Simpulan	22
5.2 Saran	22
DAFTAR PUSTAKA	23
LAMPIRAN	25
RIWAYAT HIDUP	37

DAFTAR TABEL

1	Kriteria pemeringkatan sekolah	3
2	Ketepatan klasifikasi	9
3	Peubah yang digunakan	10
4	Karakteristik sekolah berdasarkan peubah kategorik	15
5	Dugaan parameter regresi logistik multinomial	18
6	Peluang prediksi sekolah masuk ke gerombol 4	19
7	Nilai odds rasio	19
8	Persentase ketepatan klasifikasi	20
9	Validasi data tahun 2021	21

DAFTAR GAMBAR

1	Proses pembentukan <i>cf tree</i>	4
2	Rata-rata persentase siswa mendaftar ulang SNMPTN dan SBMPTN	13
3	Kriteria sekolah berdasarkan kategori perak tahun 2016-2020	14
4	Kenaikan nilai rapor tahun 2016-2020	14
5	Nilai BIC	15
6	Ukuran gerombol yang dihasilkan	16

DAFTAR LAMPIRAN

1a	Peubah yang digunakan dan kategorinya	26
1b	Peubah yang digunakan dan kategorinya	27
2	Kriteria sekolah berdasarkan peubah kategorik	28
3	Hasil penggerombolan dengan nilai BIC	29
4a	Kriteria penggerombolan	30
4b	Kriteria penggerombolan	31
5a	Dugaan parameter regresi logistik multinomial	32
5b	Dugaan parameter regresi logistik multinomial	33
5c	Dugaan parameter regresi logistik multinomial	34
5d	Dugaan parameter regresi logistik multinomial	35
5e	Dugaan parameter regresi logistik multinomial	36



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

I PENDAHULUAN

1.1 Latar Belakang

Penerimaan mahasiswa baru merupakan kegiatan rutin yang dilakukan oleh seluruh Perguruan Tinggi di Indonesia setiap tahunnya. Kegiatan ini dapat dikatakan sebagai titik awal proses pencarian calon mahasiswa baru yang berkualitas. Menerima calon mahasiswa yang berkompeten diharapkan dapat menunjang mutu dan kualitas Perguruan Tinggi itu sendiri (Kurniawan 2016). IPB University merupakan salah satu Perguruan Tinggi yang memiliki kualitas pendidikan yang baik, begitu pula dengan penerimaan mahasiswa barunya. Ada enam jalur seleksi penerimaan mahasiswa baru di IPB University, yaitu Seleksi Nasional Masuk Perguruan Tinggi Negeri (SNMPTN), Ujian Tulis Berbasis Komputer-Seleksi Bersama Masuk Perguruan Tinggi Negeri (UTBK-SBMPTN), Ketua OSIS/Prestasi Internasional (PIN), Kelas Internasional, Beasiswa Utusan Daerah (BUD), dan Ujian Tulis Mandiri Berbasis Komputer (UTMBK). Diantara jalur tersebut, ada tiga jalur seleksi penerimaan mahasiswa baru di IPB University menggunakan rapor, yaitu SNMPTN, Ketua OSIS/PIN, dan BUD sehingga perlu dibuat pemetaan sekolah yang mendaftar di IPB University.

Pada penelitian sebelumnya oleh Dewantari (2021), telah dilakukan penggerombolan sekolah pada penerimaan mahasiswa baru jalur SNMPTN di IPB University dengan metode *two-step cluster*. Penggerombolan menghasilkan gerombol dengan kriteria yang baik dan rinci, yaitu Gerombol satu merupakan gerombol sekolah yang memiliki peringkat terendah dengan kriteria yaitu *Low Commitment*, *Low Quality*, dan *Low Consistency*, Gerombol dua merupakan gerombol sekolah yang memiliki *Commitment*, *Quality*, dan *Consistency* yang cukup tinggi pada jalur SBMPTN tetapi cukup rendah pada jalur SNMPTN, sedangkan Gerombol tiga sebaliknya. Gerombol empat merupakan sekolah yang memiliki *grade* tertinggi dengan kriterianya *High Commitment*, *High Quality*, dan *High Consistency*.

Salah satu fungsi klasifikasi yang dapat digunakan jika peubah respon dan peubah penjelasnya merupakan peubah kategorik yaitu dengan regresi logistik. Regresi Logistik dapat digunakan untuk mendeskripsikan hubungan antara peubah respon yang memiliki dua kategori atau lebih dengan satu atau lebih peubah penjelas berskala kategorik atau interval (Hosmer dan Lemeshow 2013). Pada umumnya, ada beberapa analisis klasifikasi lainnya seperti analisis diskriminan dan *Chi-Square Automatic Interaction Detection* (CHAID). Akan tetapi, analisis diskriminan digunakan pada kasus dimana peubah respon berupa data kualitatif dan peubah penjelas berupa data kuantitatif. Hal ini dapat diatasi oleh regresi logistik yang mampu mengakomodir peubah bebas berupa data kualitatif, kuantitatif, atau campuran keduanya (Crystine 2018). Metode CHAID menghasilkan ketepatan yang lebih baik dibandingkan analisis regresi logistik tetapi metode regresi logistik dapat mengetahui besarnya pengaruh masing-masing peubah penjelas (Shyntia 2022).

Pada penelitian ini, digunakan metode regresi logistik multinomial karena respon yang dipakai ada empat kategori. Regresi logistik multinomial merupakan regresi logistik yang digunakan saat variabel dependen mempunyai skala yang bersifat *polichotomous* atau multinomial (Agresti 2019). Regresi logistik multinomial merupakan regresi logistik yang digunakan saat peubah respon

mempunyai skala yang bersifat multinomial dengan peubah respon berskala nominal lebih dari dua kategori (Hosmer dan Lemeshow 2013). Model regresi logistik multinomial digunakan ketika tidak ada urutan di antara kategori respon (Aditya *et al.* 2015).

2.2 Tujuan

Tujuan yang akan dicapai dalam penelitian ini adalah membuat fungsi klasifikasi sekolah pendaftar SNMPTN di IPB University untuk memperhitungkan berapa peluang siswa diterima di jalur SNMPTN menggunakan metode regresi logistik multinomial.

II TINJAUAN PUSTAKA

2.1 Pemeringkatan dan Kategori Sekolah

Pemeringkatan sekolah di SNMPTN hanya dilakukan oleh sekolah yang memiliki NPSN (Nomor Pokok Sekolah Nasional). NPSN sendiri menjadi pedoman pengenalan dari PDSP (Pusat Data dan Statistik Pendidikan) untuk sekolah sebagai sebuah satuan pendidikan melalui tiap Dinas Pendidikan kota atau kabupaten di Indonesia. Setelah sekolah dipastikan memiliki NPSN, pemeringkatan LTMPT (Lembaga Tes Masuk Perguruan Tinggi) yang dimaksud adalah urutan akreditasi sekolah dari A, B, hingga C. Jumlah sekolah yang dapat mengikuti SNMPTN ditentukan oleh akreditasi sekolahnya masing-masing (Purbowati 2022).

Tabel 1 Kriteria pemeringkatan sekolah

Frekuensi IPK $\geq 2,10$ (F%)	Ada besaran IPK	Peringkat
$F \geq 75$	$IPK \geq 3,50$	A+
	$2,75 \leq IPK < 3,50$	A
	selainnya ($IPK < 2,75$)	A-
$50 \leq F < 75$	$IPK \geq 3,50$	B+
	$2,75 \leq IPK < 3,50$	B
	selainnya ($IPK < 2,75$)	B-
$25 \leq F < 50$	$IPK \geq 3,50$	C+
	$2,75 \leq IPK < 3,50$	C
	selainnya ($IPK < 2,75$)	C-
$F < 25$	-	D
Tanpa data	Tanpa data	-

Tabel 1 menunjukkan pemeringkatan sekolah sesudah tahun akademik 1996/1997. Sekolah dengan peringkat A dapat mengajukan 40% siswa terbaiknya, sekolah dengan peringkat B dapat mengajukan 25% siswa terbaiknya, sekolah dengan peringkat C dan lainnya dapat mengajukan 5% siswa terbaiknya. Semakin baik peringkat sekolah, semakin besar pula persentase siswa yang dapat mengikuti SNMPTN. Kriteria sekolah dibedakan dalam tiga kategori, yaitu komitmen, kualitas, dan konsistensi. Ketiga kategori tersebut diperoleh berdasarkan hasil penggerombolan dengan menggabungkan peubah-peubah yang sesuai dan identik dari masing-masing gerombol (Dewantari 2021).

Kategori komitmen dapat dilihat berdasarkan siswa yang mendaftar SNMPTN jika lolos. Kategori kualitas dapat dilihat berdasarkan mahasiswa yang memiliki $IPK \geq 3,50$ dan sekolah yang termasuk dalam Top 1000 SNMPTN. Sekolah yang termasuk ke dalam Top 1000 sekolah SNMPTN adalah sekolah yang memiliki rata-rata TPS (Tes Potensi Skolastik) 1000 terbesar pertama (LTMPT 2021). Kategori konsistensi dapat dilihat berdasarkan kekonsistenan sekolah yang sering diterima oleh IPB pada jalur SNMPTN.

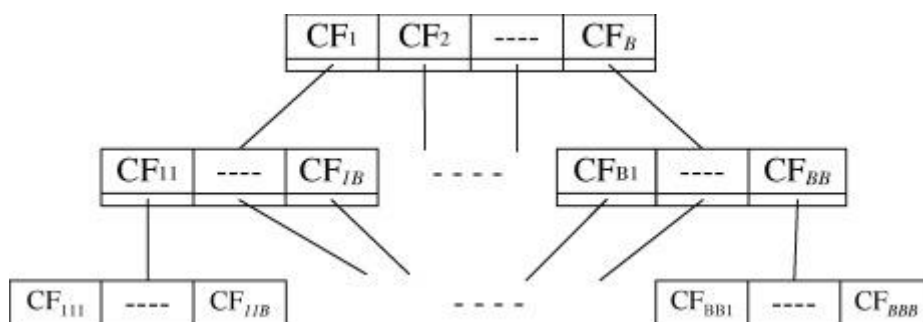
2.2 Two-Step Cluster

Metode penggerombolan *Two-Step Cluster* adalah metode untuk jenis data yang berukuran besar. Metode ini dikembangkan untuk mengolah data yang memiliki tipe peubah yang berbeda, yaitu kontinu dan kategorik (Chiu *et al.* 2001). Menurut Hair *et al.* (2010), data yang besar adalah data yang berjumlah lebih dari 500 amatan. Tahapan penggerombolan dengan metode *Two-Step Cluster* ada dua, yaitu tahap *pre-clustering* (penggerombolan awal) objek ke dalam anak gerombol dan tahap pembentukan gerombol optimal (penggerombolan akhir) (Lathifaturrahmah 2014).

a) Tahap Pre-Clustering

Menurut Anonimous (2001), tahap penggerombolan awal ini dilakukan dengan pendekatan sekuensial, yaitu objek diamati satu persatu berdasarkan ukuran jarak dan ditentukan apakah objek tersebut masuk dalam gerombol yang telah terbentuk atau harus membentuk gerombol baru. Pada tahap ini, dilakukan dengan pembentukan *Cluster Features (CF) Tree*. *CF Tree* terdiri dari beberapa tingkat cabang (*nodes*) dan masing-masing cabang berisikan individu objek (*entries*) dari gerombol awal. Tingkatan daun yang terdapat pada cabang merepresentasikan anak gerombol.

