



Analisis Diskriminan Dalam Pengklasifikasian Kemiskinan Pada Kota/Kabupaten di Provinsi Sumatera Utara

Puteri Fajar Addini^{1*}, Windy², Adi Purwanto³

^{1,2,3}Teknik Informatika, STMIK Pelita Nusantara, Medan, Indonesia
Email Korespondensi: pf.addini@gmail.com

Abstrak—Pembangunan manusia didefinisikan sebagai suatu proses untuk perluasan pilihan yang lebih banyak kepada penduduk melalui upaya-upaya pemberdayaan yang mengutamakan peningkatan kemampuan dasar manusia. Nilai kebutuhan dasar yang harus dapat diperoleh setiap orang agar tidak terkategori miskin yang kemungkinan disebut garis kemiskinan. Penelitian ini bertujuan mendapatkan model untuk mengklasifikasikan kemiskinan pada kota/kabupaten di Provinsi Sumatera Utara dengan faktor indeks pembangunan manusia, tingkat kemiskinan, tingkat pengangguran terbuka, dan upah minimum pada kota/kabupaten di Provinsi Sumatera Utara. Multivariate Analisis diskriminan merupakan salah satu metode analisis multivariat yang digunakan untuk mengklasifikasikan beberapa kelompok data yang sudah terkelompokkan dengan cara membentuk kombinasi linier fungsi diskriminan. Dalam penelitian ini, dapat disimpulkan bahwa analisis diskriminan dapat digunakan untuk menggolongkan kemiskinan di kota/kabupaten di Sumatera Utara berdasarkan variabel kemiskinan.

Kata Kunci: Pengklasifikasian, Diskriminan, Kemiskinan

Abstract—Human development is defined as a process for expanding more choices to the population through empowerment efforts that prioritize the improvement of basic human capabilities. The value of basic needs that must be obtained by everyone so that they are not categorized as poor, which might be called the poverty line. This study aims to obtain a model for classifying poverty in cities/districts in North Sumatra Province with the factors of human development index, poverty rate, open unemployment rate, and minimum wage in cities/districts in North Sumatra Province. Multivariate Discriminant analysis is a multivariate analysis method used to classify several groups of data that have been grouped by forming a linear combination of discriminant functions. In this study, it can be concluded that discriminant analysis can be used to classify poverty in cities/districts in North Sumatra based on the poverty variable.

Keywords: Classification, Discrimination, Poverty

I. PENDAHULUAN

Keberhasilan pembangunan nasional tidak hanya dilihat dari laju pertumbuhan ekonomi yang tinggi, tetapi yang paling penting adalah keberhasilan pembangunan manusia. Pembangunan manusia didefinisikan sebagai suatu proses untuk perluasan pilihan yang lebih banyak kepada penduduk melalui upaya-upaya pemberdayaan yang mengutamakan peningkatan kemampuan dasar manusia agar dapat sepenuhnya berpartisipasi disegala bidang pembangunan[1].

Nilai kebutuhan dasar yang harus dapat diperoleh setiap orang agar tidak terkategori miskin yang kemungkinan disebut garis kemiskinan (GK). Kemiskinan merupakan kondisi perekonomian seseorang atau sekelompok orang yang tidak terpenuhinya hak-hak dasarnya untuk mempertahankan dan mengembangkan kehidupan yang bermartabat. Kebutuhan dasar tersebut meliputi kebutuhan pangan, kesehatan, pendidikan, perumahan, pekerjaan, air bersih, pertanian, sumber daya alam, lingkungan hidup.

Analisis diskriminan adalah analisis yang tepat digunakan untuk memodelkan hubungan antara satu peubah respon berskala non metrik (nominal atau ordinal) dengan lebih dari satu peubah prediktor yang berskala metrik (interval atau rasio). Analisis ini bertujuan untuk mengklasifikasikan obyek ke dalam kelompok yang telah diketahui dari informasi awal dan menghitung risiko dari kemungkinan kesalahan pengklasifikasian.

