



## Analisis Diskriminan Terhadap Angka Partisipasi Kasar Perguruan Tinggi berdasarkan Faktor Pendidikan dan Sosial Ekonomi

Muhammad Arsyad Alzam<sup>1</sup>, Shifa Elmaliyasari<sup>2</sup>, Nanda Aulia Pratiwi<sup>3</sup>, Aviolla Terza Damaliana<sup>4</sup>, Shindi Shella May Wara<sup>5</sup>

<sup>2, 3, 4, 5</sup> Program Studi Sains Data, Universitas Pembangunan "Veteran" Jawa Timur, Indonesia

<sup>1</sup>[23083010082@student.upnjatim.ac.id](mailto:23083010082@student.upnjatim.ac.id)

<sup>2</sup>[23083010022@student.upnjatim.ac.id](mailto:23083010022@student.upnjatim.ac.id)

<sup>3</sup>[23083010011@student.upnjatim.ac.id](mailto:23083010011@student.upnjatim.ac.id)

<sup>4</sup>[aviolla.terza.sada@upnjatim.ac.id](mailto:aviolla.terza.sada@upnjatim.ac.id)

<sup>5</sup>[shindi.shella.fasilkom@upnjatim.ac.id](mailto:shindi.shella.fasilkom@upnjatim.ac.id)

**Abstract:** The level of participation in higher education or GER, reflects the key to developing superior and competitive human resources. GER serves as a primary measure to evaluate the effectiveness of education equity policies, considering socioeconomic factors such as income level, poverty rate, and education budget allocation. In addition, GER is influenced by educational factors, particularly the number of higher education institutions, the percentage of college graduates, and the completion rate of senior secondary education. The secondary data used in this study was sourced from the Central Statistics Agency (BPS) and covers all provinces in Indonesia. The objective is to identify characteristics that differentiate provinces based on GER levels. The primary method used is discriminant analysis, which enables classification modeling based on key educational variables. The results show that the percentage of college graduates and the senior secondary completion rate are the most influential in distinguishing provincial GER levels. The resulting discriminant function is:  $\hat{D} = -8.217 + 0.394X_3$  (percentage of college graduates) +  $0.055X_4$  (senior secondary completion rate), with a classification accuracy of 81,6%. These findings provide an essential basis for designing more targeted higher education policies to address disparities in access across regions.

**Keywords:** GER, Education, Socio-Economics, Indicator, Higher Education

**Abstrak:** Tingkat partisipasi dalam pendidikan tinggi atau APK, mencerminkan kunci dalam pengembangan SDM yang unggul dan kompetitif. APK menjadi acuan utama untuk mengevaluasi efektivitas pemerataan pendidikan, dengan mempertimbangkan faktor sosial ekonomi seperti pendapatan, kemiskinan, dan anggaran pendidikan. Selain itu, APK juga dipengaruhi oleh faktor pendidikan, terutama jumlah perguruan tinggi, persentase tamat perguruan tinggi, dan tingkat penyelesaian SMA/ sederajat. Data sekunder yang digunakan dalam penelitian ini bersumber dari Badan Pusat Statistik (BPS) dan mencakup semua provinsi di Indonesia. Tujuan penelitian ini adalah mengidentifikasi karakteristik yang membedakan provinsi berdasarkan pencapaian APK. Metode utama yang digunakan adalah analisis diskriminan, yang mampu memodelkan klasifikasi berdasarkan faktor pendidikan. Hasil penelitian menunjukkan bahwa persentase tamat perguruan tinggi dan tingkat penyelesaian SMA/ sederajat memiliki pengaruh terbesar dalam membedakan provinsi. Fungsi diskriminan yang dihasilkan adalah:  $\hat{D} = -8.217 + 0.394X_3$  (persentase tamat perguruan tinggi) +  $0.055X_4$  (tingkat penyelesaian SMA/ sederajat), dengan akurasi klasifikasi sebesar 81,6%. Temuan ini menjadi dasar penting dalam merancang kebijakan pendidikan tinggi yang lebih tepat sasaran untuk mengatasi ketimpangan akses antar wilayah.

**Kata kunci:** APK, Pendidikan, Sosial Ekonomi, Indikator, Perguruan Tinggi

### I. PENDAHULUAN

Kontribusi pendidikan tinggi sangat signifikan dalam membentuk individu kerja yang berkualitas dan daya saing yang bagus di tingkat global. Keberhasilan pembangunan pendidikan dapat diukur melalui indikator Angka Partisipasi Kasar (APK), yaitu proporsi warga yang sedang mengenyam pendidikan pada level tertentu relatif terhadap jumlah penduduk dalam golongan umur resmi jenjang tersebut [1]. APK menjadi salah satu parameter untuk mengukur efektivitas kebijakan pemerataan pendidikan di berbagai daerah. Beberapa faktor sosial ekonomi menjadi determinan penting terhadap APK, di antaranya adalah kemiskinan, pendapatan per kapita, jumlah perguruan tinggi, dan anggaran pendidikan [1]. Selain itu, kesejahteraan sosial ekonomi keluarga terlihat berdampak positif terhadap



semangat siswa dalam melanjutkan pendidikan ke perguruan tinggi [2]. Hambatan sosial ekonomi seringkali membatasi akses pendidikan, meskipun siswa memiliki potensi akademik yang tinggi.