Prosedur *CF Tree* dilakukan dengan memilih satu amatan awal secara acak yang akan diukur jaraknya satu persatu dengan amatan lainnya menggunakan ukuran jarak yang telah ditentukan. Jika besarnya jarak terletak pada daerah penerimaan, maka amatan akan menjadi anggota anak gerombol. Jika besarnya jarak terletak di luar wilayah daerah penerimaan, maka amatan tersebut masuk ke dalam gerombol yang telah dibentuk (Mongi 2015).



Gambar 1 Contoh *cf tree* (Horng *et al.* (2011) *Science Direct*. 38(1):306-313

b) Tahap Pembentukan Gerombol Optimal

Pada tahap ini, hasil dari *CF Tree* digerombolkan dengan penggerombolan hierarki menggunakan metode *agglomerative*, yaitu dimulai dengan n gerombol yang masing-masing beranggotakan satu objek. Dua gerombol yang paling dekat digabung dan ditentukan kembali jarak antar gerombol yang baru. Menentukan banyaknya gerombol optimal dapat dilakukan dengan menghitung nilai *Bayesian Information Criterion (BIC)* terkecil dengan rumus sebagai berikut:

$$BIC(J) = -2 \sum_{j=1}^J \xi_j + m_j \log(N) \quad (1)$$

$$m_j = j \left(2K^A + \sum_{k=1}^{K^B} (L_k - 1) \right) \quad (2)$$

keterangan:

ξ_j : log-likelihood untuk gerombol j

m_j : parameter model

J : banyaknya gerombol

K^A : jumlah peubah kontinu

K^B : jumlah peubah kategorik

L_k : jumlah kategori untuk peubah kategorik ke- k

N : jumlah total amatan

Jika nilai BIC selalu menurun tiap penambahan jumlah gerombol, maka dapat dilihat dengan perbandingan rasio perubahan jarak terbesar pertama dengan rasio perubahan jarak terbesar kedua. Jika perbandingan rasio perubahan jarak terbesar pertama (R_{j1}) dengan rasio perubahan jarak terbesar kedua (R_{j2}), $\left(\frac{R_{j1}}{R_{j2}}\right) > c_2$ ($c_2 = 1,15$ diperoleh berdasarkan studi simulasi SPSS (2001)), maka banyaknya gerombol optimal adalah j_1 . Jika sebaliknya, maka banyaknya gerombol optimal yang diperoleh adalah $\max\{j_1, j_2\}$.

2.3 Regresi Logistik

Analisis regresi adalah suatu teknik statistika yang digunakan untuk mengetahui hubungan antara dua variabel atau lebih. Analisis regresi logistik merupakan salah satu jenis analisis regresi dimana variabel respon bersifat kategorik dan variabel prediktor bersifat kategorik atau numerik (Hosmer dan Lemeshow 2013). Regresi logistik merupakan salah satu metode yang dapat digunakan untuk mencari hubungan variabel respon yang bersifat *dichotomous* (berskala nominal atau ordinal dengan dua kategori) atau *polychotomous* (mempunyai skala nominal atau ordinal dengan lebih dari dua kategori) dengan satu atau lebih variabel prediktor dan variabel respon bersifat kontinu atau kategorik (Agresti 2019).

Secara umum, persamaan regresi logistik adalah sebagai berikut (Hosmer dan Lemeshow 2013):

$$\pi(x_i) = \frac{\exp(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi})}{1 + \exp(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi})} \quad (3)$$

keterangan:

- $\pi(x_i)$: peluang terjadinya kategori variabel respon
 x_{ji} : variabel prediktor ke- j
 p : banyaknya variabel prediktor
 β_0 : intersep
 $\beta_1, \beta_2, \dots, \beta_p$: koefisien regresi untuk setiap variabel prediktor
 i : 1, 2, ..., n

2.4 Regresi Logistik Multinomial

Regresi logistik multinomial merupakan regresi logistik yang digunakan saat variabel dependen mempunyai skala yang bersifat *polichotomous* atau multinomial (Agresti 2019). Regresi logistik multinomial merupakan perluasan regresi logistik biner, dengan kata lain apabila variabel dependen bersifat multinomial maka analisis hubungan kausal yang digunakan adalah regresi logistik multinomial.

Misalkan peubah respon terdiri atas empat kategori, maka peubah respon y dapat diberi kode 0, 1, 2, atau 3, dimana $y = 0$ merupakan kategori acuan. Pada model regresi logistik biner, terdapat satu fungsi logit $y = 1$ terhadap $y = 0$, maka pada model regresi logistik multinomial empat kategori terdapat tiga fungsi logit, yaitu fungsi logit $y = 1$ terhadap $y = 0$, fungsi logit $y = 2$ terhadap $y = 0$, dan fungsi logit $y = 3$ terhadap $y = 0$. Menurut Hosmer dan Lemeshow (2013), persamaan regresi logistik multinomial secara umum adalah sebagai berikut:

$$P(Y = j|x) = \pi_j(x) = \frac{\exp[g_j(x)]}{1 + \sum_{k=0}^{r-1} \exp[g_k(x)]} = \frac{\exp(\beta_{j0} + \beta_{j1}x_1 + \beta_{j2}x_2 + \dots + \beta_{jp}x_p)}{1 + \sum_{k=0}^{r-1} \exp(\beta_{k0} + \beta_{k1}x_1 + \beta_{k2}x_2 + \dots + \beta_{kp}x_p)} \quad (4)$$

dimana:

- $P(Y = j|x)$: peluang bersyarat dari peubah respon y untuk kategori ke- j pada vektor x ($j = 1, 2, 3$, dan 4)
 $\pi_j(x)$: persamaan regresi logistik untuk peubah respon y untuk kategori ke- j dengan kovariat x
 $g_j(x)$: logit pada peubah respon y untuk kategori ke- j
 x : peubah penjelas
 β_{kp} : parameter model logit

Dalam regresi logistik multinomial, suatu peubah respon dengan r kategori akan membentuk persamaan logit sebanyak $r-1$. Pada penelitian ini, model regresi logistik multinomial terdiri atas empat kategori y ($y = 1, 2, 3$, dan 4) sehingga dibutuhkan tiga fungsi logit dan dipilih kategori respon yang menjadi kategori pembanding yaitu $y = 4$. Diperlukan peubah boneka (*dummy variable*) untuk peubah penjelas yang bersifat kategorik. Secara umum, jika sebuah peubah skala nominal atau ordinal mempunyai k kemungkinan nilai, maka diperlukan $k-1$ peubah boneka. Untuk membuat menjadi fungsi linear, dilakukan transformasi logit karena fungsi



regresi yang diperoleh berdasarkan persamaan regresi logistik di atas berbentuk non-linear. Persamaan transformasi logit adalah sebagai berikut (Agresti 2019):

$$g(x) = \ln \left[\frac{\pi_j(x)}{\pi_0(x)} \right] = \beta_0 + \sum_{k=1}^p \beta_k x_{ik} \quad (5)$$

dimana:

j : 1, 2, ..., $r-1$

$\pi_0(x)$: persamaan regresi logistik untuk kategori pembandingan

β_0 : intersep

β_k : koefisien regresi untuk variabel ke- k

x_{ik} : variabel prediktor ke- i

2.5 Pendugaan Parameter

Pendugaan merupakan suatu pernyataan mengenai parameter populasi yang diketahui berdasarkan informasi dari sampel acak yang diambil dari populasi bersangkutan. Dalam pendugaan parameter, digunakan metode kemungkinan maksimum (*maximum likelihood*). Karena amatan diasumsikan saling bebas, maka fungsi kemungkinan maksimumnya adalah:

$$l(\beta) = \prod_{i=1}^n \pi_i(x) \quad (6)$$

$$l(\beta) = \prod_{i=1}^n \pi_0(x)^{y_{0i}} \pi_1(x)^{y_{1i}} \pi_2(x)^{y_{2i}} \pi_3(x)^{y_{3i}} \quad (7)$$

Prinsip metode kemungkinan maksimum yaitu memaksimumkan logaritma fungsi kemungkinannya sehingga:

$$L(\beta) = \ln (l(\beta)) \quad (8)$$

$$L(\beta) = \ln \left[\prod_{i=1}^n \pi_0(x)^{y_{0i}} \pi_1(x)^{y_{1i}} \pi_2(x)^{y_{2i}} \pi_3(x)^{y_{3i}} \right] \quad (9)$$

dimana:

i : 1, 2, ..., n

y_{0i} : respon pada pengamatan ke- i

$\pi(x_i)$: peluang kejadian ke- i bernilai $y = 1$

Dengan melakukan penurunan terhadap $L(\beta)$, maka diperoleh nilai β menggunakan iterasi *Newton-Raphson* yang memaksimumkan $L(\beta)$ dengan syarat $\frac{\partial L}{\partial \beta} = 0$.

2.6 Pengujian Simultan

Pengujian simultan dilakukan dengan menggunakan *likelihood-ratio test*. Pengujian ini merupakan pengujian model dengan membandingkan likelihood untuk model lengkap (L_1) dan likelihood untuk model yang semua parameternya sama dengan nol (L_0). Uji yang dapat digunakan untuk mengetahui peran seluruh peubah penjelas di dalam model secara bersama-sama yaitu uji-G (Hosmer dan Lemeshow 2013). Hipotesisnya adalah:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1: \text{minimal ada satu } \beta_i \neq 0, i = 1, 2, \dots, p$$

Statistik uji-G atau *likelihood-ratio test* dinyatakan sebagai berikut:

$$G = -2 \ln \left[\frac{L_0}{L_p} \right] \quad (10)$$

dimana:

L_0 : fungsi kemungkinan tanpa peubah penjelas

L_p : fungsi kemungkinan dengan peubah penjelas

Statistik uji-G mengikuti distribusi *chi-square* sehingga untuk memperoleh keputusan dilakukan perbandingan dengan titik kritis $\chi^2_{p(a)}$. Kriteria penolakan yaitu tolak H_0 apabila $G > \chi^2_{p(a)}$ atau nilai- $p < \alpha$.

2.7 Pengujian Parsial

Pengujian parsial atau uji Wald dilakukan untuk menguji variabel prediktor atau parameter β_j berpengaruh signifikan atau tidak terhadap variabel respon secara parsial. Hipotesisnya adalah:

$$H_0: \beta_j = 0 \text{ (pengaruh peubah ke-} j \text{ tidak signifikan)}$$

$$H_1: \beta_j \neq 0, j = 1, 2, \dots, p \text{ (pengaruh peubah ke-} j \text{ signifikan)}$$

Menurut Agresti (2019), statistik uji-G dinyatakan sebagai berikut:

$$W_j = \left[\frac{\hat{\beta}_j}{\widehat{SE}(\hat{\beta}_j)} \right]^2 \quad (11)$$

dimana:

$\hat{\beta}_j$: nilai koefisien dugaan peubah prediktor ke- j

$\widehat{SE}(\hat{\beta}_j)$: galat baku dari $\hat{\beta}_j$

Kriteria penolakan yaitu tolak H_0 apabila $W_j > \chi^2_{p(a)}$ atau nilai- $p < \alpha$ (Hosmer dan Lemeshow 2013).

