

Penerapan Analisis Diskriminan untuk Klasifikasi Pengaruh Data Warisan Budaya Takbenda terhadap Banyaknya Wisatawan Domestik

NADIRA ANNISAFIYA, DIANNE AMOR KUSUMA, BUDI NURANI RUCHJANA

Departemen Matematika, Fakultas MIPA, Universitas Padjadjaran
Jl. Raya Bandung Sumedang KM 21 Jatinangor Sumedang 45363

Email: nadira19002@mail.unpad.ac.id, amor@unpad.ac.id, budi.nurani@unpad.ac.id

Abstrak

Analisis diskriminan adalah analisis multivariat yang dapat digunakan untuk menentukan hubungan dependensi antara variabel independen berupa data kuantitatif dan variabel dependen berupa data kualitatif. Warisan Budaya Takbenda (WBTb) dapat diasumsikan sebagai variabel independen yang memiliki hubungan dengan banyaknya wisatawan domestik yang datang ke suatu provinsi sebagai variabel dependen. WBTb terancam punah dan perlu diperhatikan, salah satu cara untuk melestarikannya adalah melalui sektor pariwisata. Semakin banyak WBTb yang dimiliki suatu provinsi diharapkan semakin banyak wisatawan domestik yang berkunjung ke provinsi tersebut. Penelitian ini mengkaji model klasifikasi menggunakan analisis diskriminan untuk banyaknya wisatawan domestik dengan kategori tinggi dan rendah yang dipengaruhi oleh variabel WBTb dengan lima kategori, yaitu : (1) Adat Istiadat Masyarakat, Ritus, dan Perayaan-perayaan (AIMRP); (2) Kemahiran dan Kerajinan Tradisional (KKT); (3) Pengetahuan dan Kebiasaan Perilaku Mengenai Alam dan Semesta (PKPMAS); (4) Seni dan Pertunjukan (SP); (5) Tradisi dan Ekspresi Lisan (TEL). Hasil penelitian menunjukkan bahwa dari kelima kategori tersebut, variabel AIMRP, PKPMAS, dan TEL merupakan kategori yang dominan memengaruhi tinggi-rendahnya banyaknya wisatawan domestik dengan tingkat akurasi klasifikasi sebesar 96,57%.

Kata kunci: analisis diskriminan, Warisan Budaya Takbenda, banyaknya wisatawan domestik, variabel dominan

Abstract

Discriminant analysis is a multivariate analysis that can be used to determine the dependency relationship between independent variables in the form of quantitative data and dependent variables in the form of qualitative data. Intangible Cultural Heritage can be assumed as an independent variable that has a relationship with the number of domestic tourists coming to a province as a dependent variable. Intangible Cultural Heritage is endangered and needs to be noted, one way to preserve it is through the tourism sector. The more Intangible Cultural Heritage owned by a province, it is expected that more domestic tourists will visit the province. This study examines the classification model using discriminant analysis for the number of domestic tourists with high and low categories influenced by Intangible Cultural Heritage variables with five categories, namely: (1) Community Customs, Rites, and Celebrations; (2) Traditional Skills and Crafts; (3) Knowledge and Behavioral Habits Regarding Nature and the Universe; (4) Arts and Performing Arts; (5) Oral Tradition and Expression. The results showed that of the five categories, AIMRP, PKPMAS, and TEL variables were the dominant categories affecting the high and low number of domestic tourists with a classification accuracy rate of 96.57%.

Keywords: *discriminant analysis, Intangible Cultural Heritage, large number of domestic tourists, dominant variables*

1. PENDAHULUAN

Warisan budaya dihasilkan dari pengetahuan masyarakat yang telah dipelajari dan dilaksanakan sebagai aktivitas secara turun temurun dari generasi ke generasi, kemudian berkembang menjadi berbagai bentuk budaya yang dikelompokkan menjadi warisan budaya benda (*tangible culture*) dan WBTb (*intangible culture*). Warisan budaya benda menjadi perhatian utama dunia terutama pada bangunan-bangunan yang menunjukkan hasil dari peradaban manusia, sedangkan WBTb belum menjadi perhatian dan ditetapkan menjadi warisan budaya dunia. Perbedaan tersebut menjadi perhatian para pemerhati budaya untuk menjadikan WBTb menjadi warisan budaya dunia juga. Oleh karena itu, *The United Nations Educational, Scientific, and Cultural Organization* (UNESCO) sebagai organisasi yang memiliki salah satu tugas untuk menjaga budaya dari berbagai negara turut mengatur WBTb dalam konvensi 2003 UNESCO yang kemudian Indonesia meratifikasi konvensi tersebut ke dalam peraturan perundang-undangan.

Berdasarkan konvensi 2003 UNESCO, WBTb merupakan warisan yang diturunkan secara terus menerus dari generasi ke generasi berupa berbagai ekspresi, representasi, praktik, pengetahuan, artefak, objek, keterampilan, dan ruang-ruang budaya yang terkait, bagian dari warisan budaya tersebut adalah masyarakat, kelompok, ataupun perorangan. Kemudian UNESCO membagi WBTb menjadi lima kategori, yaitu Tradisi dan Ekspresi Lisan (TEL), Seni Pertunjukan (SP), Adat Istiadat Masyarakat, Ritual, dan Perayaan-perayaan (AIMRP), Pengetahuan dan Kebiasaan Perilaku Mengenai Alam dan Semesta (PKPMAS), dan Kemahiran dan Kerajinan Tradisional (KKT) yang selanjutnya dilakukan penetapan WBTb oleh Menteri yang membidangi kebudayaan dikarenakan berbagai macam ranah kebudayaan yang ada dan melestarikan WBTb yang terancam punah, penetapan WBTb diperlukan untuk pemberian status yang awalnya budaya takbenda menjadi WBTb.