Penelitian oleh Mukhaiyar et al. menunjukkan bahwa APK jenjang SMA/ sederajat dipengaruhi secara signifikan oleh jumlah sekolah, tingkat kemiskinan, dan besarnya anggaran pendidikan, meski tidak semuanya signifikan secara parsial [3]. Meskipun penelitian tersebut berfokus pada jenjang menengah, implikasinya terhadap pendidikan tinggi tetap relevan. Dalam menganalisis data sosial ekonomi, analisis diskriminan sering digunakan sebagai metode statistik untuk tujuan klasifikasi. Metode ini membentuk fungsi diskriminan berdasarkan kombinasi linear variabel prediktor untuk mengklasifikasikan objek ke dalam golongan tertentu [4]. Sebagai salah satu metode klasifikasi, analisis diskriminan digunakan untuk membedakan objek berdasarkan karakteristik tertentu. Dalam penelitian ini, nilai APK dikategorikan menjadi dua golongan, yaitu “APK Tinggi” yang diberi kode 1 dan “APK Rendah” yang diberi kode 0, guna mempermudah proses klasifikasi dalam analisis diskriminan.

Dalam konteks analisis multivariat, metode discriminant analysis menawarkan pendekatan yang tepat untuk mengelompokkan daerah berdasarkan tingkat partisipasi pendidikan tinggi dengan mempertimbangkan variabel sosial ekonomi. Teknik ini memungkinkan identifikasi variabel prediktor yang paling berkontribusi terhadap pembentukan golongan yang berbeda secara statistik, serta menyediakan fungsi klasifikasi yang dapat digunakan dalam pengambilan kebijakan pendidikan. Namun, penelitian yang secara khusus menerapkan analisis diskriminan untuk mengklasifikasikan tingkat APK pendidikan tinggi antarprovinsi di Indonesia berdasarkan faktor pendidikan dan sosial ekonomi masih terbatas. Oleh karena itu, penelitian ini bertujuan untuk menerapkan analisis diskriminan dalam mengklasifikasikan provinsi-provinsi di Indonesia berdasarkan tingkat APK pendidikan tinggi. Temuan dalam penelitian ini diharapkan dapat memberikan bukti empiris yang bermanfaat untuk mendukung perumusan kebijakan pemerataan akses pendidikan tinggi yang lebih tepat sasaran.

## II. METODE PENELITIAN

### 2.1. Sumber Data

Data Angka Partisipasi Kasar (APK) dipilih yang asalnya dari BPS tahun 2023 kemudian mencakup seluruh 38 provinsi di Indonesia.

Variabel	Definisi	Skala Pengukuran
Jumlah Perguruan Tinggi	Total universitas di suatu provinsi	Rasio
Rata-rata Biaya Perguruan Tinggi	Rata - rata biaya kuliah per semester	Rasio
Persentase Tamat Perguruan Tinggi	Proporsi penduduk yang sudah menyelesaikan pendidikan tinggi	Rasio
Tingkat Penyelesaian SMA/ sederajat	Persentase penduduk yang lulus SMA/SMK/MA	Rasio
Persentase Penduduk Miskin	Proporsi penduduk di bawah garis kemiskinan	Rasio
Tingkat Pengangguran Terbuka	Persentase angkatan kerja yang tidak bekerja tetapi aktif mencari kerja	Rasio
Rata-rata Pendapatan Bersih	Rata-rata pendapatan rumah tangga setelah pajak	Rasio
Angka Partisipasi Kasar Perguruan Tinggi	Indikator yang mengukur persentase penduduk usia kuliah (19-23 tahun) yang sedang menempuh pendidikan tinggi	Kategorik

Tabel 1. Variabel Data

Pada tabel 1, Angka Partisipasi Kasar (APK) merupakan variabel yang dipengaruhi ( $Y$ ) bertugas mengukur persentase penduduk dalam rentang usia 19-23 tahun yang sedang mengenyam pendidikan tinggi. Adapun faktor yang berdampak seperti jumlah perguruan tinggi, rata-rata biaya perguruan tinggi, tingkat penyelesaian SMA/Sederajat, dan persentase penduduk miskin yang berperan sebagai variabel bebas ( $X$ ). Faktor-faktor bebas ini dapat mempengaruhi tingkat keterlibatan dalam pendidikan tinggi di provinsi tertentu