2.8 Interpretasi Koeifisien

Interpretasi koefisien dalam regresi logistik dapat dilihat menggunakan nilai *odds ratio*. *Odds ratio* adalah perbandingan nilai risiko pada dua individu dengan rumus sebagai berikut:

$$Odds = \frac{\pi_i}{1 - \pi_i} \quad (12)$$

dimana:

π_i : peluang sukses

$1 - \pi_i$: peluang gagal

Odds ratio (ψ) menunjukkan perbandingan berapa kali lipat kenaikan atau penurunan angka kejadian $Y = i$ terhadap $Y = 0$ sebagai kategori pembanding jika nilai variabel prediktor (x) berubah sebesar nilai tertentu (Agresti 2019).

$$\psi_1 = \frac{\pi_1(1)/\pi_o(1)}{\pi_1(o)/\pi_o(o)} \quad (13)$$

$$\psi_2 = \frac{\pi_2(1)/\pi_o(1)}{\pi_2(o)/\pi_o(o)} \quad (14)$$

Jika $\psi_i < 1$ menunjukkan bahwa antar kedua variabel terdapat hubungan negatif dan jika $\psi_i > 1$ menunjukkan bahwa antar kedua variabel terdapat hubungan positif (Zahro dan Zain 2018).

2.9 Ketepatan Klasifikasi Model

Menurut Hosmer dan Lemeshow (2013), salah satu ukuran kebaikan model adalah ketika memiliki peluang salah klasifikasi yang kecil. Ketepatan klasifikasi model dapat diketahui dengan menggunakan tabel ketepatan klasifikasi, yaitu tabel frekuensi dua arah antara kelompok data aktual dan prediksi.

Tabel 2 Ketepatan klasifikasi

Aktual	Prediksi				Total	Ketepatan
	1	2	3	4		
1	a	b	c	d	n_1	$\frac{a}{n_1}$
2	e	f	g	h	n_2	$\frac{f}{n_2}$
3	i	j	k	l	n_3	$\frac{k}{n_3}$
4	m	n	o	p	n_4	$\frac{p}{n_4}$
Total	n_1	n_2	n_3	n_4	n	$\frac{(a + f + k + p)}{n}$

Tingkat ketepatan klasifikasi yaitu banyaknya dugaan yang tepat berdasarkan banyak contohnya. Tingkat ketepatan klasifikasi juga dapat dilihat dari nilai spesifisitas dan sensitivitasnya. Spesifisitas merupakan kemampuan model dalam memprediksi kejadian $Y = 0$. Sensitivitas merupakan kemampuan model dalam memprediksi $Y = 1$. Semakin besar tingkat ketepatan klasifikasi maka semakin tinggi tingkat akurasi model.

III METODE

3.1 Data

Data yang digunakan pada penelitian ini berasal dari Direktorat Administrasi Pendidikan dan Penerimaan Mahasiswa Baru IPB University, yaitu data sekolah pendaftar SNMPTN IPB University tahun ajaran 2016-2020 sebagai data *training* dan tahun ajaran 2021 sebagai data *testing*. Peubah respon yang digunakan yaitu empat gerombol yang dihasilkan dari penggerombolan yang dilakukan. Sedangkan peubah penjelas yang digunakan berupa 30 peubah kontinu, yaitu X1 hingga X25 dan 12 peubah kategorik, yaitu X26 hingga X42 yang dapat dilihat pada Tabel 3. Peubah dengan kategorinya dapat dilihat pada Lampiran 1.

Tabel 3 Peubah yang digunakan

Peubah	Nama peubah	Tipe
Y	Komitmen, Kualitas, dan Konsistensi	Kategorik
X ₁ – X ₅	Persentase Siswa Mendaftar Ulang SNMPTN IPB pada Tahun 2016 – 2020	Numerik
X ₆ – X ₁₀	Persentase Siswa Mendaftar Ulang SBMPTN IPB pada Tahun 2016 – 2020	Numerik
X ₁₁ – X ₁₅	Persentase Siswa Lolos SNMPTN XY pada Tahun 2016 – 2020	Numerik
X ₁₆ – X ₂₀	Persentase Siswa Lolos SNMPTN IPB pada Tahun 2016 – 2020	Numerik
X ₂₁ – X ₂₅	Persentase Mahasiswa dengan IPK ≥ 3.5 pada Tahun 2016 – 2020	Numerik
X ₂₆ – X ₃₀	Kategori Perak Sekolah pada Tahun 2016 – 2020	Numerik
X ₃₁ – X ₃₅	Kenaikan Nilai Rapor pada Tahun 2016 – 2020	Kategorik
X ₃₆	Akreditasi Sekolah pada Tahun 2016 – 2020	Kategorik
X ₃₇	Kepemilikan Sekolah Tahun 2016 – 2020	Kategorik
X ₃₈	Kekonsistenan Sekolah pada Tahun 2016 – 2020	Kategorik
X ₃₉	Loyalitas Sekolah pada SNMPTN 2016 – 2020	Kategorik
X ₄₀	Jenis Sekolah pada Tahun 2016 – 2020	Kategorik
X ₄₁	Keikutsertaan Sekolah pada Top 1000 Sekolah SNMPTN	Kategorik
X ₄₂	Daerah Asal Sekolah	Kategorik

3.2 Prosedur Analisis Data

Prosedur analisis data pada penelitian ini menggunakan perangkat lunak *IBM SPSS Statistics 28* versi *trial* dan *RStudio* versi 4.0.2 dengan *package foreign, nnet, ggplot2, reshape2, caret, performance, pscl, readxl, lmtest*, dan *glmnet*. Tahapan prosedur analisis pada penelitian ini adalah sebagai berikut:

1. Melakukan eksplorasi data

Eksplorasi data dilakukan untuk melihat karakteristik data lebih lanjut. Eksplorasi dilakukan dengan melihat rata-rata persentase siswa mendaftar ulang SNMPTN dan SBMPTN di IPB University dan kriteria sekolah berdasarkan kategori perak pada tahun 2016 – 2020. Selanjutnya, membuat

grafik kenaikan nilai rapor pada tahun 2016 – 2020 menggunakan diagram batang.

2. Melakukan analisis gerombol:

- a. Melakukan praproses data untuk data numerik dan kategorik pada tahun 2016 – 2020.
- b. Melakukan penggerombolan menggunakan metode *two-step cluster*. Penggerombolan dilakukan tanpa penanganan pencilan dengan tahapan *pre-clustering* dan pembentukan gerombol optimal dengan nilai BIC.
- c. Menghitung kualitas gerombol dengan nilai rata-rata koefisien *Silhouette*. Koefisien *Silhouette* berada diantara selang -1 sampai 1.
- d. Melakukan interpretasi kriteria masing-masing gerombol. Ada tiga kategori utama dari empat gerombol yang terbentuk, yaitu komitmen, kualitas, dan konsistensi dengan masing-masing kriteria, yaitu *low* (rendah) dan *high* (tinggi). Kriteria sekolah akan dikelompokkan ke dalam tiga kategori tersebut. Sebagai contoh, dalam satu gerombol memiliki komitmen rendah, kualitas rendah, dan konsistensi tinggi. Selain itu, dalam satu gerombol lain memiliki komitmen tinggi, kualitas tinggi, dan konsistensi rendah, begitupun dengan gerombol lainnya.
- e. Hasil penggerombolan akan dijadikan peubah respon untuk membangun model regresi logistik multinomial.

3. Membangun model regresi logistik multinomial:

- a. Menduga parameter regresi logistik multinomial.
Pendugaan parameter regresi logistik multinomial dilakukan dengan fungsi kemungkinan maksimum.
- b. Melakukan pengujian parameter secara simultan.
Pengujian simultan dilakukan untuk mengetahui kecocokan model. Pengujian simultan menggunakan uji G dengan hipotesis:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1: \text{minimal ada satu } \beta_i \neq 0, i = 1, 2, \dots, p$$
Tolak H_0 jika $G > \chi^2_{p(\alpha)}$ atau nilai-p $< \alpha$ yang artinya minimal ada satu peubah penjelas yang berpengaruh nyata terhadap peubah respon pada taraf nyata 5%.
- c. Melakukan pengujian parameter secara parsial.
Pengujian parsial dilakukan untuk mengetahui peubah penjelas yang paling berpengaruh pada model. Pengujian parsial menggunakan uji Wald dengan hipotesis:

$$H_0: \beta_j = 0$$

$$H_1: \beta_j \neq 0, j = 1, 2, \dots, p$$
Tolak H_0 jika $W_j > \chi^2_{p(\alpha)}$ atau nilai-p $< \alpha$ yang artinya minimal ada satu peubah penjelas yang signifikan terhadap peubah respon pada taraf nyata 5%.
- d. Melakukan interpretasi terhadap koefisien yang terbentuk.
Interpretasi koefisien ditentukan berdasarkan rasio odds. Kemudian, dilakukan perbandingan antara gerombol empat dengan gerombol lainnya.

Melakukan evaluasi model

Evaluasi model dilakukan dengan menggunakan tabel ketepatan klasifikasi. Semakin besar tingkat ketepatan klasifikasi maka semakin tinggi tingkat akurasi model.

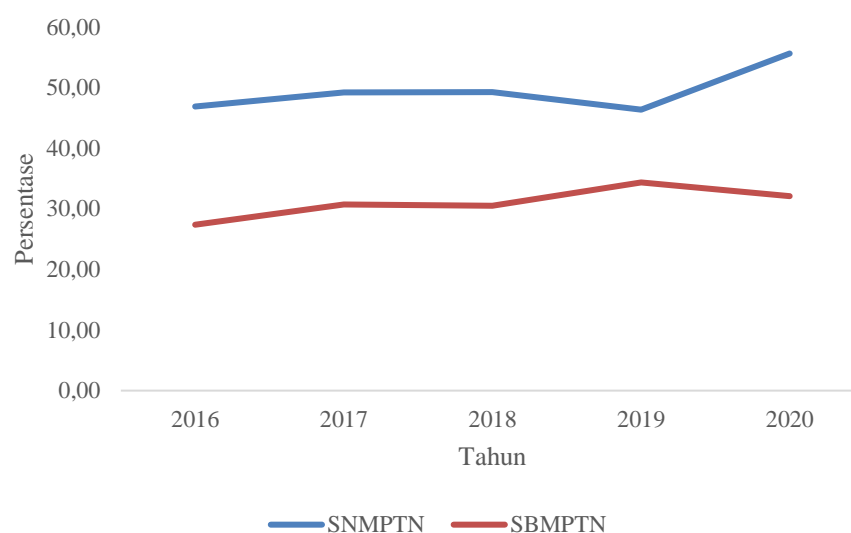
Melakukan validasi pada data *testing*

Validasi pada data *testing*, yaitu data tahun 2021 dilakukan untuk melihat apakah model yang telah dibuat pada tahun 2016 – 2020 bisa dipakai untuk mengukur kemampuan prediksi sekolah pada tahun berikutnya atau tidak.