Penelitian ini bertujuan mendapatkan model untuk mengklasifikasikan kemiskinan pada kota/kabupaten di Provinsi Sumatera Utara dengan faktor indeks pembangunan manusia, tingkat kemiskinan, tingkat pengangguran terbuka, dan upah minimum pada kota/kabupaten di Provinsi Sumatera Utara.

II. METODE PENELITIAN

2.1 Kemiskinan

Kemiskinan adalah ketidakmampuan untuk membeli barang-barang kebutuhan dasar seperti makanan, pakaian, papan, dan obat-obatan Maipita [2]Sedangkan Badan Pusat Statistik[3] mendefinisikan kemiskinan sebagai ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur dari sisi pengeluaran. Kemudian menurut Kuncoro [4]dalam Tyas [5]kemiskinan adalah ketidakmampuan untuk memenuhi standar hidup minimum. Dengan demikian, dapat disimpulkan bahwa kemiskinan adalah suatu keadaan dimana seseorang atau daerah tidak dapat meningkatkan kehidupan yang lebih layak atau dapat dikatakan tidak dapat meningkatkan standar hidup yang lebih baik.

2.2 Analisis Diskriminan

Multivariate Analisis diskriminan merupakan salah satu metode analisis multivariat yang digunakan untuk mengetahui variabel-variabel ciri yang membedakan tiap-



tiap kelompok yang terbentuk dan bertujuan untuk mengklasifikasikan beberapa kelompok data yang sudah terkelompokkan dengan cara membentuk kombinasi linier fungsi diskriminan[6],[7].

2.2.2 Model analisis diskriminan

Model yang digunakan untuk Analisis Diskriminan, sangat mirip dengan Analisis faktor dan regresi, yang merupakan suatu kombinasi linear dari berbagai variabel independen yaitu[8],[6]

$$D = b_1X_1 + b_2X_2 + b_3X_3 + \dots + b_kX_k \quad (1)$$

Dimana,

- a. D = Skor diskriminan
- b. b = Koefesien diskriminan
- c. X = Variabel independent

2.2.2 Menghitung *cutting score*

cutting score dapat dihitung dengan rumus sebagai berikut[6],

$$Z_{cu} = \frac{N_A Z_B + N_B Z_A}{N_A N_B} \quad (2)$$

Dimana,

- a. Z_{cu} = Cutting score
- b. N_A = jumlah anggota grup a
- c. N_B = jumlah anggota grup B
- d. Z_A = centroid grup A
- e. Z_B = centroid grup B

2.2.3 Menilai Keakuratan Prediksi Klasifikasi

Untuk menilai model secara keseluruhan adalah dengan menentukan tingkat keakuratan prediksi dari fungsi diskriminan. Penentuan ini dilakukan dengan menggunakan uji statistik yang dinamakan Press's Q Statistik. Adapun Press's Q Statistik dapat ditulis dengan rumus,

$$Press'Q = \frac{[N-(qk)]^2}{N(k-1)} \quad (3)$$

Dimana,

- a. N = Banyaknya sampel = $n_{11} + n_{12} + n_{21} + n_{22}$
- b. q = Banyaknya kasus yang diklasifikasi secara tepat = $n_{11} + n_{12}$
- c. k = Banyaknya Kelompok

III. HASIL DAN PEMBAHASAN

3.1 Standarisasi Data

Standarisasi digunakan ketika terdapat perbedaan satuan yang berbeda. Pada penelitian ini, data yang digunakan memiliki satuan yang berbeda maka dilakukan standarisasi data. Hasil standarisasi data Indikator dan IPM. Berdasarkan tabel dibawah, telah dilakukan standarisasi sehingga satuan dari setiap peubah telah sama. Selanjutnya, dapat dilanjutkan dengan melakukan pengelompokan peubah menggunakan cluster hirarki. Pada metode cluster, hal awal yang perlu dilakukan yaitu penentuan ukuran kemiripan antar objek.