WBTb mempunyai hubungan terhadap banyaknya wisatawan khususnya wisatawan domestik yang datang ke suatu provinsi, semakin banyak WBTb yang dimiliki suatu provinsi

maka diharapkan semakin banyak wisatawan domestiknya. Wisatawan domestik adalah penduduk yang melakukan perjalanan ke suatu tempat di negaranya, tetapi diluar lingkungan tempat tinggalnya sehari-hari dalam jangka waktu sekurang-kurangnya satu malam dan tidak lebih dari satu tahun, tujuan perjalanannya adalah bukan untuk mencari penghasilan dari tempat yang dikunjungi (WTO, 2008). Warisan budaya dan pariwisata di seluruh dunia memiliki hubungan yang sangat erat. Pariwisata pun dari aspek ekonomi membantu pelestarian benda-benda artefak dan kehidupan masyarakat di mata wisatawan (Siswanto, 2007)[11]. Pariwisata tiap provinsi memiliki daya tarik tersendiri bagi para wisatawan domestik, salah satu yang menjadi daya tarik adalah WBTb yang terdapat di provinsi tersebut.

Sektor pariwisata dapat menjadi “komoditas” melalui pertunjukan hingga produk kebudayaan untuk memenuhi kebutuhan komersil pariwisata (Siswanto, 2007)[11]. Pembangunan sosial budaya memiliki hubungan yang erat dengan pembangunan ekonomi, sehingga perlu dirancang kebijakan mengenai pelestarian dan pengembangan budaya salah satunya budaya takbenda agar menjadi pendorong laju pembangunan nasional. Perwujudan UU Nomor 5 tahun 2017 tentang pemajuan kebudayaan yang harus dilestarikan untuk membangun masa depan dan peradaban bangsa serta tujuan kesebelas pada *Sustainable Development Goals* (SDGs) 2030, yaitu kota dan pemukiman yang berkelanjutan dengan upaya melindungi dan mempromosikan WBTb dapat terpetakan dengan baik.

WBTb memengaruhi banyaknya wisatawan domestik sehingga dapat dilakukan klasifikasi. Salah satu analisis yang dapat mengklasifikasikan ialah analisis diskriminan. Analisis diskriminan merupakan analisis multivariat yang berkaitan dengan pemisahan sekelompok objek (observasi) yang berbeda dan melakukan penempatan objek (observasi) ke dalam kelompok yang telah ditentukan sebelumnya (Johnson dan Wichern, 2007)[8]. Analisis diskriminan memiliki tujuan untuk mengklasifikasikan suatu observasi ke dalam kelompok yang saling bebas dan menyeluruh berdasarkan variabel prediktor. Oleh karena itu, pada penelitian ini penulis bermaksud untuk menerapkan analisis diskriminan dalam pengklasifikasian WBTb dengan banyaknya wisatawan domestik. Hasil penelitian ini dapat menjadi rekomendasi bagi instansi terkait seperti Dinas Pariwisata, Kebudayaan, Pemuda, dan Olahraga (Disparbudpora) di tingkat provinsi maupun kabupaten kota.

2. METODE PENELITIAN

Penelitian ini menggunakan data penetapan WBTb dari 34 provinsi berupa data sekunder yang diperoleh melalui website Kemdikbud[9] dan data banyaknya wisatawan domestik berupa data sekunder yang diperoleh melalui Badan Pusat Statistik (BPS)[2] dengan nama “Jumlah Perjalanan Wisatawan Nusantara (Orang) 2021”. Data WBTb sebagai variabel independen, yaitu AIMRP (X_1), KKT (X_2), PKPMAS (X_3), SP (X_4), dan TEL (X_5). Data banyaknya wisatawan domestik menjadi variabel dependen (y).

Analisis diskriminan merupakan analisis multivariat yang dapat digunakan pada hubungan dependensi dalam kasus dimana variabel independen berupa data kuantitatif dan variabel dependen berupa data kualitatif. Analisis diskriminan memiliki tujuan untuk mengklasifikasikan suatu observasi ke dalam kelompok yang saling bebas dan menyeluruh berdasarkan variabel prediktor.

Analisis diskriminan memiliki asumsi dasar yang harus dipenuhi, yaitu (Backhaus *et al.*, 2021)[1]:

- (1) Normalitas multivariat dari variabel independen.
- (2) Tidak adanya multikolinearitas pada variabel independen.
- (3) Matriks varians kovarians dalam kelompok bersifat homogen.

Untuk dapat memenuhi asumsi dasar dari analisis diskriminan, diperlukan uji yang berhubungan, yaitu:

- (1) Uji normalitas multivariat.

- (2) Pemeriksaan multikolinearitas.
- (3) Uji homogenitas matriks varians kovarians.

Setelah asumsi terpenuhi, dalam analisis diskriminan diperlukan perhitungan korelasi kanonik, *Wilks' lambda*, *confusion matrix*, dan analisis diskriminan bertahap.

2.1. Uji Normalitas Multivariat. Asumsi pada analisis diskriminan adalah variabel independen merupakan variabel berskala metrik dan berdistribusi normal. Uji ini bertujuan untuk mengetahui sebuah data multivariat berdistribusi normal atau tidak. Data harus berdistribusi normal agar analisis diskriminan menjadi optimal. Pengujian normalitas multivariat dapat dilakukan dengan uji Henze-Zirkler.

Uji hipotesis:

H_0 : data berdistribusi normal multivariat.

H_1 : data tidak berdistribusi normal multivariat.