IV HASIL DAN PEMBAHASAN

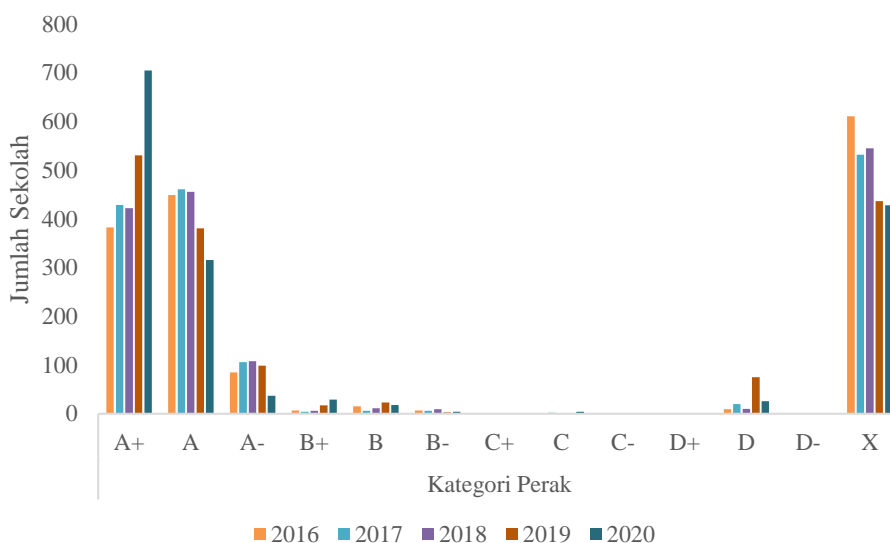
4.1 Deskripsi Data

Amatan dalam penelitian ini adalah sekolah peserta SNMPTN IPB University yang termasuk sekolah perak, yaitu sebanyak 1568 sekolah dari tahun ajaran 2015/2016 – 2020/2021. Sekolah perak merupakan pemeringkatan sekolah berdasarkan rekam jejak siswa dan sekolah yang diterima di IPB University. Kategori sekolah yang digunakan ada tiga, yaitu komitmen, kualitas, dan kekonsistenan sekolah. Kategori tersebut diwakilkan oleh peubah-peubah yang digunakan dalam penelitian ini untuk melihat apakah sekolah pendaftar bisa menjadi bagian dari civitas akademika IPB University. Mayoritas sekolah pendaftar SNMPTN IPB University adalah sekolah negeri yang memiliki akreditasi A dan berasal dari Pulau Jawa yang dapat dilihat pada Lampiran 2.



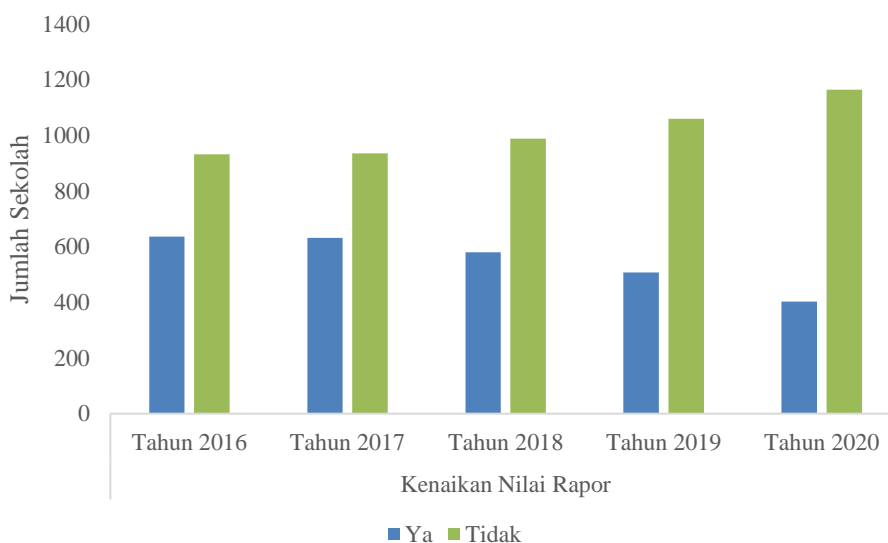
Gambar 2 Rata-rata persentase siswa mendaftar ulang SNMPTN dan SBMPTN

Komitmen sekolah dapat dilihat dari siswa yang mendaftar ulang SNMPTN atau SBMPTN di IPB University. Suatu sekolah dikatakan memiliki komitmen apabila siswa yang lolos SNMPTN atau SBMPTN di IPB University melakukan daftar ulang. Gambar 2 menunjukkan rata-rata persentase siswa yang mendaftar ulang baik SNMPTN maupun SBMPTN di IPB University. Rata-rata persentase siswa yang mendaftar ulang SNMPTN lebih tinggi dan meningkat pada tahun 2020. rata-rata persentase siswa yang mendaftar ulang SBMPTN lebih rendah dan menurun pada tahun 2020. Hal ini berarti komitmen sekolah pada jalur SNMPTN lebih tinggi dibandingkan dengan jalur SBMPTN.



Gambar 3 Kriteria sekolah berdasarkan kategori perak tahun 2016-2020

Gambar 3 menunjukkan mayoritas sekolah memiliki kategori perak X, A, dan A+. Sekolah dengan kategori perak X merupakan sekolah yang tidak terlibat pada jalur SNMPTN dan SBMPTN pada tahun tersebut. Selain itu, sekolah dengan kategori perak A+ dan A merupakan sekolah yang terlibat pada jalur SNMPTN dan SBMPTN serta memiliki predikat yang tinggi pada tahun tersebut.



Gambar 4 Kenaikan nilai rapor tahun 2016-2020

Nilai rapor siswa menjadi salah satu hal yang ada dalam penerimaan mahasiswa baru, khususnya jalur SNMPTN. Gambar 4 menunjukkan data kenaikan nilai rapor sekolah yang diterima di IPB University. Nilai rapor sekolah berdasarkan tahun sebelumnya mengalami penurunan untuk setiap tahunnya. Hal ini berarti sekolah tersebut belum cukup menjaga mutu sekolah karena sekolah yang baik adalah sekolah yang selalu menjaga mutu sekolahnya.

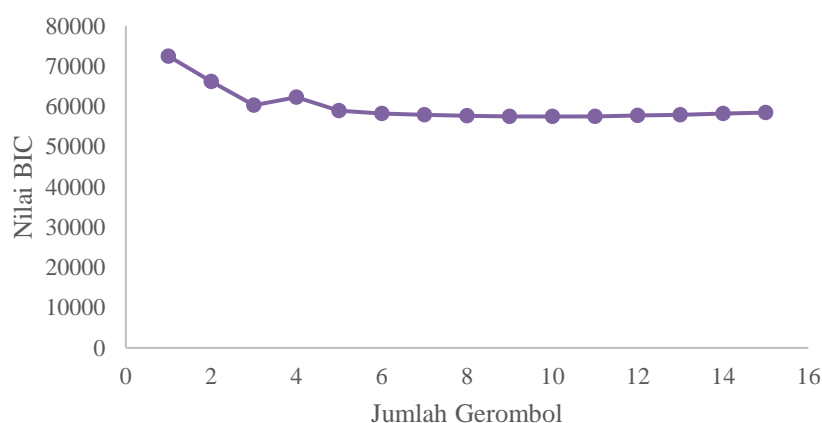
Tabel 4 Karakteristik sekolah berdasarkan peubah kategorik

No	Karakteristik	Kategori	Jumlah	Persentase(%)
1	Top 1000 Sekolah SNMPTN 2021	Ya	593	37,82
		Tidak	975	62,18
2	Kekonsistenan Sekolah	Ya	323	20,60
		Tidak	1245	79,40
3	Loyalitas Sekolah	Ya	1443	92,03
		Tidak	125	7,97

Sekolah yang termasuk ke dalam Top 1000 Sekolah SNMPTN 2021 memiliki persentase lebih kecil daripada yang tidak termasuk dalam Top 1000 Sekolah SNMPTN 2021 yaitu sebesar 37,82% atau sebanyak 593 sekolah. Kekonsistenan sekolah berasal dari penerimaan siswa pada jalur SNMPTN dan diperoleh persentase yang lebih kecil dibandingkan ke-tidak-konsistenan sekolah yaitu sebesar 20,60% atau sebanyak 323 sekolah. Hal ini berarti sekolah masih belum konsisten diterima di IPB University selama lima tahun berturut-turut. Semakin tinggi tingkatan sekolah yang diterima, semakin konsisten sekolah tersebut dalam menjaga mutu sekolahnya. Loyalitas sekolah dilihat dari intensitas suatu sekolah yang terlibat dalam pendaftaran jalur SNMPTN di IPB University baik yang diterima maupun tidak. Tabel 4 menunjukkan bahwa sebesar 92,03% atau sebanyak 1443 sekolah memiliki loyalitas tinggi terhadap IPB University.

4.2 Analisis Gerombol

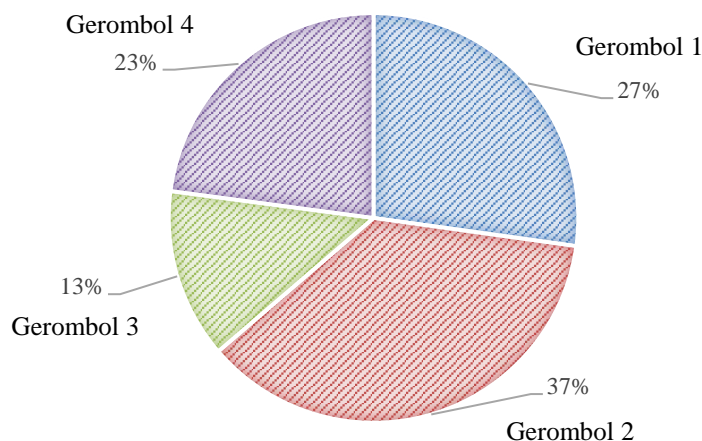
Penggerombolan sekolah dilakukan dengan menggunakan metode *Two-Step Cluster*. Penggerombolan tersebut ditentukan dengan memperhatikan nilai koefisien *Silhouette* untuk menghitung kualitas gerombol. Berdasarkan hasil penggerombolan, nilai koefisien *Silhouette* yang dihasilkan baik, yaitu sebesar 0,2 dengan empat gerombol optimal.



Gambar 5 Nilai BIC

Penentuan jumlah gerombol yang dihasilkan dapat dilihat dari nilai BIC terkecil. Gambar 5 menunjukkan nilai BIC mengalami penurunan untuk setiap jumlah gerombolnya. Oleh karena itu, penentuan jumlah gerombol optimal diperoleh menggunakan nilai rasio perubahan BIC dan nilai rasio perubahan jarak terbesar. Nilai rasio perubahan BIC digunakan untuk menentukan jumlah gerombol maksimum sedangkan nilai rasio perubahan jarak terbesar digunakan untuk menentukan jumlah optimal dari gerombol yang terbentuk.

Berdasarkan hasil penggerombolan yang terbentuk pada Lampiran 3, rasio jarak terbesar ada dua, yaitu dengan jumlah 2 gerombol sebesar 1,495 dan 4 gerombol sebesar 1,693. Hasil perbandingan dari kedua rasio jarak tersebut adalah $\left(\frac{R_{j1}}{R_{j2}}\right) = \left(\frac{1,495}{1,693}\right) = 0,883$. Karena hasil perbandingan rasio jarak kurang dari c_2 ($c_2 = 1,15$), maka penentuan jumlah gerombol optimal adalah $\max\{2,4\}$, yaitu empat gerombol.