## 2.2. Metode Analisis

Analisis diskriminan termasuk dalam metode regresi di mana variabel dependennya bersifat kategorik atau non-metrik. Teknik ini digunakan untuk menghitung skor komposit berdasarkan variabel-variabel pada sampel, guna mengelompokkan sampel ke dalam kategori tertentu [5]. Penerapan analisis diskriminan dalam studi ini memiliki peran krusial dalam mengelompokkan provinsi-provinsi di Indonesia berdasarkan tingkat APK pendidikan tinggi, dengan mempertimbangkan variabel sosial ekonomi seperti tingkat kemiskinan, pendapatan, dan alokasi anggaran pendidikan. Dengan metode ini, peneliti bisa mengetahui faktor mana yang paling berpengaruh dalam membedakan tingkat partisipasi pendidikan antar daerah. Untuk mencapai tujuan penelitian, dilakukan serangkaian tahapan yang sistematis sesuai dengan metode analisis diskriminan. Setiap langkah bertujuan untuk memastikan validitas hasil klasifikasi berdasarkan faktor sosial ekonomi terhadap tingkat partisipasi pendidikan tinggi. Adapun tahapan-tahapan dalam alur penelitian dapat dijelaskan sebagai berikut:

### 1. Pengumpulan Data

Temuan ini memanfaatkan data yang asalnya dari Badan Pusat Statistik (BPS) Republik Indonesia melalui website resmi [www.bps.go.id](http://www.bps.go.id). Data ini berisikan indikator pendidikan tinggi di 38 provinsi Indonesia tahun 2023.

### 2. Persiapan Data

Tahap persiapan data mencakup proses pengolahan data APK Perguruan Tinggi serta indikator-indikator sosial ekonomi di suatu daerah. Proses ini dimulai dengan memastikan data lengkap dan menangani nilai yang hilang. Selanjutnya, data dinormalisasi untuk memiliki skala yang sama, lalu dibagi menjadi dua kategori: "Rendah" (APK di bawah 37%) dan "Tinggi" (APK di atas 37%).

### 3. Uji Asumsi Diskriminan

Sebelum melakukan analisis diskriminan, diperlukan beberapa uji untuk terpenuhinya beberapa asumsi penting. Ketiga uji yang dilakukan sebagai berikut:

#### a. Uji Normalitas

Pengujiannya agar tahu data ini sudah berdistribusi normal atau belum, khususnya dalam konteks distribusi normal multivariat. Salah satu teknik yang diaplikasikan adalah uji Mardia, yang menilai dua elemen penting: kemiringan dan keruncingan. Dengan menggunakan uji ini, kita dapat menentukan apakah data cenderung mengikuti distribusi normal atau jika terdapat deviasi dari pola tersebut. Berikut adalah hipotesis serta rumus sistematis dari uji yang dilakukan:

$H_0$ : Data termasuk distribusi normal multivariat

$H_1$ : Data tidak termasuk distribusi normal multivariat

- Skewness

$$b_{M,1} = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \left[ \left( x_i - \bar{x} \right)^T S^{-1} \left( x_j - \bar{x} \right) \right]^3 \quad (1)$$

Keterangan:

$x_i, x_j$  = Vektor observasi ke- $i$  dan ke- $j$

$\bar{x}$  = Vector rerata

$s^{-1}$  = Invers matriks kovarian sampel

$T$  = Transpose

$n$  = Jumlah observasi



Dengan Pengambilan Keputusan:

$H_0$  ditolak jika nilai skewness  $>$  z-tabel pada taraf signifikan tertentu, misalnya  $\alpha = 0,05$

- Kurtosis

$$b_{M,2} = \frac{1}{n^2} \sum_{i=1}^n \left[ \left( x_i - \bar{x} \right)^T S^{-1} \left( x_j - \bar{x} \right) \right]^2 \quad (2)$$

Dengan Pengambilan Keputusan:

$H_0$  ditolak jika nilai kurtosis berbeda signifikan dari nilai kurtosis normal.

### b. Uji Homogenitas

Uji homogenitas adalah teknik statistik yang dipakai untuk tahu apakah dua atau lebih golongan sampel memiliki varians yang seragam. Salah satu metode yang sering dipakai itu uji Box's M, pengembangan dari uji Bartlett. Uji ini mengevaluasi kesamaan matriks kovariansi di antara golongan dan memberikan kemudahan lebih jika dibandingkan dengan uji normalitas multivariat. Proses analisisnya dilakukan dengan memanfaatkan distribusi F untuk menguji asumsi tentang kesamaan varians di antara berbagai populasi [5]. Berikut adalah hipotesis serta rumus sistematisnya:

$H_0$ : Matriks kovarian antar golongan adalah homogen.