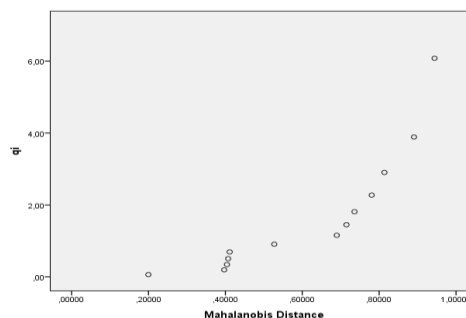
3.2 Proses Analisis Diskriminan

3.2.1 Uji Normal Multivariat

Uji normalitas multivariate dengan teknik scatter-plot dan korelasi antara jarak mahalnobis dengan Chi Square. teknik ini dapat digunakan untuk uji asumsi normalitas multivariat, dengan hipotesisnya[9],[10]:

H_0 : Data berasal dari populasi berdistribusi normal

H_a : Data berasal dari populasi yang tidak berdistribusi normal



Gambar 1. Data Populasi

Apabila secara visual menampilkan setidaknya 50% plot nya membentuk sebuah garis linear, maka dapat dikatakan telah memenuhi asumsi Data berasal dari populasi berdistribusi normal. Dari table mahalnobis distance dapat dilihat plotnya membentuk sebuah garis linier, sehingga data-data memenuhi asumsi dan berdistribusi normal.

Tabel 1. Data Memenuhi Asumsi dan berdistribusi Normal

Correlations			
		Mahalanobis Distance	Qi
Mahalanobis Distance	Pearson Correlation	1	,865**
	Sig. (2-tailed)		,000
	N	33	13
qi	Pearson Correlation	,865**	1
	Sig. (2-tailed)	,000	
	N	13	13

** . Correlation is significant at the 0.01 level (2-tailed).

Nilai koefisien korelasinya 0,865 bandingkan dengan nilai r table. Untuk praktisnya lihat saja nilai sig 0.000 < 0,05. Dengan demikian korelasinya signifikan dan dapat diartikan bahwa data berasal dari populasi berdistribusi normal.



3.2.2 Uji Multivariate

Tabel 2. Uji Multivariate

Descriptive Statistics				
	K	Mean	Std. Deviation	N
Upah Minimum (Rupiah)	0	2464181,39	500691,248	18
	1	2791570,20	328008,530	15
	Total	2612994,48	455692,092	33
Penggangan Terbuka	0	3754,56	3307,950	18
	1	6040,53	10590,101	15
	Total	4793,64	7497,671	33
Jumlah Penduduk Miskin 2022 (Jiwa)	0	39,6500	23,37470	18
	1	36,9467	45,42048	15
	Total	38,4212	34,56447	33

Tabel 3. Box's Test of Equality

Box's Test of Equality of Covariance Matrices^a

Box's M	29,588
F	4,403
df1	6
df2	6318,516
Sig.	,000

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

a. Design: Intercept + K

Tabel 4. Design Intercept + K

Multivariate Tests ^a							
Effect	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	
Intercept	Pillai's Trace	,979	444,246 ^b	3,000	29,000	,000	,979
	Wilks' Lambda	,021	444,246 ^b	3,000	29,000	,000	,979
	Hotelling's Trace	45,956	444,246 ^b	3,000	29,000	,000	,979
	Roy's Largest Root	45,956	444,246 ^b	3,000	29,000	,000	,979
K	Pillai's Trace	,256	3,320 ^b	3,000	29,000	,033	,256
	Wilks' Lambda	,744	3,320 ^b	3,000	29,000	,033	,256
	Hotelling's Trace	,343	3,320 ^b	3,000	29,000	,033	,256
	Roy's Largest Root	,343	3,320 ^b	3,000	29,000	,033	,256

a. Design: Intercept + K
b. Exact statistic

Tabel 5. Levene's Test

Levene's Test of Equality of Error Variances^a

	F	df1	df2	Sig.
Upah Minimum (Rupiah)	3,535	1	31	,070
Penggangan Terbuka	4,728	1	31	,037
Jumlah Penduduk Miskin 2022 (Jiwa)	,818	1	31	,373