Gambar 6 Ukuran gerombol yang dihasilkan

Gambar 6 menunjukkan persentase sekolah dari masing-masing gerombol yang dihasilkan. Gerombol 1 beranggotakan 27,2% atau sebanyak 426 sekolah, Gerombol 2 beranggotakan 36,5% atau sebanyak 573 sekolah, Gerombol 3 beranggotakan 13,4% atau sebanyak 210 sekolah, dan Gerombol 4 beranggotakan 22,9% atau sebanyak 359 sekolah.

Suatu kriteria dikatakan rendah jika rata-rata suatu peubah kontinu pada suatu gerombol lebih rendah dari rata-rata gabungan antar gerombol dan termasuk kategori dengan persentase terendah pada suatu gerombol. Suatu kriteria dikatakan tinggi jika rata-rata suatu peubah kontinu pada suatu gerombol lebih tinggi dari rata-rata gabungan antar gerombol dan termasuk kategori dengan persentase tertinggi pada suatu gerombol sesuai dengan Lampiran 4a dan 4b.

a. Gerombol 1 (Komitmen, kualitas, dan konsistensi rendah)

Kriteria Gerombol 1 merupakan sekolah yang memiliki komitmen rendah terhadap IPB University karena persentase siswa yang mendaftar ulang SNMPTN dan SBMPTN di IPB University berada di bawah rata-rata. Kekonsistenan dalam gerombol ini memiliki persentase yang rendah dalam mempertahankan siswanya untuk dapat diterima di IPB University. Sekolah pada gerombol ini juga memiliki kualitas yang cukup rendah seperti persentase siswa yang lolos SNMPTN ITB/ITS, siswa yang lolos SNMPTN

IPB University, dan siswa yang memperoleh $IPK \geq 3,5$. Sekolah pada gerombol ini juga tidak memiliki kategori perak sekolah, kenaikan nilai rapor, dan tidak termasuk sekolah pada Top 1000 sehingga sekolah yang termasuk dalam gerombol ini memiliki tingkat paling rendah. Sekolah yang memiliki komitmen, kualitas, dan konsistensi rendah cukup banyak, yaitu sebanyak 426 sekolah.

b. Gerombol 2

Sekolah yang tergolong gerombol ini merupakan sekolah yang memiliki komitmen yang cukup rendah pada jalur SNMPTN dan SBMPTN tetapi tidak lebih rendah dari Gerombol 1. Kualitas yang dimiliki pada gerombol ini juga cukup rendah tetapi persentase siswa yang lolos SNMPTN IPB University cukup tinggi dibandingkan persentase siswa yang lolos SNMPTN ITB/ITS dan siswa yang memperoleh $IPK \geq 3,5$ juga cukup tinggi. Kekonsistenan sekolah pada gerombol ini juga rendah sehingga sebanyak 573 sekolah yang termasuk gerombol ini memiliki peringkat rendah kedua setelah gerombol 1. Sekolah yang memiliki komitmen, kualitas, dan konsistensi cukup rendah pada gerombol ini paling banyak, yaitu sebanyak 573 sekolah.

c. Gerombol 3

Sekolah yang tergolong gerombol ini merupakan sekolah yang memiliki komitmen tinggi pada jalur SBMPTN tetapi cukup rendah pada jalur SNMPTN. Kualitas pada gerombol ini juga cukup tinggi dilihat dari persentase siswa yang lolos SNMPTN ITB/ITS, siswa yang memperoleh $IPK \geq 3,5$, dan termasuk sekolah dengan kategori A+. Namun, sekolah pada gerombol ini tidak termasuk dalam sekolah Top 1000. Konsistensi sekolah pada gerombol ini juga cukup rendah. Secara keseluruhan, sekolah yang termasuk dalam gerombol ini memiliki peringkat tertinggi kedua setelah Gerombol 4. Sekolah yang memiliki komitmen, kualitas, dan konsistensi cukup tinggi pada gerombol ini paling sedikit, yaitu sebanyak 210 sekolah.

d. Gerombol 4 (Komitmen, kualitas, dan konsistensi tinggi)

Sekolah yang tergolong gerombol ini merupakan sekolah yang memiliki komitmen, kualitas, dan kekonsistenan tinggi. Persentase yang dihasilkan dari gerombol ini berada diatas rata-rata, baik siswa yang mendaftar ulang SNMPTN dan SBMPTN di IPB University, siswa yang lolos SNMPTN dan SBMPTN di IPB University, siswa yang memperoleh $IPK \geq 3,5$, dan mayoritas sekolah pada gerombol ini adalah sekolah dengan kategori perak A+. Secara keseluruhan, sekolah pada gerombol ini merupakan sekolah yang memiliki peringkat tertinggi. Sekolah yang memiliki komitmen, kualitas, dan konsistensi tinggi cukup banyak, yaitu sebanyak 359 sekolah.

4.3 Analisis Regresi Logistik Multinomial

Analisis regresi logistic multinomial menghasilkan tiga model logit seperti pada Lampiran 5a, 5b, 5c, 5d, dan 5e. Untuk melihat peubah mana saja yang memiliki pengaruh besar dengan satu model logit, dilakukan menggunakan metode *forward logistic regression*. Metode *forward logistic regression* adalah langkah maju dimana peubah bebas dimasukkan satu persatu menurut urutan besar

pengaruhnya terhadap model, dan berhenti jika semua yang memenuhi syarat telah masuk. Peubah bebas yang berpengaruh nyata yang masuk ke dalam model ada 20 peubah, yaitu X3, X6, X11, X12, X13, X16, X18, X22, X24, X26, X27, X28, X29, X30, X33, X34, X35, X38, X39, dan X40. Hasil dari pemodelan ini akan membentuk model logit. Pengujian koefisien regresi menghasilkan nilai uji G, yaitu sebesar 2893,376 dan $\chi^2_{(0,05,43)} = 59,303$ yang artinya tolak H_0 , yaitu minimal ada satu peubah penjelas yang berpengaruh nyata terhadap peubah respon. Pengujian koefisien regresi berdasarkan uji parsial dengan uji Wald menyatakan tolak H_0 , yaitu minimal ada satu peubah penjelas yang signifikan terhadap peubah respon. Namun, ada satu peubah penjelas yang tidak signifikan pada taraf nyata 5%, yaitu X16 atau persentase siswa lolos SNMPTN di IPB pada tahun 2016 seperti pada Tabel 5.

Tabel 5 Dugaan parameter regresi logistik multinomial

Peubah bebas	$\hat{\beta}$	Galat baku	Wald	Nilai-p
Konstanta	0,469	0,060	61,324	$8,90 \times 10^{-15}$
X3	-0,001	0,000	6,843	0,009
X6	0,001	0,000	10,857	0,001
X11	0,002	0,001	5,004	0,025
X12	0,003	0,001	14,722	0,000
X13	0,005	0,001	27,573	$1,72 \times 10^{-7}$
X16	0,002	0,001	3,617	0,057*
X18	0,002	0,001	9,840	0,002
X22	-0,009	0,002	20,958	$5,08 \times 10^{-6}$
X24	-0,005	0,002	8,014	0,005
X26	0,018	0,003	39,312	$4,67 \times 10^{-10}$
X27	0,096	0,018	29,713	$5,81 \times 10^{-8}$
X28	0,009	0,003	9,941	0,002
X29	0,059	0,015	15,468	$8,76 \times 10^{-5}$
X30	0,091	0,003	1313,192	$< 2 \times 10^{-16}$
X33(1)	0,075	0,036	4,389	0,036
X34(1)	-0,076	0,027	7,695	0,006
X35	0,328	0,032	108,035	$< 2 \times 10^{-16}$
X38(1)	1,002	0,040	672,001	$< 2 \times 10^{-16}$
X39(1)	-0,136	0,045	8,976	0,003
X40	0,099	0,040	6,041	0,014

Keterangan: * nilai-p > 0,05

Hasil dari uji G dan uji Wald menyatakan bahwa ada 19 peubah hasil reduksi menggunakan metode *forward logistic regression* yang signifikan dan menghasilkan model logit sebagai berikut:

$$\begin{aligned}\hat{g}(x) = & 0.469 - 0.001X_3 + 0.001X_6 + 0.002X_{11} + 0.003X_{12} \\ & + 0.005X_{13} + 0.002X_{18} - 0.009X_{22} - 0.005X_{24} \\ & + 0.018X_{26} + 0.096X_{27} + 0.009X_{28} + 0.059X_{29} \\ & + 0.091X_{30} + 0.075X_{33(1)} - 0.076X_{34(1)} + 0.328X_{35} \\ & + 1.002X_{38(1)} - 0.136X_{39(1)} + 0.099X_{40}\end{aligned}$$

Tabel 6 Peluang prediksi sekolah masuk ke gerombol 4

Peubah	Sekolah A	Sekolah B	Sekolah C
X3	33,33	80	88,89
X6	33,33	66,67	79,17
X11	0,8	48,6	86,7
X12	1,3	48,8	88,9
X13	1,1	48,1	87,5
X18	1,3	38,9	77,8
X22	33,3	76,9	91,7
X24	33,3	83,3	91,7
X26	1	8	13
X27	1	8	13
X28	1	8	13
X29	1	8	13
X30	1	8	13
X33	1	1	1
X34	1	1	1
X35	0	1	1
X38	1	1	1
X39	1	1	1
X40	0	1	1
Peluang	0,66032	0,9442	0,98866

Berdasarkan model yang didapatkan, dugaan peluang sekolah masuk ke dalam kategori gerombol dapat dihitung. Tabel 6 menunjukkan peluang sekolah masuk ke dalam Gerombol 4. Sebagai contoh, sekolah A adalah sekolah yang memiliki kriteria rendah, sekolah B adalah sekolah yang memiliki kriteria cukup tinggi, dan sekolah C adalah sekolah yang memiliki kriteria tinggi dengan nilai seperti pada Tabel 6. Sekolah dengan kriteria rendah memiliki peluang masuk ke dalam Gerombol 4 yang relatif kecil. Sekolah dengan kriteria tinggi memiliki peluang masuk ke dalam Gerombol 4 yang relatif besar. Hal ini terjadi karena Gerombol 4 adalah gerombol yang memiliki komitmen, kualitas, dan konsistensi yang tinggi.

Tabel 7 Nilai odds rasio

Peubah bebas	Odds rasio
Konstanta	1,599
X3	0,999
X6	1,001
X11	1,002
X12	1,004

Peubah bebas	Odds rasio
X13	1,005
X16	1,002
X18	1,002
X22	0,991
X24	0,995
X26	1,018
X27	1,101
X28	1,009
X29	1,060
X30	1,096
X33(1)	1,078
X34(1)	0,927
X35	1,388
X38(1)	2,725
X39(1)	0,873
X40	1,104

Interpretasi koefisien dapat dilihat dari nilai odds rasionya. Nilai odds rasio dapat dilihat pada Tabel 7. Dilakukan perbandingan antara gerombol lain dengan Gerombol 4. Interpretasi dari peubah kenaikan nilai rapor pada tahun 2020 adalah kecenderungan siswa tidak mengalami kenaikan nilai rapor akan masuk ke gerombol lain lebih besar 1,388 kali daripada siswa yang mengalami kenaikan nilai rapor pada Gerombol 4.

Rasio odds pada peubah kekonsistenan sekolah sebesar 2,725. Kecenderungan kekonsistenan sekolah akan masuk ke gerombol lain lebih besar 2,725 kali daripada ke-tidak-konsistenan pada Gerombol 4. Rasio odds untuk peubah jenis sekolah kategori MA sebesar 1,104 menggambarkan bahwa kecenderungan sekolah MA masuk ke gerombol lain daripada Gerombol 4 adalah 1,104 kali dari sekolah SMA.