$H_1$ : Matriks kovarian antar golongan tidak homogen.

- Statistik Uji Box's M

$$M = (n - k) \log |S| - \sum_{i=1}^k (n_i - 1) \log |S_i| \quad (3)$$

Keterangan:

$n$  = Jumlah total observasi

$k$  = Jumlah golongan

$|S|$  = Determinan dari matriks kovarian gabungan

$|S_i|$  = Determinan dari matriks kovarian golongan ke- $i$

$n_i$  = Jumlah data pada golongan ke- $i$

Dengan Pengambilan Keputusan:

$H_0$  ditolak ketika  $p\text{-value} < 0,05$ , makanya matriks kovarian antar golongan tidak homogen

### c. Uji Independensi

Tujuan dari uji Bartlett adalah untuk menentukan apakah tiga atau lebih golongan data memiliki varians yang homogen. Uji ini paling akurat jika ukuran sampel antar golongan sama, dan tetap bisa digunakan untuk ukuran berbeda meskipun kurang presisi. Namun, uji ini sangat sensitif terhadap ketidaknormalan data, sehingga uji normalitas perlu dilakukan sebelumnya [6]. Berikut adalah hipotesis serta rumusnya:

$H_0$ : Variabel independen tidak saling berhubungan.

$H_1$ : Variabel independen menunjukkan hubungan signifikan.

- Variansi Gabungan



$$S_p^2 = \frac{\sum_{i=1}^k (n_i - 1) S_i^2}{N - k} \quad (4)$$

Keterangan:

$S_i^2$  = Varians golongan ke-i

$n_i$  = Jumlah data di golongan ke-i

$N$  = Jumlah total data

$k$  = Jumlah golongan

$S_p^2$  = Varians gabungan dari semua golongan

- Statistik Uji Barlett

$$h = \frac{\left( (S_1^2)^{n_1-1} \cdot (S_2^2)^{n_2-1} \cdot \dots \cdot (S_k^2)^{n_k-1} \right)^{1/(N-k)}}{S_p^2} \quad (5)$$

Keterangan:

Rumus statistik uji barlett membandingkan variasi antar golongan dengan varians gabungan untuk tahu homogenitas varians. Keputusan diambil dengan menolak  $H_0$  apabila  $p\text{-value} < 0,05$ , yang melihat adanya perbedaan yang nyata antar varians (tidak homogen).

#### 4. Analisis Diskriminan

Untuk memperoleh model klasifikasi yang optimal serta memahami faktor-faktor yang membedakan antar golongan, dilakukan serangkaian tahapan dalam analisis diskriminan. Adapun langkah-langkah yang dilakukan meliputi:

- a. Pemilihan Variabel Prediktor Analisis Diskriminan

Tahapan awal dalam analisis diskriminan adalah mengidentifikasi variabel prediktor yang memiliki relevansi serta signifikansi secara statistik. Fungsi diskriminan dinyatakan dalam bentuk: [7]

$$\hat{D}(X) = b_0 + b_1x_1 + b_2x_2 + \dots + b_kx_k \quad (6)$$

Keterangan:

$\hat{D}(X)$  = Nilai diskriminan

$b_0$  = Konstanta

$b_1, b_2, \dots, b_k$  = Koefisien diskriminan

$x_1, x_2, \dots, x_k$  = Variabel predictor

- b. Estimasi Koefisien Fungsi Diskriminan

Setelah variabel prediktor ditentukan, tahap berikutnya adalah menyusun fungsi diskriminan. Koefisien diskriminan dapat dibentuk seperti:

$$\hat{b} = S_W^{-1}(\bar{x}_1 - \bar{x}_2) \quad (7)$$

Keterangan:

$S_W^{-1}$  = Matriks kovarian dalam golongan

$\bar{x}_1, \bar{x}_2$  = Vektor rata-rata masing-masing golongan

- c. Identifikasi Titik Sentroid Golongan

Titik sentroid ini mencerminkan nilai rata-rata dari fungsi diskriminan setiap golongan dan berperan sebagai panduan dalam proses pengelompokan.

$$C = \frac{D_1 + D_2}{2} \quad (8)$$

Dengan Pengambilan Keputusan:

- Jika  $D(X) > CD(X) \rightarrow$  masuk golongan 1 (misalnya APK tinggi)
- Jika  $D(X) \leq CD(X) \rightarrow$  masuk golongan 2 (APK rendah)

### 5. Evaluasi Model

Model diskriminan yang telah dikembangkan kemudian dianalisis untuk mengetahui tingkat keakuratannya menggunakan matriks klasifikasi atau *confusion matrix* dan *cross validation*. [7]

#### a. Confusion Matrix

$$Akurasi = \frac{\text{Jumlah klasifikasi benar}}{\text{Total seluruh data}} = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