Taraf signifikansi $\alpha = 5$

Henze dan Zirkler (1990)[7] menyatakan bahwa menguji normal multivariat dapat digunakan statistik uji Henze-Zirkler sebagai berikut :

$$HZ = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n e^{-\frac{\beta^2}{2} D_{ij}} - 2(1 + \beta^2)^{-\frac{p}{2}} \frac{1}{n} \sum_{i=1}^n e^{\frac{-\beta^2}{2(1+\beta^2)} D_i} + (1 + 2\beta^2)^{\frac{p}{2}} \quad (1)$$

dengan,

$$\beta = \frac{1}{\sqrt{2}} \left(\frac{n(2p+1)}{4} \right)^{\frac{1}{p+4}}$$

$$D_{ij} = (x_i - x_j)' S^{-1} (x_i - x_j)$$

$$D_i = (x_i - \hat{x})' S^{-1} (x_i - \hat{x})$$

dengan,

p : banyak variabel

x_i : pengamatan ke- i

x_j : pengamatan ke- j

\hat{x} : rata-rata tiap variabel

S^{-1} : invers dari matriks varians kovarians S

D_{ij} : jarak mahalanobis antara x_i dan x_j

D_i : jarak mahalanobis pada x_i

H_0 tidak ditolak apabila nilai p -value $> \alpha$, maka dapat disimpulkan bahwa data berdistribusi normal (Henze dan Zirkler, 1990)[7]. Secara visual dapat dilihat dengan *Quantile-Quantile* (Q-Q) plot antara D_i dengan $\chi_p^2(\frac{n-i+\frac{1}{2}}{n})$, jika plot menyebar mendekati model (garis lurus), maka data multivariat berdistribusi normal (Johnson dan Wichern, 2007)[8]. Apabila asumsi normalitas tidak terpenuhi, maka dilakukan deteksi outlier (Hamidah *et al.*, 2022). Deteksi outlier dilakukan dengan menghapus data yang memiliki nilai $D_i > \chi_p^2(\frac{n-i+\frac{1}{2}}{n})$ atau berada diluar garis mahalanobis.

2.2. 2.2 Pemeriksaan Multikolinearitas. Multikolinearitas adalah hubungan linear antar variabel independen. Pemeriksaan multikolinearitas memiliki tujuan untuk melihat apakah model regresi terdapat korelasi yang tinggi antar variabel independen. Jika terdapat korelasi yang tinggi diantara variabel independen, maka hubungan antara variabel independen dan variabel dependen akan terganggu. Pemeriksaan multikolinearitas antar variabel independen yang paling sederhana dapat dilakukan menggunakan matriks korelasi. Koefisien korelasi yang digunakan adalah Pearson product-moment yang dapat dihitung menggunakan persamaan berikut:

$$r_{gh} = \frac{n\Sigma x_{ig}x_{ih} - (\Sigma x_{ig})(\Sigma x_{ih})}{\sqrt{\{n\Sigma x_{ig}^2 - (\Sigma x_{ig})^2\}\{n\Sigma x_{ih}^2 - (\Sigma x_{ih})^2\}}} \tag{2}$$

Apabila koefisien korelasi antar dua variabel $> 0,80$, maka terdapat multikolinearitas (Bertan, 2016)[4].

2.3. Uji Homogenitas Matriks Varians Kovarians. Pada analisis diskriminan, matriks varians kovarians diasumsikan homogen. Uji yang paling umum untuk menilai asumsi ini adalah dengan menggunakan uji Box M (Backhaus *et al.*, 2021)[1]. Uji Box M sensitif terhadap penyimpangan normalitas multivariat. Oleh karena itu, harus dilakukan uji normalitas multivariat terlebih dahulu.

Uji hipotesis:

$$H_0 : \Sigma_1 = \Sigma_2 = \dots = \Sigma_k.$$

H_1 : paling sedikit dua Σ_l tidak sama.

Taraf signifikansi $\alpha = 5$

Statistik uji Box M dengan pendekatan χ^2 (Johnson dan Wichern, 2007)[8]:

$$\chi_{hitung}^2 = -2(1 - c_1) M \tag{3}$$

dengan,

$$S_{pooled} = \frac{\sum_{l=1}^k v_l S_l}{\sum_{l=1}^k v_l}$$

$$M = \frac{1}{2} \sum_{l=1}^k v_l \ln |S_l| - \frac{1}{2} \left(\sum_{l=1}^k v_l \right) \ln |S_{pooled}|$$

$$c_1 = \left[\sum_{l=1}^k \frac{1}{v_l} - \frac{1}{\sum_{l=1}^k v_l} \right] \left[\frac{2p^2 + 3p - 1}{6(p+1)(k-1)} \right]$$

$$v_l = n_l - 1$$

dengan,

k : banyaknya kelompok

n_l : banyaknya pengamatan pada kelompok ke- l

S_l : matriks varians kovarians pada kelompok ke- l

S_{pooled} : matriks varians kovarians gabungan

H_0 tidak ditolak apabila $\chi_{hitung}^2 \leq \chi_{\alpha; \frac{1}{2}(k-1)p(p+1)}^2$ atau dapat disimpulkan bahwa matriks varians kovarians homogen.

2.4. Korelasi Kanonik. Korelasi kanonik mengukur sejauh mana hubungan antara nilai diskriminan dan kelompok (Backhaus *et al.*, 2021)[1]. Tiap sampel pada n observasi dibentuk matriks varians kovarians. Kombinasi linear dari dua kelompok dapat ditulis dengan persamaan berikut (Johnson dan Wichern, 2007)[8]:

$$U = a_1X_1 + a_2X_2 + \dots + a_pX_p = a'X$$

$$V = b_1Y_1 + b_2Y_2 + \dots + b_qY_q = b'Y$$

sehingga

$$Var(U) = a' Cov(X) a = a' S_{XX} a$$

$$Var(V) = b' Cov(Y) b = b' S_{YY} b$$

$$Cov(U, V) = a' Cov(X, Y) b = a' S_{XY} b$$

Vektor koefisien \mathbf{a} dan \mathbf{b} diperoleh dengan nilai eigen dari matriks $S_{XX}^{-1}S_{XY}S_{YY}^{-1}S_{YX}$ yang berpadanan dengan vektor eigen e_k dapat dihitung menggunakan persamaan berikut:

$$|S_{XX}^{-1}S_{XY}S_{YY}^{-1}S_{YX} - \lambda I| = 0$$

sehingga vektor koefisien \mathbf{a} dan \mathbf{b} dapat dihitung dengan rumus sebagai berikut:

$$a_k = S_{XX}^{-1}.e_k$$

$$b_k = S_{YY}^{-1}.S_{YX}.a_k$$

Menurut Johnson dan Wichern (2007)[8] korelasi kanonik dapat diperoleh melalui perhitungan berikut:

$$\rho = \text{Corr}(U, V) = \frac{\text{Cov}(U, V)}{\sqrt{\text{Var}(U)}\sqrt{\text{Var}(V)}} = \frac{a' S_{XY} b}{\sqrt{a' S_{XX} a} \sqrt{b' S_{YY} b}} \quad (4)$$

Nilai maksimum dan terbaik dari korelasi kanonik adalah 1 (Backhaus *et al.*, 2021)[1]. Tujuan utama dari korelasi kanonik adalah untuk mengukur asosiasi antara variabel dependen dengan variabel independen.

2.5. Fungsi Diskriminan. Fungsi diskriminan adalah kombinasi linear dari variabel-variabel yang dapat dicari dengan rumus berikut (Backhaus *et al.*, 2021)[1]:

$$Y = b_1 X_1 + b_2 X_2 + \dots + b_g X_g \quad (5)$$

dengan

Y : skor diskriminan

b_g : koefisien diskriminan; $g = 1, 2, \dots, p$

X_g : variabel independen; $g = 1, 2, \dots, p$

Koefisien diskriminan diperoleh dengan mencari nilai eigen dan dilanjutkan menggunakan persamaan berikut:

$$(W^{-1}B - \hat{\lambda}_i I) \hat{b}_i = 0$$

dengan,

$$W = \sum_{l=1}^k \sum_{i=1}^{n_l} (X_{ij} - \bar{X}_l) (X_{ij} - \bar{X}_l)'$$

$$B = \sum_{l=1}^k n_l (\bar{X}_l - \bar{X})(\bar{X}_l - \bar{X})'$$

2.6. Wilks' Lambda. Untuk menilai signifikansi fungsi diskriminan yang diperkirakan dapat menggunakan bantuan *Wilks' lambda* (Johnson dan Wichern, 2021)[8].

$$\Lambda = \frac{|W|}{|B + W|} \quad (6)$$

dengan,

Λ : *Wilks' lambda*

Wilks' lambda adalah ukuran dengan kualitas terbalik, yaitu jika semakin kecil nilainya, maka semakin baik daya pembeda dari fungsi diskriminan (Backhaus *et al.*, 2021)[1]. Nilai *Wilks' lambda* berkisar 0 sampai 1.

2.7. Confusion Matrix. *Confusion matrix* dapat memberikan informasi berdasarkan perbandingan hasil klasifikasi yang dilakukan oleh sistem dengan hasil klasifikasi sebenarnya. *Confusion matrix* berbentuk matriks dengan empat nilai prediksi dan nilai aktual yang berbeda (Harliana dan Widayani, 2019)[6] seperti pada Tabel 1.

		Actual Values	
		Positive	Negative
Prediction	Positive	True Positive (TP)	False Positive (FP)
	Negative	False Negative (FN)	True Negative (TN)

TABEL 1. *Confusion matrix*

Terdapat empat istilah pada klasifikasi dengan menggunakan *confusion matrix*, yaitu *True Positive* (TP), *False Positive* (FP), *False Negative* (FN), dan *True Negative* (TN). Penjelasan dari masing-masing istilah adalah sebagai berikut:

- TP merupakan data positif yang memiliki prediksi benar.
- FP merupakan data negatif yang diprediksi sebagai data positif (error tipe I).
- FN merupakan data positif yang diprediksi sebagai data negatif (error tipe II).
- TN merupakan data negatif yang memiliki prediksi benar.

Nilai tersebut digunakan untuk menghitung sensitivitas, spesifisitas, Nilai Prediksi Positif (NPP), dan Nilai Prediksi Negatif (NPN) yang dapat dihitung menggunakan rumus berikut (Harliana dan Widayani, 2019)[6]:

$$Sensitivitas = \frac{TP}{TP + FN} \times 100 \tag{7}$$

$$Spesifisitas = \frac{TN}{TN + FP} \times 100 \tag{8}$$

$$NPP = \frac{TP}{TP + FP} \times 100 \tag{9}$$

$$NPN = \frac{TN}{TN + FN} \times 100 \tag{10}$$

$$Akurasi = sensitivitas \frac{P}{(P + N)} + spesifisitas \frac{N}{(P + N)} \tag{11}$$

$$Nilai\ error = \frac{FP + FN}{(P + N)} \times 100 \tag{12}$$

2.8. Analisis Diskriminan Bertahap. Analisis diskriminan bertahap digunakan untuk mengetahui variabel yang paling dominan dalam mengelompokkan data. Menurut Backhaus *et al.* (2021)[1] prosedur analisis diskriminan bertahap mengikuti proses yang berurutan dengan menambah atau menghapus variabel independen. Pemilihan variabel independen yang memiliki pengaruh dominan terhadap diskriminasi kelompok dapat dilihat dari nilai *Akaike Information Criterion* (AIC). Menurut Harlyan *et al.* (2020) AIC dapat dihitung menggunakan persamaan sebagai berikut:

$$AIC = 2k - 2 \ln(L) \tag{13}$$

dengan

k : banyaknya variabel independen

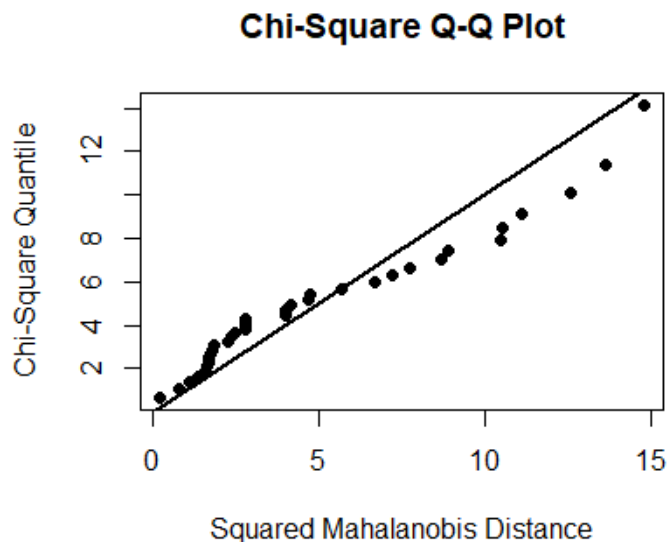
L : log-likelihood estimator yang mampu memproduksi variabel y

Model yang memiliki nilai AIC paling kecil adalah model terbaik (Harlyan *et al.*, 2020).

3. HASIL DAN PEMBAHASAN

3.1. **Uji Normalitas Multivariat untuk Data WBTb di Indonesia.** Langkah yang dilakukan untuk melakukan uji normalitas multivariat, antara lain :

- (1) Uji Normalitas Multivariat Uji normalitas multivariat menggunakan bantuan Q-Q plot dan uji Henze-Zikler dengan $\alpha = 0,05$. Data yang diuji hanya variabel independen (X_1, X_2, X_3, X_4, X_5), yaitu data penetapan WBTb dari 34 provinsi. Q-Q plot antara D_i dengan $\chi_p^2 \left(\frac{n-i+\frac{1}{2}}{n} \right)$ yang dihasilkan terdapat pada Gambar 1.

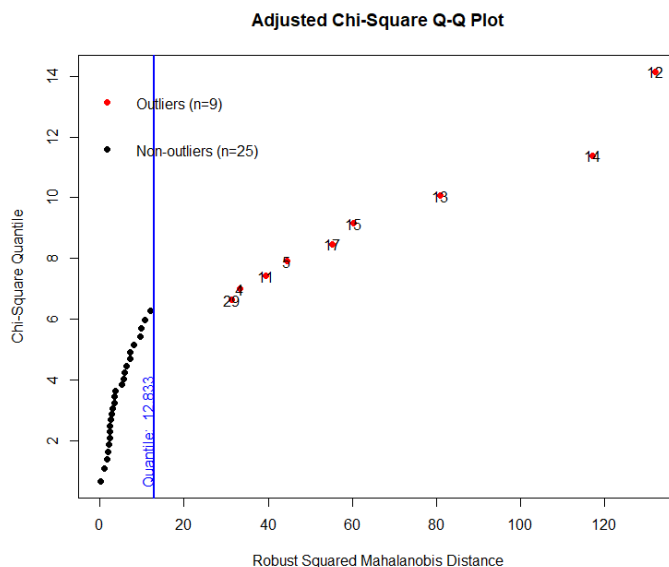


GAMBAR 1. Q-Q Plot

Gambar 1 menunjukkan bahwa plot yang dihasilkan dapat dilihat secara visual tidak mendekati garis horizontal atau data tidak berdistribusi normal multivariat. Q-Q plot yang diperoleh divalidasi menggunakan uji Henze-Zirkler dengan output $HZ = 1,539$ dan $p\text{-value} = 1,1021 \times 10^{-11}$ sehingga $p\text{-value} = 1,1021 \times 10^{-11} < 0,05$ yang berarti H_0 ditolak. Berdasarkan output Q-Q plot dan uji Henze-Zirkler yang diperoleh dapat disimpulkan bahwa data tidak berdistribusi normal multivariat. Menurut Hamidah *et al.* (2022) jika asumsi normalitas tidak terpenuhi, maka dilakukan deteksi outlier.

- (2) Data Outlier

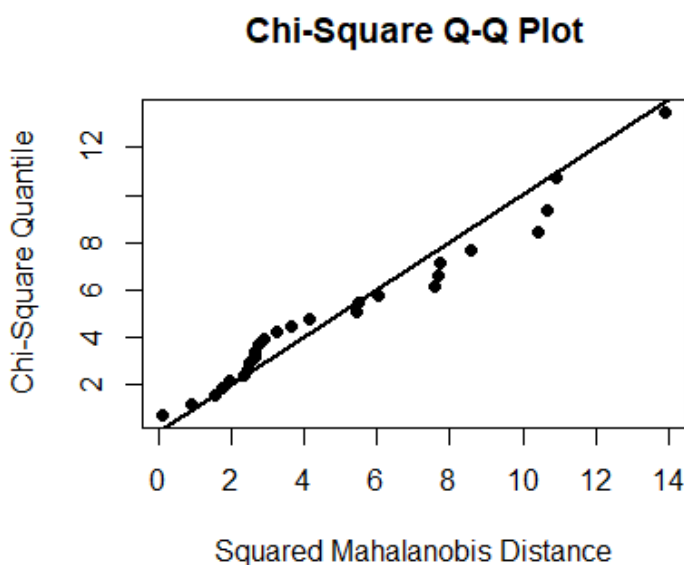
Data outlier merupakan data pengamatan yang berada jauh dibandingkan dengan pengamatan-pengamatan lainnya (Dewi *et al.*, 2016)[5]. Menurut Dewi *et al.* (2016)[5] keberadaan data outlier dapat mengganggu proses analisis data. Data outlier dapat diketahui berdasarkan plot berwarna merah yang melewati garis Mahalanobis. Deteksi outlier terdapat pada Gambar 2.