Tabel 8 Persentase ketepatan klasifikasi

Amatan	Dugaan				Persentase ketepatan (%)
	1	2	3	4	
1	426	0	0	0	100
2	0	569	1	3	99,30
3	1	1	206	2	98,09
4	0	6	1	351	98,04
Total					98,86

Tabel 8 menunjukkan persentase ketepatan klasifikasi model secara keseluruhan adalah sebesar 98,86%. Hasil pengklasifikasian berdasarkan Tabel 7 menunjukkan bahwa tidak terdapat kesalahan klasifikasi karena ketepatan klasifikasi yang didapatkan memiliki persentase yang tinggi. Hal ini berarti model yang dihasilkan dari pengklasifikasian sudah baik.

4.4 Validasi Model

Model yang didapatkan dari hasil klasifikasi sebelumnya akan dilakukan validasi untuk data tahun 2021. Validasi ini dilakukan untuk melihat apakah model ini bisa dipakai untuk memprediksi sekolah pada tahun berikutnya.

Tabel 9 Validasi data tahun 2021

Statistik	Nilai
<i>Accuracy</i>	0,618
<i>Kappa</i>	0,457
<i>Mean Sensitivity</i>	0,624
<i>Mean Specificity</i>	0,864

Tabel 9 menunjukkan hasil validasi pada data tahun 2021 memiliki akurasi sebesar 61,8% dengan nilai *kappa* sebesar 45,7% dan *mean sensitivity* sebesar 62,4%. Validasi ini memiliki nilai *mean specificity* yang cukup besar, yaitu 86,4%. Validasi model yang dilakukan menggunakan data tahun 2021 masih belum cukup bagus memprediksi karena nilai akurasi dikatakan bagus atau baik jika berada pada selang 80-110% (Gonzales dan Herrador 2007). Hal ini dikarenakan setiap tahun terjadi perubahan sekolah yang mendaftar SNMPTN di IPB University. Secara keseluruhan, model tahun 2016 – 2020 belum bisa dipakai untuk memprediksi sekolah pada tahun berikutnya sehingga perlu dilakukan penggerombolan ulang untuk tahun berikutnya.

V SIMPULAN DAN SARAN

5.1 Simpulan

Penggerombolan sekolah pada penerimaan mahasiswa baru jalur SNMPTN di IPB University pada tahun 2016-2020 menggunakan metode *Two-Step Cluster* menghasilkan empat gerombol optimal dengan nilai rata-rata koefisien *Silhouette* sebesar 0,2. Kriteria empat gerombol yang dihasilkan yaitu Gerombol 1 adalah sekolah dengan peringkat terendah (komitmen, kualitas, dan konsistensi rendah). Gerombol 2 adalah sekolah yang memiliki komitmen, kualitas, dan konsistensi yang cukup rendah tetapi tidak lebih rendah dari Gerombol 1. Gerombol 3 adalah sekolah yang memiliki komitmen, kualitas, dan konsistensi yang cukup tinggi pada jalur SBMPTN tetapi cukup rendah pada jalur SNMPTN. Gerombol 4 adalah sekolah yang memiliki peringkat tertinggi (komitmen, kualitas, dan konsistensi tinggi).

Peubah-peubah yang berpengaruh signifikan terhadap sekolah pada penerimaan mahasiswa baru jalur SNMPTN di IPB University yaitu persentase siswa mendaftar ulang SNMPTN dan SBMPTN IPB University, persentase siswa lolos SNMPTN ITB/ITS dan IPB University, persentase mahasiswa dengan IPK $\geq 3,5$, kategori perak sekolah, kenaikan rapor, kekonsistenan sekolah, loyalitas sekolah, dan jenis sekolah. Model regresi logistik multinomial memberikan ketepatan klasifikasi sebesar 98,86%. Peluang sekolah pada setiap kategori dapat diperoleh dari model logit yang terbentuk. Terdapat kategori atau nilai masing-masing peubah penjelas yang dapat memprediksi peluang sekolah masuk ke dalam masing-masing gerombol. Sekolah yang memiliki persentase siswa mendaftar ulang SNMPTN dan SBMPTN di IPB University, persentase mahasiswa dengan IPK $\geq 3,5$, loyalitas sekolah, dan jenis sekolah yang besar memiliki peluang tinggi untuk masuk dalam Gerombol 4 (Komitmen, kualitas, dan konsistensi tinggi).

Model yang dihasilkan dari hasil klasifikasi pada tahun 2016-2020 belum bisa dipakai untuk tahun berikutnya karena validasi yang dilakukan pada tahun 2021 memiliki nilai akurasi yang belum cukup bagus, yaitu sebesar 61,80%. Hal ini dikarenakan setiap tahun terjadi perubahan sekolah yang mendaftar SNMPTN di IPB University.

5.2 Saran

Saran yang perlu disampaikan oleh peneliti untuk penelitian selanjutnya yaitu perlu dilakukan penggerombolan dan pengklasifikasian ulang untuk tahun selanjutnya agar bisa memprediksi kembali sekolah pada penerimaan mahasiswa baru jalur SNMPTN di IPB University. Saran yang lain, yaitu bisa dicobakan menggunakan metode klasifikasi lain untuk penelitian selanjutnya.

DAFTAR PUSTAKA

- Aditya AR, Suparti, Sudarno. 2015. Ketepatan Klasifikasi Pemilihan Metode Kontrasepsi di Kota Semarang Menggunakan Bootstrap Aggregating Regresi Logistik Multinomial. *Jurnal Gaussian*. 4(1):11-20.
- Agresti A. 2019. *Categorical Data Analysis*. Ed ke-3. New Jersey (NJ): John Wiley and Sons.
- The SPSS TwoStep Cluster Component. A scalable component to segment your costumers more effectifely*. 2001. White paper-technical report, SPSS Inc Chicago.
- Chiu T, Fang D, Chen J, Wang Y, Jeris C. 2001. A Robust and Scalable Clustering Algorithm for Mixed Type Attributes in Large Database Environment. Di dalam: Jeris C, editor. The 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2001); 2001 Agt 26-29; San Francisco, United States. New York (NY): ACM Press. Hlm 263-264; [diunduh 2022 Feb 18]. <https://dl.acm.org/doi/10.1145/502512.502549>.
- Crystine A. 2018. Perbandingan Regresi Logistik Ordinal dan Diskriminan Kernel [tesis]. Bogor: IPB University.
- Dewantari NKM. 2021. Penggerombolan sekolah pada penerimaan mahasiswa baru jalur snmptn di IPB menggunakan metode two-step cluster. *Xplore*. 10(3):238-248.
- Gonzales AG, Herrador MA. 2007. A practical guide to analytical method validation, including measurement uncertainty and accuracy profiles. *Trends in Analytical Chemistry*. 26(3):227-238.
- Habibie N, Rizki A, Silvianti P. 2020. Penggerombolan hasil ujian nasional sma menggunakan k-rataan samar. *Xplore*. 10(1):41-54.
- Hair JF, Anderson RE, Babin BJ, Black WC. 2010. *Multivariate Data Analysis*. Volume ke – 7. New Jersey (US): Prentice Hall.
- Horng SJ, Su MY, Chen YH, Kao TW, Chen RJ, Lai JL, Perkasa CD. 2011. A novel intrusion detection system based on hierarchical clustering and support vector machines. *Science Direct*. 38(1):306-313.
- Hosmer DW, Lemeshow S. 2013. *Applied Logistic Regression*. Ed ke-3. New Jersey (NJ): John Wiley and Sons.
- Johnson RA, Wichern DW. 2002. *Applied Multivariate Statistical Analysis*. Ed ke-5. New Jersey: Prentice Hall.
- Kurniawan E. 2016. Metode Topsis untuk Menentukan Penerimaan Mahasiswa Baru Pendidikan Dokter di Universitas Muhammadiyah Purwokerto [tesis]. Purwokerto: Universitas Muhammadiyah Purwokerto.
- Lathifaturrahman. 2014. Perbandingan hasil penggerombolan k-means, fuzzy k-means, dan two step cluster. *JPM IAIN Antasari*. 2(1):39-62.
- [LTMPPT] Top 1000 Sekolah dengan Nilai Rerata TPS UTBK Tertinggi Tahun 2021. 2021.
- Mattjik AA, Sumertajaya IM. 2011. Sidik Peubah Ganda dengan Menggunakan SAS. Wibawa GNA, Hadi AF, editor. Bogor: IPB Press.
- Mongi CE. 2015. Penggunaan analisis *two-step clustering* untuk data campuran. *JdC*. 4(1):9-19.
- Purbowati D. 2022. Pemeringkatan sekolah dan siswa di SNMPTN. [diakses 2022

Juli 20]. <https://akupintar.id/info-pintar/-/blogs/pemeringkatan-sekolah-dan-siswa-di-snmptn>

Shyntia FD. 2022. Pendekatan metode chaid dan regresi logistik dalam menganalisis faktor berpengaruh pada kejadian stunting di provinsi Jawa Barat. *Xplore*. 11(1):48-58.

SPSS Inc. 2001. *The SPSS TwoStep Cluster Component. A scalable component enabling more efficient customer segmentation* [Technical Report]. New York.

Zahro ZZ, Zain I. 2018. Analisis regresi logistik multinomial pada faktor-faktor yang mempengaruhi sumber air bersih rumah tangga di Jawa Timur. *Jurnal Sains dan Seni ITS*. 7(2):2337-3520.

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
- b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

LAMPIRAN

@Hak cipta milik IPB University

IPB University



- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Lampiran 1a Peubah yang digunakan dan kategorinya

Peubah	Nama peubah	Kategori
Y	Komitmen, Kualitas, dan Konsistensi	1=Gerombol1, 2=Gerombol2, 3=Gerombol3, 4=Gerombol4
X ₁ – X ₅	Persentase Siswa Mendaftar Ulang SNMPTN IPB pada Tahun 2016 – 2020	-
X ₆ – X ₁₀	Persentase Siswa Mendaftar Ulang SBMPTN IPB pada Tahun 2016 – 2020	-
X ₁₁ – X ₁₅	Persentase Siswa Lolos SNMPTN XY pada Tahun 2016 – 2020	-
X ₁₆ – X ₂₀	Persentase Siswa Lolos SNMPTN IPB pada Tahun 2016 – 2020	-
X ₂₁ – X ₂₅	Persentase Mahasiswa dengan IPK ≥ 3.5 pada Tahun 2016 – 2020	-
X ₂₆ – X ₃₀	Kategori Perak Sekolah pada Tahun 2016 – 2020	Skor: 13=A+, 12=A, 11=A-, 10=B+, 9=B, 8=B-, 7=C+, 6=C, 5=C-, 4=D+, 3=D, 2=D-, 1=X
X ₃₁ – X ₃₅	Kenaikan Nilai Rapor pada Tahun 2016 – 2020	1=jika terdapat kenaikan rata-rata nilai rapor pada tahun tsb 0=jika tidak
X ₃₆	Akreditasi Sekolah pada Tahun 2016 – 2020	3=A, 2=B, 1=C, 0=lainnya
X ₃₇	Kepemilikan Sekolah Tahun 2016 – 2020	1= Negeri, 0=Swasta
X ₃₈	Kekonsistenan Sekolah pada Tahun 2016 – 2020	1=jika sekolah tsb diterima SNMPTN selama 5 tahun berturut-turut, 0=jika minimal 1 tahun siswa asal sekolah tsb tidak diterima SNMPTN IPB
X ₃₉	Loyalitas Sekolah pada SNMPTN 2016 – 2020	1=jika minimal 3 dari 5 tahun sekolah tsb mendaftar ke IPB, 0=lainnya
X ₄₀	Jenis Sekolah pada Tahun 2016 – 2020	1=SMA, 0=MA
X ₄₁	Keikutsertaan Sekolah pada Top 1000 Sekolah SNMPTN	1=jika masuk Top 1000 sekolah, 0=lainnya