Matriks ini menunjukkan jumlah observasi yang:

- Diklasifikasikan dengan benar (True Positive dan True Negative)
- Diklasifikasikan dengan salah (False Positive dan False Negative)

#### b. Uji Validasi Silang

$$CV_{accuracy} = \frac{1}{k} \sum_{i=1}^k Accuracy \quad (10)$$

Keterangan:

$k$  = Jumlah lipatan (fold)

$Error_i$  = Nilai error (misalnya MSE, MAE, atau akurasi) pada fold ke- $i$

$CV_{error}$  = Nilai rata-rata error dari seluruh fold

### 6. Interpretasi Hasil

Hasil analisis menunjukkan faktor-faktor sosial ekonomi yang menyebabkan perbedaan Angka Partisipasi Kasar (APK) perguruan tinggi, seperti pendapatan rumah tangga, tingkat pendidikan orang tua, dan jenis pekerjaan.

## III. HASIL DAN PEMBAHASAN

### 3.1. Statistika Deskriptif

Statistik deskriptif digunakan untuk menggambarkan distribusi data dan mengenali pola-pola yang muncul. Rincian nilai statistik deskriptif untuk indikator ini tercantum pada tabel 2.

Variabel	Rata-rata	Minimum	Maksimum
Jumlah Perguruan Tinggi	154.58	16	734
Rata-rata Biaya Perguruan Tinggi	16002.43	9280	31270
Persentase Tamat Perguruan Tinggi	11.4582	3.47	18.23
Tingkat Penyelesaian SMA/ sederajat	65.8124	39.50	89.69
Persentase Penduduk Miskin	10.6647	3.80	29.66
Tingkat Pengangguran Terbuka	4.3797	1.32	6.75
Rata-rata Pendapatan Bersih	2091.0697	1209.50	3360.50

Tabel 2. Statistika Deskripsi Data

Berdasarkan tabel 2 menunjukkan variabel jumlah perguruan tinggi ( $X_1$ ) memiliki rata-rata 154,58, dengan rentang besar antara 16 hingga 734 institusi, menunjukkan distribusi yang tidak merata. Rata-rata biaya perguruan tinggi ( $X_2$ ) juga bervariasi cukup tajam, yakni antara Rp9.280.000 hingga Rp31.270.000, dengan nilai rata-rata Rp16.002.000,43. Persentase penduduk yang menamatkan pendidikan tinggi ( $X_3$ ) tergolong rendah, rata-rata hanya 11,46%, sedangkan tingkat penyelesaian SMA/ sederajat ( $X_4$ ) lebih tinggi dengan rata-rata 65,81%, namun tetap menunjukkan wilayah yang tertinggal. Persentase penduduk miskin ( $X_5$ ) rata-rata 10,66% dengan rentang cukup lebar, dari 3,80%

hingga 29,66%. Tingkat pengangguran terbuka ( $X_6$ ) tergolong sedang, dengan rata-rata 4,38%, dan rata-rata pendapatan bersih ( $X_7$ ) sebesar Rp2.091.000,07 juga menunjukkan kesenjangan yang cukup besar, berkisar antara Rp1.209.000,50 hingga Rp3.360.000,50.

Klasifikasi APK	Jumlah	Proporsi(%)
APK Tinggi	12	31.58
APK Rendah	26	68.42

**Tabel 3.** Proposi APK

Tabel 3 menunjukkan bahwa sebagian besar provinsi yaitu sebesar 62,5% berada dalam kategori APK Rendah, sedangkan hanya 37,5% yang termasuk APK Tinggi. Hal ini mencerminkan masih adanya tantangan dalam pemerataan akses pendidikan di wilayah tersebut.

### 3.2. Uji Asumsi

Sebelum analisis diskriminan dilakukan, perlu diuji beberapa asumsi dasar, yaitu normalitas multivariat, homogenitas matriks kovarian, dan independensi antar variabel. Uji yang digunakan meliputi Mardia untuk normalitas, Box's M untuk homogenitas, dan Bartlett's Test untuk independensi.

#### 1. Uji Normal Multivariat

Uji Normalitas bertujuan untuk memastikan bahwa asumsi normalitas terpenuhi sebelum dilakukan analisis diskriminan. Adapun hipotesis untuk uji normalitas:

$H_0$ : Data faktor-faktor pendidikan dan sosial ekonomi antarprovinsi berdistribusi normal secara multivariat.

$H_1$ : Data faktor-faktor pendidikan dan sosial ekonomi antarprovinsi tidak berdistribusi normal secara multivariat.

