

Perbandingan Analisis Diskriminan dan Analisis Regresi Logistik Ordinal dalam Prediksi Klasifikasi Kondisi Kesehatan Bank

Fajri Zufa

Alumni Program Studi Statistika, FMIPA, Universitas Bengkulu
e-mail : fajrizufa@gmail.com

Sigit Nugroho

Staf Pengajar Program Studi Statistika, FMIPA, Universitas Bengkulu
e-mail : snugroho@unib.ac.id

Mudin Simanihuruk

Staf Pengajar Program Studi Matematika, FMIPA, Universitas Bengkulu
e-mail : mudinsimanihuruk@yahoo.com.au

Abstract: *The purpose of this research is to compare the accuracy of bank classification prediction based on Capital Adequacy Ratio (CAR), Earning Asset Quality (EAQ), Non Performing Loan (NPL), Return on Assets (ROA), Net Interest Margin (NIM), Short Term Mismatch (STM) and Loan to Deposit Ratio (LDR). Discriminant analysis and ordinal logistic regression analysis are compared in classifying the prediction. The data used are secondary data, namely data classification of bank conditions in Indonesia in 2014 obtained from research institute PT Infovesta Utama. Based on Apparent Error Rate (APER) score obtained, it can be said that discriminant analysis is better in predicting the classification of bank conditions in Indonesia than that of ordinal logistic regression analysis. Discriminant analysis has the average prediction accuracy of 80%, while ordinal logistic regression analysis has the average prediction accuracy of 74,38%.*

Keywords: *Discriminant Analysis, Ordinal Logistic Regression Analysis, Prediction.*

1. Pendahuluan

Substansi permasalahan prediksi klasifikasi kondisi bank ditentukan oleh hubungan antara variabel prediktor dengan variabel respon. Beberapa variabel prediktor dapat dihubungkan dengan 1 variabel respon atau lebih. Untuk menganalisis hubungan beberapa variabel prediktor dengan 1 variabel respon yang terdiri dari 2 kategori atau lebih, dapat digunakan analisis multivariat, yaitu analisis diskriminan dan analisis regresi logistik ordinal.

Tujuan dari analisis diskriminan adalah untuk menggambarkan ciri-ciri suatu pengamatan dari bermacam-macam populasi yang diketahui, baik secara grafis maupun aljabar dengan membentuk fungsi diskriminan[1]. Dengan kata lain, analisis diskriminan digunakan untuk mengklasifikasikan individu ke dalam salah satu dari dua kelompok atau lebih dan dapat digunakan untuk mengetahui seberapa besar terjadi misklasifikasi. Kegunaan dari analisis diskriminan, yaitu: (1) penjelasan (*explanation*) terhadap fenomena yang dipelajari atau permasalahan yang diteliti, (2) prediksi alternatif kategori dari variabel respon berdasarkan nilai variabel prediktor, (3) penentuan variabel prediktor mana yang merupakan pembeda terkuat, (4) pengelompokan obyek.

Sedangkan analisis regresi logistik ordinal merupakan salah satu metode statistik yang menggambarkan hubungan antara suatu variabel respon (Y) dengan lebih dari satu variabel prediktor (X) di mana variabel respon lebih dari dua kategori dan skala pengukuran bersifat tingkatan [2]. Regresi logistik ordinal merupakan perluasan dari regresi logistik biner di mana variabel respon berupa skala ordinal yang terdiri dari tiga kategori atau lebih dan variabel prediktor berupa *covariate* (jika menggunakan skala interval atau rasio) atau *factor* (jika menggunakan skala nominal atau ordinal).

Pada penelitian ini, akan dibandingkan perhitungan prediksi klasifikasi kondisi bank berdasarkan rasio CAR, EAQ, NPL, ROA, NIM, STM dan LDR menggunakan analisis diskriminan dan analisis regresi logistik ordinal. Kemudian, melalui perbandingan nilai ukuran ketepatan prediksi dengan menggunakan *Apparent Error Rate* (APER) akan diketahui analisis manakah yang lebih baik dalam memprediksi klasifikasi kondisi bank di Indonesia.

2. Metode Penelitian

2.1 Sumber Data

Data yang digunakan pada penelitian ini adalah data sekunder, yaitu data klasifikasi kondisi bank di Indonesia menurut metode CAMEL pada tahun 2014 yang diambil dari situs resmi Lembaga Riset Data PT Infovesta Utama.