@Hak cipta milik IPB University

Peubah	Nama peubah	Kategori
X ₄₂	Daerah Asal Sekolah	1=Pulau Sumatera, 2=Pulau Jawa, 3=Kepulauan Nusa Tenggara, 4=Pulau Kalimantan, 5=Pulau Sulawesi, 6=Kepulauan Maluku, 7=Pulau Papua

@Hak cipta milik IPB University

IPB University

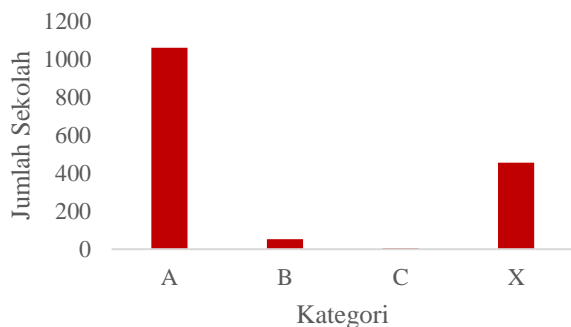


Lampiran 2 Kriteria sekolah berdasarkan peubah kategorik

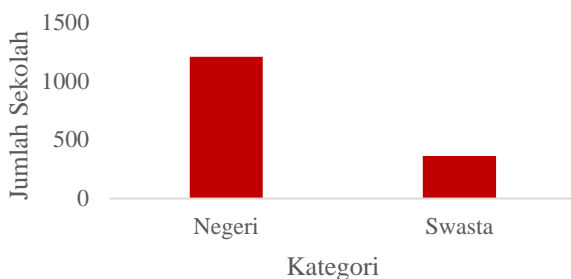
@Hak cipta milik IPB University

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

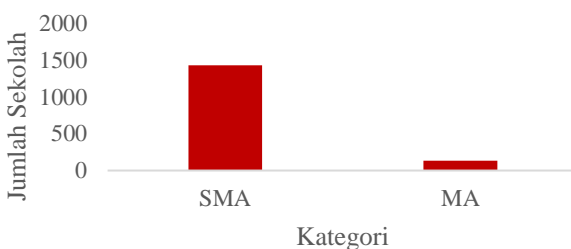
Akreditasi Sekolah Tahun 2016-2020



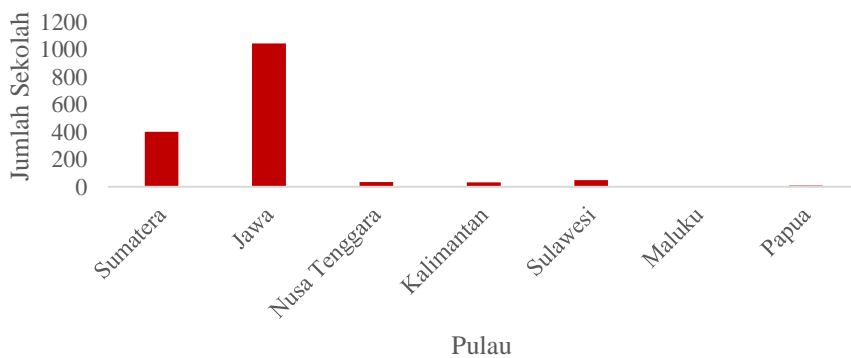
Kepemilikan Sekolah Tahun 2016-2020



Jenis Sekolah Tahun 2016-2020



Daerah Asal Sekolah Tahun 2016-2020



Lampiran 3 Hasil penggerombolan dengan nilai BIC

<i>Number of Clusters</i>	<i>Schwarz's Bayesian Criterion (BIC)</i>	<i>BIC Change^a</i>	<i>Ratio of BIC Change^b</i>	<i>Ratio of Distance Measures^c</i>
1	72486,176			
2	66191,974	-6294,202	1,000	1,495
3	60281,542	-1978,789	0,314	1,282
4	62260,331	-3931,643	0,625	1,693
5	58922,517	-1359,025	0,216	1,471
6	58266,871	-655,646	0,104	1,234
7	57894,714	-372,157	0,059	1,138
8	57669,593	-225,121	0,036	1,109
9	57549,126	-120,467	0,019	1,071
10	57492,619	-56,506	0,009	1,120
11	57531,907	39,287	-0,006	1,247
12	57729,541	197,634	-0,031	1,001
13	57927,611	198,070	-0,031	1,139
14	58203,700	276,089	-0,044	1,049
15	58505,847	302,147	-0,048	1,021

- a. *The changes are from the previous number of clusters in the table.*
b. *The ratios of changes are relative to the change for the two clusters solution.*
c. *The ratios of distance measures are based on the current number of clusters against the previous number of clusters.*

Lampiran 4a Kriteria penggerombolan

Kriteria Gerombol	Keterangan	Peubah	Pengukuran	Gerombol				
				1	2	3	4	Combined
Commitment	Persentase siswa yang mendaftar ulang SNMPTN IPB	X1	Mean	27,99	33,04	46,72	91,5	49,8125
		X2	Mean	31,44	35,56	44,85	94,51	51,59
		X3	Mean	25,8	42,47	34,98	96,34	49,8975
		X4	Mean	20,31	40,26	36,07	93,01	47,4125
		X5	Mean	3,05	76,32	38,25	95,11	53,1825
	Persentase siswa yang mendaftar ulang SBMPTN IPB	X6	Mean	10,8	16,41	59,43	45,78	33,105
		X7	Mean	17,66	20,03	57,58	47,57	35,71
		X8	Mean	17,88	21,29	49,89	48,87	34,4825
		X9	Mean	18	24,37	70,47	48,54	40,345
		X10	Mean	0,59	29,63	67,14	55,39	38,1875
Quality	Persentase siswa yang lolos SNMPTN ITB/ITS	X11	Mean	2,69	2,35	33,17	4,2	10,6025
		X12	Mean	3,47	2,72	41,44	5,33	13,24
		X13	Mean	4,16	2,5	41,35	6,62	13,6575
		X14	Mean	3,49	2,25	37,08	5,11	11,9825
		X15	Mean	11,5	6,01	72,82	15,77	26,525
	Persentase siswa yang lolos SNMPTN IPB	X16	Mean	4,82	6,82	4,62	23,25	9,8775
		X17	Mean	6,45	8,24	5,53	25,13	11,3375
		X18	Mean	7,29	10,06	3,77	25,43	11,6375
		X19	Mean	4,71	10,03	3,89	25,07	10,925
		X20	Mean	3,23	70,3	13,7	78,87	41,525
	Persentase mahasiswa yang memperoleh IPK >= 3,5	X21	Mean	37,93	43,45	79,37	97,22	64,4925
		X22	Mean	45,66	49,84	81,05	97,33	68,47
		X23	Mean	41,14	56,12	68,92	98,15	66,0825
		X24	Mean	39,24	59,82	84,07	95,47	69,65
		X25	Mean	0,12	97,28	82,31	95,94	68,9125
	Kategori Perak Sekolah	X26	Persentase	60,80%	54,50%	52,40%	49,90%	
				X	X	A+	A+	
		X27	Persentase	52,60%	46,80%	56,70%	54,60%	
				X	X	A+	A+	
		X28	Persentase	56,80%	42,10%	42,40%	61,00%	
				X	X	A+	A+	
		X29	Persentase	51,60%	34,40%	64,30%	64,60%	
	Kenaikan Nilai Rapor			X	X	A+	A+	
		X30	Persentase	94,60%	47,30%	67,10%	81,60%	
				X	A+	A+	A+	
		X31	Persentase	73,20%	68,60%	61,90%	73,00%	
				Tidak	Tidak	Tidak	Ya	
		X32	Persentase	70,70%	68,60%	56,70%	65,70%	
				Tidak	Tidak	Tidak	Ya	

@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Lampiran 4b Kriteria penggerombolan

Kriteria Gerombol	Keterangan	Peubah	Pengukuran	Gerombol			
				1	2	3	4
Quality	Kenaikan Nilai Rapor	X33	Persentase	77,70%	63,40%	75,70%	62,40%
				Tidak	Tidak	Tidak	Ya
		X34	Persentase	81,90%	68,20%	72,40%	53,20%
				Tidak	Tidak	Tidak	Ya
	Keikutsertaan Sekolah pada Top 1000 Sekolah 2020	X35	Persentase	99,80%	80,10%	90,50%	74,70%
				Tidak	Tidak	Tidak	Ya
Consistency	Kekonsistenan Sekolah	X41	Persentase	67,10%	61,60%	55,20%	61,30%
				Tidak	Tidak	Tidak	Tidak
Kriteria Umum	Akreditasi Sekolah	X38	Persentase	97,20%	99,50%	90,50%	80,20%
				Tidak	Tidak	Tidak	Konsisten
	Kepemilikan Sekolah	X36	Persentase	99,80%	94,90%	81,40%	95,50%
				X	A	A	A
	Loyalitas Sekolah	X37	Persentase	70,20%	75,20%	70,00%	91,90%
				Negeri	Negeri	Negeri	Negeri
	Jenis Sekolah	X39	Persentase	83,80%	92,80%	92,90%	100,0%
				Loyal	Loyal	Loyal	Loyal
	Daerah Asal sekolah	X40	Persentase	88,50%	88,80%	96,70%	95,80%
				SMA	SMA	SMA	SMA
		X42	Persentase	50,90%	66,80%	81,90%	75,80%
				Jawa	Jawa	Jawa	Jawa