Statistika Uji	P-value	Keputusan
Skewness: 21.1265	1.0000	Gagal Tolak $H_0$
Kurtosis: -56.1842	1.0000	

**Tabel 4.** Hasil Uji Normal Multivariat Mardia

Hasil uji normalitas multivariat dengan metode Mardia pada tabel 4 menunjukkan nilai *skewness* sebesar 21.1265 dan *kurtosis* -56.1842, dengan *p-value* masing-masing 1.0000. Karena *p-value* lebih besar dari 0.05, maka  $H_0$  gagal ditolak. Dengan demikian, tidak terdapat cukup bukti untuk menyatakan bahwa data menyimpang dari distribusi normal. Maka dari itu, dapat disimpulkan bahwa data terkait faktor pendidikan dan sosial ekonomi antarprovinsi memiliki distribusi normal secara multivariat.

#### 2. Uji Homogenitas

Uji homogenitas bertujuan untuk memastikan bahwa matriks kovarian dari setiap golongan adalah sama, yang merupakan salah satu asumsi penting sebelum dilakukan analisis diskriminan. Adapun hipotesis untuk uji homogenitas:

$H_0$ : Matriks kovarian dari variabel-variabel pendidikan dan sosial ekonomi antar golongan provinsi adalah homogen.

$H_1$ : Matriks kovarian dari variabel-variabel pendidikan dan sosial ekonomi antar golongan provinsi tidak homogen.

Statistika Uji	Df	P-value	Keputusan
Box'M: -81.1745	28	1.0000	Gagal Tolak $H_0$

**Tabel 5.** Hasil Uji Homogenitas Box's M

Berdasarkan tabel 5 hasil uji homogenitas dengan metode Box's M, diperoleh nilai statistik uji sebesar 81.1745 dengan derajat kebebasan (*df*) 28 dan *p-value* sebesar 1.0000. Karena *p-value* jauh melebihi 0.05, maka keputusan uji menunjukkan gagal menolak  $H_0$ . Maknanya, tidak terdapat perbedaan yang signifikan dalam matriks kovarian antar golongan, sehingga dapat

disimpulkan bahwa matriks kovarian dari variabel-variabel pendidikan dan sosial ekonomi antar golongan provinsi adalah homogen.

### 3. Uji Independensi

Uji independensi bertujuan untuk mengetahui apakah variabel-variabel independen saling berkorelasi, sehingga layak untuk dianalisis lebih lanjut dalam analisis multivariat seperti analisis faktor. Adapun hipotesis untuk uji independensi:

$H_0$ : Tidak terdapat hubungan signifikan antar variabel pendidikan dan sosial ekonomi yang digunakan untuk mengklasifikasikan provinsi berdasarkan APK.

$H_1$ : Terdapat hubungan signifikan antar variabel pendidikan dan sosial ekonomi yang digunakan untuk mengklasifikasikan provinsi berdasarkan APK.

Statistika Uji	Df	P-value	Keputusan
Barlett: 56.6130	21	0.0000	Tolak $H_0$

Tabel 6. Hasil Uji Independensi Barlett

Berdasarkan tabel 6 hasil uji Bartlett, diperoleh nilai statistik uji sebesar 56.6130 dengan derajat kebebasan ( $df$ ) 21 dan  $p$ -value sebesar 0.0000. Mengacu  $p$ -value yang jauh di bawah 0.05, keputusan pengujian yang diambil adalah menolak  $H_0$ . Maknanya, terdapat perbedaan yang signifikan dalam varians antar variabel, Kesimpulannya, terdapat keterkaitan yang bermakna antara variabel. pendidikan dan sosial ekonomi yang digunakan untuk mengklasifikasikan provinsi berdasarkan APK.

### 3.3. Model Analisis Diskriminan

Proses awal dalam menyusun fungsi diskriminasi adalah menetapkan indikator yang sesuai, yang dipilih menggunakan metode *stepwise*. Metode ini secara otomatis memilih variabel berdasarkan kontribusi statistiknya dalam membedakan kelompok, sehingga hanya variabel yang paling signifikan yang dimasukkan ke dalam model.

Variabel	Wilks' Lambda	F-statistic	P-value
$X_3$	0.622	19.461	0
$X_4$	0.541	13.130	0

Tabel 7. Variabel dalam Model Analisis

Tabel 7 menunjukkan hasil analisis diskriminan terhadap kategori Angka Partisipasi Kasar (APK) tiap provinsi dengan acuan pencapaian target nasional, dua variabel yang signifikan dalam membedakan golongan provinsi adalah Persentase Tamat Perguruan Tinggi dan Tingkat Penyelesaian SMA/ sederajat. Nilai Wilks' Lambda masing-masing sebesar 0.622 dan 0.541 menunjukkan bahwa kedua variabel tersebut mampu menjelaskan variasi antar golongan dengan cukup baik.

Setelah melalui proses seleksi variabel, hanya variabel-variabel yang memiliki kontribusi signifikan terhadap fungsi diskriminan yang dimasukkan ke dalam model. Berikut adalah koefisien fungsi diskriminan dari variabel terpilih.