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + K

Tabel 6. Test Of Betwevs

Tests of Between-Subjects Effects							
Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	Upah Minimum (Rupiah)	876955366151,565 ^a	1	876955366151,565	4,713	,038	,132
	Penggangan Terbuka	42755681,459 ^b	1	42755681,459	,755	,392	,024
	Jumlah Penduduk Miskin 2022 (Jiwa)	59,793 ^c	1	59,793	,049	,827	,002
Intercept	Upah Minimum (Rupiah)	226005748069975,160	1	226005748069975,160	1214,660	,000	,975
	Penggangan Terbuka	784994451,883	1	784994451,883	13,857	,001	,309
	Jumlah Penduduk Miskin 2022 (Jiwa)	48003,131	1	48003,131	,38985	,000	,557
K	Upah Minimum (Rupiah)	876955366151,564	1	876955366151,564	4,713	,038	,132
	Penggangan Terbuka	42755681,459	1	42755681,459	,755	,392	,024
	Jumlah Penduduk Miskin 2022 (Jiwa)	59,793	1	59,793	,049	,827	,002
Upah Minimum (Rupiah)	Upah Minimum (Rupiah)	5768013689762,677	31	186064957734,280			
	Penggangan Terbuka	1756126382,178	31	56649238,135			
	Jumlah Penduduk Miskin 2022 (Jiwa)	38170,682	31	1231,312			
Total	Upah Minimum (Rupiah)	231960394924918,000	33				
	Penggangan Terbuka	2557187400,000	33				
	Jumlah Penduduk Miskin 2022 (Jiwa)	86944,730	33				
Corrected Total	Upah Minimum (Rupiah)	6644969055914,242	32				
	Penggangan Terbuka	1798882063,636	32				
	Jumlah Penduduk Miskin 2022 (Jiwa)	38230,475	32				

a. R Squared = ,132 (Adjusted R Squared = ,104)
b. R Squared = ,024 (Adjusted R Squared = -,008)

Tabel 7. Uji Varian Kovarian

Covariance Matrices^a

K	Upah Minimum (Rupiah)	Penggangan Terbuka	Jumlah Penduduk Miskin 2022 (Jiwa)
0	Upah Minimum (Rupiah)	250691726179,899	904818117,536
	Penggangan Terbuka	904818117,536	10942533,203
	Jumlah Penduduk Miskin 2022 (Jiwa)	4210051,032	61731,876
1	Upah Minimum (Rupiah)	107589596050,315	2196230901,957
	Penggangan Terbuka	2196230901,957	112150236,981
	Jumlah Penduduk Miskin 2022 (Jiwa)	9487830,397	451867,359
Total	Upah Minimum (Rupiah)	207655282997,320	163288824,151
	Penggangan Terbuka	163288824,151	56215064,489
	Jumlah Penduduk Miskin 2022 (Jiwa)	6161226,496	228906,977

a. The total covariance matrix has 32 degrees of freedom.

Nilai di sepanjang diagonal matriks kovarian hanyalah varian dari setiap subjek. Misalnya:

Varian dari upah minimum adalah 250691726179,899

Varian skor dari penggangan terbuka adalah 56215064,49

Varian skor jumlah penduduk miskin adalah 6161226,49 .

Nilai-nilai lain dalam matriks mewakili kovarian antara berbagai subjek. Misalnya:

Kovariansi antara skor upah minimum dan penggangan terbuka adalah 904818117,536 .



Kovariansi antara skor jumlah penduduk miskin dan upah minimum adalah 6161226,496 .

Angka positif untuk kovarians menunjukkan bahwa dua variabel cenderung naik atau turun secara bersamaan. Sebagai contoh, upah minimum dan pengangguran terbuka memiliki kovarians positif, yang menunjukkan bahwa penduduk yang mendapat nilai tinggi pada upah minimum juga cenderung mendapat nilai tinggi pada pengangguran terbuka. Demikian pula, penduduk yang mendapat nilai rendah pada upah minimum juga cenderung memiliki nilai rendah pada pengangguran terbuka.

Uji diskriminan

Terdapat empat peubah yang terseleksi masuk dalam analisis, yaitu angka harapan hidup, harapan lama sekolah, rata-rata lama sekolah, dan paritas daya beli. Sebelum melanjutkan ke analisis diskriminan, terlebih dahulu akan dimunculkan rata-rata nilai peubah dari masing-masing kelompok yang terbentuk.