Berikut uraian variabel yang digunakan pada penelitian ini:

Variabel respon:

- Kondisi Bank (Y), terdiri dari tiga kategori (kelompok), yaitu Sangat Sehat, Sehat, dan Cukup Sehat (berdasarkan data klasifikasi kondisi Bank di Indonesia).

Variabel prediktor:

- *Capital Adequacy Ratio* (X_1),
- *Earning Asset Quality* (X_2),
- *Non Performing Loan* (X_3),

- *Return on Assets* (X_4),
- *Net Interest Margin* (X_5),
- *Short Term Mismatch* (X_6),
- *Loan to Deposit Ratio* (X_7).

Data dibagi menjadi dua bagian, yaitu data *training* dan data *testing*. Data *training* digunakan untuk membentuk model dalam analisis, baik model analisis diskriminan ataupun model analisis regresi logistik ordinal. Sedangkan data *testing* digunakan untuk menguji ketepatan prediksi dari model yang terbentuk. Dengan pertimbangan banyaknya anggota kategori yang tak sama, jumlah data *training* dan *testing* masing-masing dipilih 80% dan 20% dari keseluruhan data. Dalam hal ini kategori ketiga (Cukup Sehat) hanya berjumlah 3 sampel. Sehingga, maksimal kombinasi untuk kategori ketiga (Cukup Sehat) adalah 2 sampel untuk data *training* dan 1 sampel untuk data *testing* digunakan perbandingan 80% untuk data *training* dan 20% untuk data *testing*. Pembagian data *training* dan *testing* pada data klasifikasi kondisi bank di Indonesia, yaitu:

$$N_{\text{training}} = \frac{80}{100} \times 100 = 80$$

$$N_{\text{testing}} = \frac{20}{100} \times 100 = 20$$

Data klasifikasi kondisi bank di Indonesia pada tahun 2014 terdiri dari 100 bank, maka untuk pembagian data *training* dan *testing* pada data klasifikasi kondisi bank di Indonesia masing-masing sebanyak 80 bank data *training* dan 20 bank data *testing*.

Teknik pengambilan sampel yang digunakan adalah *Proportionate Stratified Random Sampling*. Teknik ini hampir sama dengan *Simple Random Sampling* namun penentuan sampelnya memperhatikan strata (tingkatan) yang ada dalam populasi. Jumlah bank yang diteliti ada 100 bank dengan 3 tingkatan strata atau kategori kondisi bank (Y) berdasarkan kesehatan bank, yaitu bank dengan kondisi Sangat Sehat ada 60 bank, bank dengan kondisi Sehat ada 37 bank, dan bank dengan kondisi Cukup Sehat ada 3 bank. Sampling dilakukan untuk memilih data *training* dan data *testing*.

Jumlah sampel yang diambil untuk data *training* adalah sebanyak 80 bank, dengan rincian sebagai berikut:

$$\text{Bank Kondisi Sangat Sehat} = \frac{60}{100} \times 80 = 48 \text{ bank}$$

$$\text{Bank Kondisi Sehat} = \frac{37}{100} \times 80 = 29,6 \approx 30 \text{ bank}$$

$$\text{Bank Kondisi Cukup Sehat} = \frac{3}{100} \times 80 = 2,4 \approx 2 \text{ bank}$$

Jumlah sampel yang diambil untuk data *testing* adalah sebanyak 20 bank, dengan rincian sebagai berikut:

$$\text{Bank Kondisi Sangat Sehat} = \frac{60}{100} \times 20 = 12 \text{ bank}$$

$$\text{Bank Kondisi Sehat} = \frac{37}{100} \times 20 = 7,4 \approx 7 \text{ bank}$$

$$\text{Bank Kondisi Cukup Sehat} = \frac{3}{100} \times 20 = 0,6 \approx 1 \text{ bank}$$