GAMBAR 2. Deteksi Outlier

Dapat dilihat pada Gambar 2, data outlier dari 34 provinsi terdapat 9 data. Rincian dari data outlier adalah Provinsi Gorontalo dengan nomor 29, Provinsi Riau dengan nomor 4, Provinsi DKI Jakarta dengan nomor 11, Provinsi Jambi dengan nomor 5, Provinsi Bali dengan nomor 17, Provinsi Jawa Timur dengan nomor 15, Provinsi Jawa Tengah dengan nomor 13, Provinsi DI Yogyakarta dengan nomor 14, dan Provinsi Jawa Barat dengan nomor 12. Oleh karena itu, 9 data tersebut dihapus dari data penelitian yang digunakan dan kembali dilakukan uji normalitas multivariat.

- (3) Uji Normalitas Multivariat Uji normalitas multivariat tanpa adanya kesembilan provinsi yang menjadi data outlier sehingga datanya menjadi 25 provinsi. Uji normalitas multivariat yang kembali dilakukan menghasilkan output yang terdapat pada Gambar 3 dan uji Henze-Zirkler yang diperoleh adalah nilai $HZ = 0,8599$ dan $p - value = 0,1824$.



GAMBAR 3. Q-Q Plot

Gambar 3 menunjukkan plot yang secara visual telah mendekati garis diagonal, maka data berdistribusi normal multivariat (Johnson dan Wichern, 2007)[8]. Nilai p -value = 0,1824 > 0,05 menunjukkan bahwa H_0 tidak ditolak (Henze dan Zirkler, 1990)[7]. Dapat disimpulkan bahwa data berdistribusi normal multivariat sehingga asumsi pertama dalam analisis diskriminan, yaitu normalitas multivariat dari variabel independen telah terpenuhi. Pengolahan data selanjutnya menggunakan 25 provinsi.

3.2. Pemeriksaan Multikolinearitas untuk Data WBTb di 25 Provinsi. Pemeriksaan multikolinearitas dilakukan menggunakan matriks korelasi dengan melihat korelasi antar variabel independen. Hasil yang diperoleh terdapat pada Tabel 2.

	AIMRP	KKT	PKPMAS	SP	TEL
AIMRP	1,0000000	0,4032016	0,1190638	0,5069861	0,6379187
KKT	0,4032016	1,0000000	0,4085814	0,5918868	0,6207175
PKPMAS	0,1190638	0,4085814	1,0000000	0,0362084	0,2773109
SP	0,5069861	0,5918868	0,0362084	1,0000000	0,6171994
TEL	0,6379187	0,6207175	0,2773109	0,6171994	1,0000000

TABEL 2. Matriks Korelasi

Berdasarkan Tabel 3 diketahui bahwa korelasi antar dua variabel tidak > 0,80, hal tersebut menunjukkan bahwa tidak terdapat multikolinearitas (Bertan, 2016)[4]. Oleh karena itu, asumsi tidak adanya multikolinearitas dapat terpenuhi.

3.3. Uji Homogenitas Matriks Varians Kovarians untuk Data WBTb di 25 Provinsi. Uji homogenitas matriks varians kovarians yang dilakukan menggunakan uji Box M adalah sebagai berikut:

Box M	DF	p -value
15,517	15	0,4148

TABEL 3. Uji Box M

Berdasarkan Tabel 3 diperoleh nilai p -value = 0,4148 > 0,05 menunjukkan bahwa matriks varians kovarians adalah homogen dan nilai uji Box M = 15,517 < $\chi_{0,05;15}^2 = 24,996$ sehingga H_0 tidak ditolak karena $\chi_{hitung}^2 \leq \chi_{\alpha; \frac{1}{2}(k-1)p(p+1)}^2$ atau dapat disimpulkan bahwa matriks varians kovarians homogen (Johnson dan Wichern, 2007)[8]. Disimpulkan bahwa matriks varians kovarians antar kelompok, yaitu kelompok kategori R (rendah) dan kelompok kategori T (tinggi) bersifat homogen. Hal tersebut menunjukkan bahwa asumsi ketiga pada analisis diskriminan, yaitu matriks varians kovarians dalam kelompok bersifat homogen sudah terpenuhi.

3.4. Korelasi Kanonik untuk Data WBTb di 25 Provinsi. Hubungan antara nilai diskriminan dan kelompok dapat diukur dengan korelasi kanonik (Backhaus *et al.*, 2021)[1]. Hasil yang diperoleh adalah nilai korelasi kanonik = 0,4259953 atau 42,6 yang berarti bahwa variabel diskriminan (y) atau banyaknya wisatawan domestik dapat dijelaskan oleh variabel independen (X_1, X_2, X_3, X_4, X_5), yaitu WBTb di 25 provinsi sebesar 42,6.

3.5. Fungsi Diskriminan untuk Data WBTb di 25 Provinsi. Fungsi diskriminan adalah kombinasi linear dari variabel-variabel (Backhaus *et al.*, 2021)[1]. Koefisien linear diskriminan yang diperoleh terdapat pada Tabel 4.