Variabel	Koefisien Fungsi
$X_3$	0.394
$X_4$	0.055
Constant	-8.217

Tabel 8. Parameter Model Analisis

Melalui nilai bobot fungsi diskriminan dari Tabel 8, maka dapat disusun fungsi diskriminan yaitu:

$$\hat{D} = -8.217 + 0.394X_3 + 0.055X_4$$

dengan:



$X_3$  = Persentase Tamat Perguruan Tinggi

$X_4$  = Tingkat Penyelesaian SMA/ sederajat

Setelah diperoleh fungsi diskriminan, langkah selanjutnya adalah menghitung nilai fungsi centroid untuk masing-masing kategori APK. Nilai ini menunjukkan posisi rata-rata masing-masing golongan dalam fungsi diskriminan dan digunakan sebagai acuan klasifikasi.

Klasifikasi APK	Function
APK Tinggi	-0.659
APK Rendah	1.209

Tabel 9. Fungsi Centroid

Nilai rata-rata skor fungsi diskriminan untuk provinsi yang memenuhi target APK nasional adalah -0.659, sedangkan untuk provinsi yang tidak memenuhi target adalah 1.209. Ini menunjukkan bahwa terdapat perbedaan yang jelas antara kedua golongan, dan model dapat memisahkan keduanya berdasarkan skor diskriminan. Setelah model klasifikasi diterapkan, langkah selanjutnya adalah mengevaluasi tingkat ketepatan hasil klasifikasinya.

Kelas Aktual	Kelas Prediksi		Total
	APK Tinggi	APK Rendah	
APK Tinggi	21	5	26
APK Rendah	2	10	12
APK Rendah (%)	80.8	19.2	100.0
APK Rendah (%)	16.7	83.3	100.0

Tabel 10. Hasil Klasifikasi

Hasil evaluasi validitas model diskriminan melalui tabel klasifikasi menunjukkan bahwa model memiliki tingkat akurasi yang cukup baik. Dari 26 provinsi yang sebenarnya memiliki APK Memenuhi Target, sebanyak 21 provinsi (80,8%) berhasil diklasifikasikan dengan benar, sementara 5 provinsi (19,2%) salah diklasifikasikan sebagai APK Tidak Memenuhi Target. Sedangkan dari 12 provinsi dengan APK Tidak Memenuhi Target, 10 provinsi (83,3%) diklasifikasikan dengan benar, dan hanya 2 provinsi (16,7%) yang salah diklasifikasikan.

Secara keseluruhan, model diskriminan mampu mengklasifikasikan data dengan baik, dengan total 31 dari 38 provinsi berhasil diklasifikasikan secara benar. Hal ini menunjukkan bahwa tingkat akurasi model secara keseluruhan adalah sebesar 81,6%, yang diperoleh dari perbandingan jumlah klasifikasi benar terhadap total observasi ( $31/38 \times 100\%$ ). Nilai akurasi ini mengindikasikan bahwa model memiliki kemampuan prediktif yang cukup tinggi dalam membedakan antara provinsi yang APK-nya memenuhi target dan yang tidak.

#### IV. KESIMPULAN

Penelitian ini menunjukkan bahwa Angka Partisipasi Kasar (APK) pendidikan tinggi di Indonesia dipengaruhi oleh faktor pendidikan, khususnya persentase tamat perguruan tinggi dan tingkat penyelesaian SMA/ sederajat. Faktor Pendidikan tersebut terbukti memiliki pengaruh yang dominan dalam membedakan provinsi berdasarkan tingkat APK. Dengan menggunakan metode analisis diskriminan, diperoleh model klasifikasi yang mampu memisahkan provinsi ke dalam kategori APK tinggi dan rendah secara akurat. Fungsi diskriminan yang dihasilkan adalah:  $\hat{D} = -8.217 + 0.394X_3$  (persentase tamat perguruan tinggi) +  $0.055X_4$  (tingkat penyelesaian SMA/ sederajat), dengan tingkat akurasi klasifikasi sebesar 81,6%. Temuan ini memperkuat pentingnya peningkatan capaian pendidikan menengah dan tinggi sebagai strategi utama dalam pemerataan akses pendidikan tinggi.