@Hak cipta milik IPB University

Lampiran 5a Dugaan parameter regresi logistik multinomial

Peubah bebas	$\hat{\beta}$	Galat baku	Wald	Nilai-p	Rasio odds
Y = 2					
Gerombol 2)					
Konstanta	-6,858	1,503	20,802	$5,06 \times 10^{-6}$	$1,05 \times 10^{-3}$
X1	0,003	0,017	0,027	0,867	1,002
X2	0,002	0,018	0,011	0,914	1,001
X3	-0,003	0,018	0,025	0,871	0,997
X4	-0,010	0,016	0,279	0,596	0,991
X5	0,031	0,022	1,915	0,166	1,031
X6	-0,004	0,011	0,131	0,716	0,995
X7	-0,008	0,010	0,638	0,424	0,991
X8	$-4,43 \times 10^{-3}$	0,011	0,148	0,699	0,995
X9	0,003	0,011	0,065	0,798	1,003
X10	-0,003	0,011	0,073	0,785	0,996
X11	0,017	0,025	0,435	0,508	1,016
X12	-0,018	0,025	0,509	0,475	0,981
X13	-0,022	0,024	0,808	0,368	0,977
X14	-0,026	0,026	0,946	0,330	0,974
X15	-0,024	0,012	3,751	0,052	0,976
X16	-0,023	0,032	0,504	0,477	0,977
X17	0,002	0,036	0,004	0,946	1,002
X18	-0,016	0,030	0,261	0,608	0,984
X19	-0,013	0,032	0,151	0,697	0,987
X20	-0,017	0,026	0,440	0,506	0,982
X21	-0,002	0,017	0,015	0,900	0,997
X22	0,018	0,013	1,723	0,189	1,017
X23	0,008	0,014	0,281	0,595	1,007
X24	-0,012	0,015	0,636	0,424	0,987
X25	0,062	0,024	6,687	$9,69 \times 10^{-3}$	1,064
X26 (3)	0,318	1,678	0,035	0,849	1,375
X26 (5)	1,028	0,183	31,482	$2,00 \times 10^{-8}$	2,794
...
X27 (3)	-1,016	1,681	0,364	0,545	0,362
X27 (5)	-0,389	0,021	331,312	0,000	0,677
...
X28 (3)	0,057	2,861	0,000	0,984	1,058
X28 (5)	2,875	0,010	69760,695	0,000	17,731
...
X29 (3)	-0,809	1,832	0,914	0,658	0,445
X29 (5)	-0,415	$3,88 \times 10^{-3}$	11446,004	0,000	0,660
...
X30 (3)	-2,294	0,165	191,490	0,000	0,100
X30 (6)	5,121	1,025	24,950	$5,85 \times 10^{-7}$	167,514

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Lampiran 5b Dugaan parameter regresi logistik multinomial

Peubah bebas	$\hat{\beta}$	Galat baku	Wald	Nilai-p	Rasio odds
...
X31 (1)	-0,672	1,332	0,254	0,614	0,510
X32 (1)	-1,255	1,483	0,714	0,397	0,285
X33 (1)	0,100	1,336	0,005	0,940	1,105
X34 (1)	1,201	1,235	0,944	0,330	3,323
X35 (1)	-0,459	1,389	0,108	0,741	0,632
X36 (1)	4,871	1,085	20,151	$7,14 \times 10^{-6}$	130,497
X36 (2)	6,592	1,188	30,780	$2,88 \times 10^{-8}$	728,940
X36 (3)	5,150	0,996	26,718	$2,34 \times 10^{-7}$	172,383
X37 (1)	-0,064	0,999	0,004	0,948	0,937
X38 (1)	-5,364	1,703	9,903	0,001	0,004
X39 (1)	2,537	1,399	1,397	0,069	12,637
X40 (1)	-0,481	1,172	0,168	0,681	0,617
X41 (1)	-0,321	0,804	0,159	0,689	0,725
X42 (2)	0,750	0,892	0,705	0,400	2,116
X42 (3)	-0,911	1,779	0,262	0,608	0,401
...
Y = 3 (Gerombol 3)					
Konstanta	-12,116	1,773	46,635	$8,5 \times 10^{-12}$	$5,47 \times 10^{-6}$
X1	0,007	0,017	0,167	0,682	1,007
X2	-0,012	0,018	0,469	0,492	0,987
X3	-0,016	0,018	0,717	0,396	0,984
X4	-0,011	0,017	0,427	0,513	0,988
X5	-0,002	0,021	0,017	0,894	0,997
X6	0,010	0,011	0,837	0,360	1,010
X7	-0,006	0,010	0,372	0,541	0,993
X8	$-5,29 \times 10^5$	0,011	$1,6 \times 10^{-5}$	0,996	0,999
X9	0,004	0,012	0,154	0,693	1,004
X10	-0,003	0,012	0,097	0,754	0,996
X11	0,039	0,019	4,247	0,039	1,040
X12	0,047	0,022	4,541	0,033	1,048
X13	0,030	0,017	3,020	0,082	1,030
X14	0,026	0,022	1,456	0,227	1,027
X15	0,013	0,010	1,700	0,192	1,013
X16	-0,033	0,040	0,715	0,397	0,966
X17	-0,013	0,041	0,104	0,745	0,986
X18	-0,008	0,032	0,060	0,806	0,991
X19	0,011	0,037	0,094	0,758	1,011
X20	0,011	0,026	0,001	0,964	1,001
X21	0,027	0,018	2,310	0,128	1,027
X22	0,024	0,012	3,892	0,048	1,025

Lampiran 5c Dugaan parameter regresi logistik multinomial

Peubah bebas	$\hat{\beta}$	Galat baku	Wald	Nilai-p	Rasio odds
X23	0,005	0,013	0,194	0,659	1,005
X24	-0,011	0,015	0,614	0,453	0,988
X25	0,011	0,024	22,972	$1,63 \times 10^{-6}$	1,126
X26 (3)	0,414	2,289	0,032	0,856	1,512
X26 (5)	0,114	0,000	242057,11	0,000	1,121
...
X27 (3)	-0,884	1,981	0,198	0,655	0,412
X27 (5)	0,314	0,000	1511633,3	0,000	1,121
...
X28 (3)	-0,432	1,252	0,119	0,729	0,648
X28 (5)	-1,291	0,007	31939,408	0,000	0,274
...
X29 (3)	-0,739	1,626	0,206	0,649	0,477
X29 (5)	0,282	$2,88 \times 10^{-5}$	95671306	0,000	1,326
...
X30 (3)	0,682	2,358	0,083	0,772	1,978
X30 (6)	0,850	1,321	0,413	$5,20 \times 10^{-1}$	2,340
...
X31 (1)	-1,013	1,295	0,611	0,433	0,362
X32 (1)	0,452	1,479	0,093	0,759	1,571
X33 (1)	-0,536	1,369	0,152	0,695	0,585
X34 (1)	-0,156	1,325	0,013	0,906	0,855
X35 (1)	-0,909	1,487	0,373	0,541	0,402
X36 (1)	2,087	1,441	2,096	$1,47 \times 10^{-1}$	8,065
X36 (2)	0,306	1,574	0,037	$8,45 \times 10^{-1}$	1,358
X36 (3)	-3,504	1,334	6,791	$9,14 \times 10^{-3}$	0,030
X37 (1)	-0,369	1,047	0,123	0,724	0,691
X38 (1)	-1,733	1,583	1,196	0,273	0,176
X39 (1)	0,802	1,383	0,335	0,562	2,230
X40 (1)	1,092	1,291	0,715	0,397	2,982
X41 (1)	0,419	0,799	0,274	0,600	1,520
X42 (2)	1,530	0,922	2,784	0,097	4,621
X42 (3)	0,297	1,683	0,030	0,859	1,347
...
Y = 4 (Gerombol 4)					
Konstanta	-12,548	1,601	61,418	$8,5 \times 10^{-12}$	$3,54 \times 10^{-6}$
X1	0,011	0,017	0,483	0,682	1,012
X2	0,001	0,018	0,006	0,492	1,001
X3	0,003	0,018	0,833	0,396	1,003
X4	0,000	0,016	0,000	0,513	1,000
X5	0,014	0,021	0,488	0,894	1,014
X6	0,007	0,011	0,459	0,360	1,007
X7	-0,003	0,010	0,138	0,541	0,996

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Lampiran 5d Dugaan parameter regresi logistik multinomial

Peubah bebas	$\hat{\beta}$	Galat baku	Wald	Nilai-p	Rasio odds
X8	$6,734 \times 10^{-3}$	0,011	0,346	0,996	1,006
X9	0,001	0,012	0,173	0,693	1,001
X10	-0,000	0,011	0,005	0,754	0,999
X11	0,001	0,027	0,005	0,039	1,002
X12	0,019	0,031	0,383	0,033	1,019
X13	0,016	0,025	0,440	0,082	1,016
X14	-0,000	0,030	0,000	0,227	0,999
X15	-0,007	0,013	0,331	0,192	0,992
X16	-0,008	0,032	0,071	0,397	0,991
X17	0,000	0,035	0,000	0,745	1,000
X18	0,024	0,031	0,622	0,806	1,024
X19	-0,018	0,032	0,310	0,758	0,981
X20	0,008	0,025	0,098	0,964	1,008
X21	0,017	0,018	0,861	0,128	1,017
X22	-0,003	0,015	0,063	0,048	0,996
X23	0,014	0,015	0,893	0,659	1,014
X24	0,030	0,017	3,017	0,453	1,031
X25	0,110	0,025	18,725	$1,63 \times 10^{-6}$	1,117
X26 (3)	0,434	2,550	0,028	0,856	1,544
X26 (5)	-0,459	0,025	331,058	0,000	0,631
...
X27 (3)	1,530	1,400	1,194	0,655	4,621
X27 (5)	0,071	0,000	17928,138	0,000	1,047
...
X28 (3)	-0,442	1,377	0,103	0,729	0,642
X28 (5)	-1,448	0,004	125499,43	0,000	0,234
...
X29 (3)	-0,307	2,222	0,019	0,649	0,735
X29 (5)	-0,044	$8,13 \times 10^{-3}$	29,735	0,000	0,956
...
X30 (3)	2,300	2,408	0,912	0,772	9,979
X30 (6)	1,265	1,102	1,315	$5,20 \times 10^{-1}$	3,546
...
X31 (1)	0,095	1,248	0,005	0,433	1,100
X32 (1)	-0,229	1,376	0,027	0,759	0,795
X33 (1)	0,455	1,274	0,127	0,695	1,577
X34 (1)	0,508	1,163	0,190	0,906	1,663
X35 (1)	1,217	1,385	0,70	0,541	3,377
X36 (1)	1,124	1,109	1,028	$1,47 \times 10^{-1}$	3,079
X36 (2)	-1,362	1,154	1,392	$8,45 \times 10^{-1}$	0,255
X36 (3)	-3,419	1,383	6,110	$9,14 \times 10^{-3}$	0,032
X37 (1)	0,089	1,120	0,005	0,724	1,093
X38 (1)	0,843	1,425	0,349	0,273	2,324
X39 (1)	-2,918	1,881	2,405	0,562	0,054

Lampiran 5e Dugaan parameter regresi logistik multinomial

Peubah bebas	$\hat{\beta}$	Galat baku	Wald	Nilai-p	Rasio odds
X40 (1)	1,519	1,278	1,411	0,397	4,570
X41 (1)	-0,373	0,807	0,213	0,600	0,688
X42 (2)	0,787	0,914	0,741	0,097	2,198
X42 (3)	-0,478	1,781	0,071	0,859	0,619
...

Hak cipta milik IPB University

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

RIWAYAT HIDUP

Penulis dilahirkan di kota Metro pada tanggal 1 November 2001 sebagai anak pertama dari pasangan bapak M. Sulthon Sulaiman dan ibu Dian Purnamasari. Pendidikan sekolah menengah atas (SMA) ditempuh di sekolah SMA Negeri 5 Palembang, dan lulus pada tahun 2018. Pada tahun 2018, penulis diterima sebagai mahasiswa program sarjana (S-1) di Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Pertanian Bogor.

Selama mengikuti program S-1, penulis pernah menjadi staff bidang Sosial dan Masyarakat KSR PMI, sekretaris 1 Organisasi Mahasiswa Daerah IKAMUSI, staff divisi Liaison Officer Statistika Ria 14, dan bendahara divisi Kestari Kompetisi Statistika Junior 2021. Penulis juga pernah melakukan Praktik Lapang di Rapido Research pada tahun 2021.

@Hak cipta milik IPB University

IPB University