	LD1
AIMRP	0,087
KKT	-0,072
PKPMAS	-0,107
SP	-0,038
TEL	0,092

TABEL 4. Koefisien Linear Diskriminan

Berdasarkan Tabel 4 dapat diperoleh fungsi diskriminan sebagai berikut :

$$Y = 0,087X_1 - 0,072X_2 - 0,107X_3 - 0,038X_4 + 0,092X_5 \quad (14)$$

dengan

X_1 : AIMRP

X_2 : KKT

X_3 : PKPMAS

X_4 : SP

X_5 : TEL

Menurut Saputri *et al.* (2016)[10] jika diperoleh hasil negatif, maka dapat menurunkan skor fungsi diskriminan dan sebaliknya jika diperoleh hasil positif, maka dapat meningkatkan skor fungsi diskriminan. Fungsi diskriminan yang diperoleh pada persamaan (14) memiliki arti tersendiri pada tiap variabel independennya, yaitu :

- (1) Variabel X_1 , yaitu AIMRP memiliki koefisien sebesar 0,087 dan memiliki tanda positif. Hal ini berarti variabel X_1 mempunyai pengaruh positif, yaitu dapat menaikkan skor diskriminan sebesar 0,087 pada setiap kenaikan satu satuan nilai variabel X_1 . Artinya, variabel X_1 dapat meningkatkan wisatawan domestik sebesar 8,7%. Pada kenyataannya, saat ini AIMRP dapat menarik para wisatawan domestik untuk datang, melihat, dan merasakan langsung AIMRP yang berbeda pada setiap provinsinya.
- (2) Variabel X_2 , yaitu KKT memiliki koefisien sebesar 0,072 dan memiliki tanda negatif. Hal ini berarti variabel X_2 mempunyai pengaruh negatif, yaitu dapat menurunkan skor diskriminan sebesar 0,072 pada setiap kenaikan satu satuan variabel X_2 . Artinya, variabel X_2 dapat menurunkan wisatawan domestik sebesar 7,2%. Dalam hal ini, KKT sedikit kurang menarik banyaknya wisatawan domestik karena secara logika wisatawan umumnya tidak mengetahui KKT sebelum datang ke provinsi tersebut, KKT umumnya menjadi oleh-oleh yang dibeli wisatawan sebelum pulang ataupun hanya dapat dilihat di provinsi tersebut.
- (3) Variabel X_3 , yaitu PKPMAS memiliki koefisien sebesar 0,107 dan memiliki tanda negatif. Hal ini berarti variabel X_3 mempunyai pengaruh negatif, yaitu dapat menurunkan skor diskriminan sebesar 0,107 pada setiap kenaikan satu satuan variabel X_3 . Artinya, variabel X_3 dapat menurunkan wisatawan domestik sebesar 10,7%. Pada kenyataannya hanya sedikit wisatawan yang mengetahui PKPMAS pada suatu provinsi.
- (4) Variabel X_4 , yaitu SP memiliki koefisien sebesar 0,038 dan memiliki tanda negatif. Hal ini berarti variabel X_4 mempunyai pengaruh negatif, yaitu dapat menurunkan skor diskriminan sebesar 0,038 pada setiap kenaikan satu satuan variabel X_4 . Artinya, variabel X_4 dapat menurunkan wisatawan domestik sebesar 3,8%. Dalam hal ini, SP sedikit kurang menarik banyaknya wisatawan domestik yang dapat diartikan sedikit wisatawan tidak mengetahui SP yang ditampilkan sebelum mengunjungi provinsi tersebut atau bukan menjadi faktor utama untuk wisatawan domestik datang.
- (5) Variabel X_5 , yaitu TEL memiliki koefisien sebesar 0,092 dan memiliki tanda positif. Hal ini berarti variabel X_5 mempunyai pengaruh positif, yaitu dapat menaikkan skor diskriminan sebesar 0,092 pada setiap kenaikan satu satuan nilai variabel X_5 . Artinya, variabel X_5 dapat meningkatkan wisatawan domestik sebesar 9,2%. Pada kenyataannya, TEL sangat mudah menyebar dan dapat digunakan diantara masyarakat meskipun

mereka bukan berasal dari provinsi tersebut. Hal tersebut dapat menjadi daya tarik bagi wisatawan domestik untuk mengunjungi suatu provinsi yang kemungkinan besar mereka telah mengetahui dan menggunakan TEL provinsi tersebut.

Oleh karena itu, penerapan analisis diskriminan dari data WBTb terhadap banyaknya wisatawan domestik dapat dilakukan, tetapi dengan hasil koefisien diskriminan pada fungsi diskriminan tidak terlalu tinggi sehingga perlu dikaji ulang. Contohnya variabel yang digunakan bukan hanya data penetapan WBTb, tetapi juga layanan pariwisata, pendapatan wisatawan, hingga keterlibatan para wisatawan.

3.6. Wilks' Lambda untuk Data WBTb di 25 Provinsi. Menilai signifikansi fungsi diskriminan yang diperkirakan dapat menggunakan bantuan *Wilks' lambda* (Johnson dan Wichern, 2007)[8]. Nilai *Wilks' lambda* yang diperoleh adalah 0,53119. Menurut Backhaus *et al.* (2021)[1] jika nilai *Wilks' lambda* semakin kecil, maka semakin baik daya pembeda dari fungsi diskriminan. Itu artinya, daya pembeda dari fungsi diskriminan sudah cukup baik. Walaupun hanya dua koefisien diskriminan yang positif, tetapi apabila dihitung menggunakan *Wilks' Lambda*, maka masih dapat digunakan.

3.7. Confusion Matrix untuk Data WBTb di 25 Provinsi. *Confusion matrix* memberikan informasi berdasarkan perbandingan hasil klasifikasi yang dilakukan oleh sistem dengan hasil klasifikasi sebenarnya. Hasil yang diperoleh terdapat pada Tabel 5.