## REFERENSI

1. R. D. A. Putri, “DETERMINAN ANGKA PARTISIPASI KASAR PENDIDIKAN TINGGI PADA 14 PROVINSI DI INDONESIA PROPOSAL SKRIPSI,” 2024.
2. M. Hasan, M. Ihsan Said, T. Supatminingsih, and T. Tahir, “PENGARUH SOSIAL EKONOMI ORANG TUA TERHADAP MOTIVASI SISWA DALAM MELANJUTKAN PENDIDIKAN KE PERGURUAN TINGGI,” 2022.
3. U. Mukhaiyar, F. Rontos, K. Handoko, and S. Kardiyan, “Analisis Faktor-Faktor yang Memengaruhi Angka Partisipasi Kasar SMA/Sederajat di Indonesia Menggunakan Regresi Ridge,” *Euler : Jurnal Ilmiah Matematika, Sains dan Teknologi*, vol. 10, no. 2, pp. 222–234, Nov. 2022, doi: 10.34312/euler.v10i2.15903.
4. R. Ariefudin *et al.*, “Analisis Diskriminan untuk Klasifikasi Tingkat Kemiskinan di Perkotaan Menurut Provinsi Berdasarkan Bagian Wilayah di Indonesia Tahun 2022 Discriminant Analysis for Classification of Poverty Levels in Urban Areas by Provinces Based on Regions in Indonesia in 2022,” 2023.
5. M. D. Purnama, C. D. Farrel, S. F. Wahyuniarti, I. P. Nuraini, and A. Oktaviarina, “Analisis Perbedaan Data Komponen Kesehatan Bayi Provinsi Jawa Timur Tahun 2022 Menggunakan MANOVA,” 2023.
6. Usmadi, “PENGUJIAN PERSYARATAN ANALISIS (UJI HOMOGENITAS DAN UJI NORMALITAS),” *Inovasi Pendidikan*, vol. 7, no. 1, pp. 50–62, Mar. 2020.
7. N. Syafiq Ibrahim, “Analisis Diskriminan Linear Robust dengan Penduga Minimum Covariance Determinant (Studi Kasus: Indeks Kerentanan Pangan Menurut Kabupaten/Kota di Indonesia Tahun 2023),” *Emerging Statistics and Data Science Journal*, vol. 2, no. 2, 2024.
8. R. S. Lubis, R. Aprilia, A. Wardani, and S. Harleni, “PENENTUAN FAKTOR-FAKTOR POTENSIAL YANG MEMPENGARUHI KEMISKINAN DI SUMATERA UTARA MENGGUNAKAN ANALISIS DISKRIMINAN,” *Jurnal Serunai Matematika*, vol. 15, no. 2, pp. 55–60, Oct. 2023.
9. N. Annisafiya, D. A. Kusuma, and B. N. Ruchjana, “Penerapan Analisis Diskriminan untuk Klasifikasi Pengaruh Data Warisan Budaya Takbenda terhadap Banyaknya Wisatawan Domestik,” *Jurnal Matematika Integratif*, vol. 19, no. 2, p. 149, Dec. 2023, doi: 10.24198/jmi.v19.n2.46791.149-161.
10. A. Sofia, “Pengelompokan Kesejahteraan Berdasarkan Faktor-Faktor Sosial Ekonomi Terhadap Klasifikasi Indeks Pembangunan Masyarakat di Provinsi Aceh Menggunakan Analisis Diskriminan,” *Original Article Indonesian Journal of Applied Mathematics*, vol. 4, no. 1, pp. 11–18, 2024, doi: 10.35472/indoja.
11. N. Nurhasanah, N. Salwa, L. Ornilla, F. AR, and A. Hasan, “Analisis Diskriminan dalam Menentukan Fungsi Pengelompokan Kabupaten/Kota di Indonesia berdasarkan Indikator Indeks Pembangunan Manusia,” *Jurnal EMT KITA*, vol. 5, no. 1, p. 37, Apr. 2021, doi: 10.35870/emt.v5i1.320.
12. S. Lorinda, “MODEL REGRESI DATA PANEL ANGKA PARTISIPASI PENDIDIKAN TINGGI BERDASARKAN ANGKA PARTISIPASI SEKOLAH DAN ANGKA PARTISIPASI KASAR TAHUN 2018-2021 DI INDONESIA,” 2024.
13. Y. Suryana and R. S. Pradana, “ANGKA PARTISIPASI KASAR PERGURUAN TINGGI, SMA DAN SMP MENURUT PROVINSI DAN MENURUT GENDER SERTA HUBUNGANNYA DENGAN TINGKAT PENGANGGURAN TERBUKA,” *Jurnal Kewidyaiswaraan*, vol. 8, no. 1, pp. 29–37, 2023.
14. F. Virdam, M. Bernadette, and N. Ariani, “ANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI ANGKA PARTISIPASI SEKOLAH PADA PROVINSI DI PULAU SULAWESI,” *Jurnal Of Development Economic And Digitalization*, vol. 2, no. 1, pp. 20–35, 2023.
15. N. P. N. Hendayanti and M. Nurhidayati, “KLASIFIKASI TINGKAT KEPARAHAN KEMISKINAN PROVINSI DI INDONESIA DENGAN ANALISIS DISKRIMINAN,” *Math Educa Journal*, vol. 5, no. 1, pp. 14–21, Apr. 2021.