Unweighted Cases		N	Percent
Valid		33	94,3
Excluded	Missing or out-of-range group codes	0	.0
	At least one missing discriminating variable	0	.0
	Both missing or out-of-range group codes and at least one missing discriminating variable	2	5,7
	Total	2	5,7
Total		35	100,0

K		Mean	Std. Deviation	Valid N (listwise)	
				Unweighted	Weighted
0	Upah Minimum (Rupiah)	2464181,3889	500691,24836	18	18,000
	Pengangguran Terbuka	3754,5556	3307,95000	18	18,000
	Jumlah Penduduk Miskin 2022 (Jiwa)	39,6500	23,37470	18	18,000
1	Upah Minimum (Rupiah)	2791570,2000	328008,53045	15	15,000
	Pengangguran Terbuka	6040,5333	10590,10090	15	15,000
	Jumlah Penduduk Miskin 2022 (Jiwa)	36,9467	45,42048	15	15,000
Total	Upah Minimum (Rupiah)	2612994,4848	455692,09231	33	33,000
	Pengangguran Terbuka	4793,6364	7497,67060	33	33,000
	Jumlah Penduduk Miskin 2022 (Jiwa)	38,4212	34,56447	33	33,000

Uji Kesamaan Vektor rata-rata

	Wilks' Lambda	F	df1	df2	Sig.
Upah Minimum (Rupiah)	.868	4,713	1	31	.038
Pengangguran Terbuka	.976	.755	1	31	.392
Jumlah Penduduk Miskin 2022 (Jiwa)	.998	.049	1	31	.827

Dilihat dari nilai p-value nya, masing-masing variabel mempunyai rata-rata yang berbeda untuk kedua kelompok. Ingat, yang diuji adalah kesamaan rata-rata pada tiap kelompok (kelompok kode nol dan kode satu), bukan rata-rata antar variabel.

Pengecekan multikolinieritas,

	Upah Minimum (Rupiah)	Pengangguran Terbuka	Jumlah Penduduk Miskin 2022 (Jiwa)	
Covariance	Upah Minimum (Rupiah)	186064957734,280	1488036794,371	6593564,294
	Pengangguran Terbuka	1488036794,371	56649238,135	237922,094
	Jumlah Penduduk Miskin 2022 (Jiwa)	6593564,294	237922,094	1231,312
Correlation	Upah Minimum (Rupiah)	1,000	.458	.436
	Pengangguran Terbuka	.458	1,000	.901
	Jumlah Penduduk Miskin 2022 (Jiwa)	.436	.901	1,000

a. The covariance matrix has 31 degrees of freedom.

Dari matriks korelasi di atas, tidak ada angka yang mencapai 0,5 atau di atasnya sehingga kita mengidentifikasi tidak ada multikolinieritas pada data.

Tabel 8. Uji Kesamaan matriks varians-covarians

Box's M	2,601
Approx.	2,520
df1	1
F	2811,671
Sig.	.113

Tests null hypothesis of equal population covariance matrices.

Tabel di atas memperlihatkan bahwa menolak hipotesis nol karena nilai p-valuenya kurang dari 0,05 (dalam hal ini penelitian menggunakan tingkat kepercayaan 95%). Dari hasil pengujian ini, kita dapat mengatakan bahwa data kita berasal dari populasi yang mempunyai matriks varians-covarians yang sama.

Tabel 9. Penghitungan cutting score,

K	Function
	1
0	-.345
1	.414

Unstandardized canonical discriminant functions evaluated at group means

Dari tabel di atas, dapat dihitung cutting score nya (-0,345)(0,414) = - + Untuk observasi pertama, karena discriminant score nya kurang dari cutting score, maka dimasukkan ke dalam kelompok kode 0 (pengklasifikasian tepat karena sebenarnya observasi pertama sebelumnya memang termasuk ke dalam anggota kelompok nol atau bunga A).