	R	T
R	18	1
T	0	6

TABEL 5. Hasil *Confusion Matrix*

Berdasarkan hasil yang diperoleh pada Tabel 5 terdapat 18 provinsi yang rendah masuk sesuai kategorinya, 6 provinsi yang tinggi masuk sesuai kategorinya, dan 1 provinsi tidak sesuai dengan kategori yang dimiliki. Menghitung sensitifitas menggunakan persamaan (7) diperoleh hasil 100%, spesifisitas diperoleh menggunakan persamaan (8) diperoleh hasil 85,71%, NPP dihitung menggunakan persamaan (9) diperoleh hasil 94,74%, NPN dihitung dengan persamaan (10) dan diperoleh hasilnya adalah 100%, diperoleh nilai akurasi dengan persamaan (11) adalah 96,57%, dan menghitung nilai error dengan persamaan (12) diperoleh hasil 4%. Walaupun kontribusi dari koefisien diskriminan tidak terlalu besar, setelah dilakukan pemeriksaan penerapan analisis diskriminan pada data WBTb terhadap banyaknya wisatawan domestik diperoleh nilai error yang sangat kecil.

3.8. Analisis Diskriminan Bertahap untuk Data WBTb di 25 Provinsi. Analisis diskriminan bertahap dilakukan untuk mengetahui variabel independen yang memiliki pengaruh paling besar dalam mengkategorikan data. Hasil yang diperoleh adalah nilai AIC terkecil sebesar 791,97 dengan variabel independen yang dapat menghasilkan nilai AIC minimum dan memiliki penurunan AIC yang signifikan adalah variabel AIMRP dengan nilai AIC 794,79, variabel PKPMAS dengan nilai AIC 794,45, dan variabel TEL dengan nilai AIC 796,66. Demikian sehingga, dari kelima variabel independen yang dimasukkan hanya terdapat tiga variabel yang paling berpengaruh dalam diskriminasi kategori banyaknya wisatawan domestik tinggi atau rendah, yaitu AIMRP, PKPMAS, dan TEL. Hasil yang diperoleh disesuaikan dengan fungsi diskriminan yang diperoleh, yaitu AIMRP dan TEL memiliki tanda positif terbesar, sedangkan PKPMAS memiliki tanda negatif terbesar. Hal tersebut menunjukkan AIMRP dan TEL berpengaruh dominan terhadap tingginya wisatawan domestik, sedangkan PKPMAS berpengaruh dominan terhadap rendahnya wisatawan domestik.

4. SIMPULAN

Analisis diskriminan dapat mengklasifikasikan data WBTb terhadap kategori wisatawan tinggi dan rendah. Variabel yang dominan untuk mengklasifikasikan tinggi rendahnya wisatawan domestik adalah AIMRP, PKPMAS, dan TEL. Selain itu, penerapan analisis diskriminan pada data WBTb di Indonesia memberikan hasil bahwa kategori yang berpengaruh positif pada banyaknya wisatawan domestik adalah AIMRP dan TEL, sedangkan kategori yang berpengaruh negatif adalah KKT, PKPMAS, dan SP.

Ucapan Terimakasih.

Penulis mengucapkan terima kasih kepada Rektor Universitas Padjadjaran yang telah memberikan dana untuk diseminasi hasil penelitian dosen dan mahasiswa melalui *Academic Leadership Grant* tahun 2023 dengan nomor kontrak: 1549/UN6.3.1/PT.00/2023 serta kepada reviewer yang telah membantu memberikan saran untuk melengkapi penulisan paper ini.

DAFTAR PUSTAKA

- [1] Backhaus, K., Erichson, B., Gensler, S., Weiber, R., dan Weiber, T., 2021, *Multivariate Analysis an Application-Oriented introduction*, Springer Gabler, Wiesbaden.
- [2] Badan Pusat Statistik, 2022, Jumlah Perjalanan Wisatawan Nusantara (Orang) 2021: An Overview, *Online*, <https://www.bps.go.id/indicator/16/1189/1/jumlah-perjalanan-wisatawan-nusantara.html> on September 9, 2022.
- [3] Badan Pusat Statistik, 2021, *Statistik Sosial Budaya 2021*, BPS-Statistics Indonesia.
- [4] Bertan, C. V., 2016, Pengaruh Pendayagunaan Sumber Daya Manusia (Tenaga Kerja) terhadap Hasil Pekerjaan (Studi Kasus Perumahan Taman Mapanget Raya(TAMARA)), *Jurnal Sipil Statistik*, 4(1), pp. 15-17 (2016).
- [5] Dewi, E. T. K., Agoestanto, A., dan Sunarmi, 2016, Metode Least Trimmed Square (LTS) dan MM-Estimation untuk Mengestimasi Parameter Regresi Ketika Terdapat Outlier, *UNNES Journal of Mathematics*, 5(1), pp. 48-54 (2016).
- [6] Harliana dan Widayani, W., 2009, Analisis Dempster Shafer pada Sistem Pakar Pendeteksi Penyakit Ispa, *FAHMA-Jurnal Teknologi dan Ilmu Komputer*, 17(2), pp. 63-68 (2009).
- [7] Henze, N. dan Zirkler, B., 1990, A Class of Invariant Consistent Tests for Multivariate Normality, *Communications in Statistics - Theory and Methods*, 19(10), pp. 3595-3617 (1990).
- [8] Johnson, R. A. dan Wichern, D. W., 2007, *Applied Multivariate Statistical Analysis*, Pearson Prentice Hall, New Jersey.
- [9] Kemdikbud, 2022, Warisan Budaya Takbenda Indonesia: An Overview, *Online*, <https://warisanbudaya.kemdikbud.go.id/?penetapan&&list&limitto=5&vieww=per10&tab=1> on October 14, 2022.
- [10] Saputri, A. R., Virgantari, F., dan Wijayanti, H., 2016, Penerapan Analisis Diskriminan untuk Mengkaji Perbedaan Konsumsi Pangan Hewani di Wilayah Perkotaan dan Perdesaan di Indonesia, *Ekologia*, 16(2), pp. 18-23 (2016).
- [11] Siswanto, N. F. N., 2007, Pariwisata dan Pelestarian Warisan Budaya, *Berkala Arkeologi*, 27(1), pp. 118-130 (2007).