Tabel 10. Hit Ratio,

		K	Predicted Group Membership		Total
			0	1	
Original	Count	0	10	8	18
		1	4	11	15
	%	0	55,6	44,4	100,0
Cross-validated ^b	Count	0	10	8	18
		1	5	10	15
	%	0	55,6	44,4	100,0
			33,3	66,7	100,0

a. 63,6% of original grouped cases correctly classified.

b. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

c. 60,6% of cross-validated grouped cases correctly classified.

Angka hit ratio di atas sudah mencapai 100% (pada kenyataannya sulit mencapai angka sebesar ini, ingat ini hanya data fiktif yang dibangkitkan dengan bantuan komputer). Pengklasifikasian observasi baru dapat diprediksi akan termasuk dalam kelompok mana berdasarkan karakteristik yang dimilikinya dengan fungsi linier yang sudah terbentuk. Inilah yang menjadi tujuan pembentukan fungsi diskriminan.

IV. KESIMPULAN

Dalam penelitian ini, dapat disimpulkan bahwa analisis diskriminan dapat digunakan untuk menggolongkan kemiskinan di kota/kabupaten di Sumatera utara berdasarkan variabel kemiskinan. Dari 33 kota/kabupaten di Sumatera Utara, 15 kota/kabupaten dikategorikan sebagai kota/kabupaten maju dan 18 kota/kabupaten dikategorikan sebagai kota/kabupaten berkembang. Variabel yang memberikan kontribusi paling besar adalah pengangguran terbuka dan variabel yang memberi kontribusi paling kecil adalah upah minimum.

UCAPAN TERIMA KASIH

Terima kasih disampaikan kepada pihak-pihak yang telah mendukung terlaksananya penelitian ini.

V. REFERENSI

- [1] M. Sianturi and N. Andika, "Peningkatan Efisiensi Penelusuran Aset melalui Sistem Manajemen Aset dan Analytical Hierarchy Process," *J. Sist. Inf.*, vol. 2, 2022.
- [2] M. R. Syahwana and R. M. Simanjorang, "Analisa Sistem Pakar Metode Bayes Dalam Mendiagnosa Penyakit Tuberculosis," *J. Sist. Inf.*, vol. 1, 2022.
- [3] R. Sitepu, "Implementasi Algoritma K-Nearest Neighbor Untuk Klasifikasi Pengajuan Kredit," *J. Sist. Inf.*, vol. 1, 2022.
- [4] A. R. Faqih and A. A. Widya, "Implementasi Aplikasi E-Ticket pada Bumdes Desa Sumbermulyo Kec. Jogoroto Kab. Jombang sebagai Solusi Digitalisasi Pengelolaan Tiket.," vol. 2, 2023.
- [5] F. F. Nugraha and E. A. Firdaus, "Implementasi Permainan Instruksional sebagai Media Pembelajaran untuk Meningkatkan Kualitas Pendidikan di SMA," *J. Sist. Inf.*, vol. 2, 2022.
- [6] A. Simangunsong, R. M. Simanjorang, and H. Fahmi, "Penerapan Metode Composite Performance

Index Dalam Seleksi Penerimaan Calon Laboran," vol. 1, 2022.

- [7] A. Suaib and I. I. Tritosmoro, "Perbandingan Performa Metode Local Binary Pattern dan Random Forest dalam Identifikasi COVID-19 pada Citra X-ray Paru-paru.," vol. 2, 2023.
- [8] I. M. Sianturi and D. Harinto, "Perbandingan Kinerja Algoritma Random Forest pada Prediksi Penetapan Tarif Penerbangan dengan Menggunakan Auto-ML," *J. Sist. Inf.*, vol. 2, 2022.
- [9] E. Murniyasih and A. Jamlean, "Perancangan Prototype Sistem Kartu Pelajar Cerdas Berbasis RFID di MA Insan Kamil Kota Sorong," *J. Sist. Inf.*, vol. 1, 2022.
- [10] D. E. Frans, "Peningkatan Produksi Budidaya Perikanan dengan Penerapan Algoritma Apriori dan Association Rule," vol. 2, 2023.