2.2 Metode Analisis

A. Analisis Diskriminan

1. Uji asumsi normal multivariat dan uji kesamaan matriks ragam-peragam antar kelompok pada data.
2. Penghitungan besarnya nilai eigen dan vektor eigen, dengan menggunakan matriks ragam-peragam dari vektor peubah acak X (variabel prediktor).
3. Pembentukan fungsi diskriminan dengan menggunakan metode *simultaneous estimation*.
4. Penghitungan peranan relatif fungsi diskriminan.
5. Uji signifikansi fungsi diskriminan dengan menggunakan uji *Wilk's Lambda*.
6. Penghitungan besarnya korelasi kanonik.
7. Penghitungan skor diskriminan dan rata-rata skor tiap kelompok (*centroid*) pada variabel respon.
8. Prediksi klasifikasi dengan cara Fisher pada data *training* dan data *testing*.
9. Penghitungan nilai ketepatan klasifikasi antara nilai sebenarnya dari sumber data dengan nilai prediksi yang diperoleh dari model diskriminan yang dibentuk dengan menggunakan APER (*Apparent Error Rate*).

B. Analisis Regresi Logistik Ordinal

1. Uji asumsi regresi logistik ordinal pada data, yaitu uji kebebasan antar variabel prediktor (Uji Non-Multikolinearitas).
2. Pembentukan model regresi logistik ordinal $P(Y \leq k|x_i)$

Model yang dapat dipakai untuk regresi logistik ordinal adalah model logit kumulatif. Pada model ini terdapat sifat ordinal dari respon Y yang dituangkan dalam peluang kumulatif, sehingga model logit kumulatif merupakan model yang didapatkan dengan cara membandingkan peluang kumulatif yaitu peluang kurang dari atau sama dengan ketegori respon ke- k pada p variabel prediktor yang dinyatakan dalam vektor x_i , $P(Y \leq k |x_i)$ dengan peluang lebih besar daripada kategori respon ke- k , $P(Y > k|x_i)$. Peluang kumulatif, $P(Y \leq k|x_i)$ didefinisikan sebagai berikut [2]:

$$P(Y \leq k|x_i) = \frac{\exp(\beta_{0k} + \sum_{j=1}^p \beta_j x_{ij})}{1 + \exp(\beta_{0k} + \sum_{j=1}^p \beta_j x_{ij})},$$

dengan $i = 1, 2, 3, \dots, n$; $j = 1, 2, \dots, p$ dan $k = 1, 2, \dots, m$,

di mana:

β_{0k} : konstanta kelompok (kategori) variabel respon ke- k ,

β_j : koefisien variabel prediktor ke- j ,

x_{ij} : data pengamatan ke- i variabel prediktor ke- j .

3. Uji kesesuaian model dengan menggunakan uji *Deviance*, dengan hipotesis sebagai berikut:
 H_0 : model logit layak untuk digunakan,
 H_1 : model logit tidak layak untuk digunakan.
4. Uji simultan dengan menggunakan uji G, dengan hipotesis sebagai berikut:
 H_0 : $\beta_1 = \beta_2 = \dots = \beta_p = 0$, artinya variabel prediktor secara bersama-sama tidak mempengaruhi variabel respon,
 H_1 : minimal ada satu $\beta_j \neq 0$, $j = 1, 2, \dots, p$, artinya variabel prediktor secara bersama-sama mempengaruhi variabel respon.
5. Uji parsial dengan menggunakan uji Wald, dengan hipotesis sebagai berikut:
 H_0 : $\beta_j = 0$, $j = 1, 2, \dots, p$, artinya variabel prediktor secara individu tidak mempengaruhi variabel respon,
 H_1 : $\beta_j \neq 0$, artinya variabel prediktor secara individu mempengaruhi variabel respon
6. Penghitungan koefisien determinasi McFadden, Cox dan Snell, dan Nagelkerke.
7. Prediksi klasifikasi dengan menggunakan model regresi logistik ordinal yang terbentuk pada data *training* dan data *testing*.
10. Penghitungan nilai ketepatan klasifikasi antara nilai sebenarnya dari sumber data dengan nilai prediksi yang diperoleh dari model regresi logistik ordinal yang dibentuk (Ketepatan klasifikasi model) dengan menggunakan APER (*Apparent Error Rate*).

Setelah dilakukan *fitting* model masing-masing analisis, dapat diketahui analisis manakah yang lebih baik dalam melakukan prediksi terhadap klasifikasi kondisi bank, yaitu dengan melihat nilai APER yang terkecil dari masing-masing analisis. Analisis yang memiliki nilai APER terkecil merupakan analisis yang lebih baik dalam melakukan prediksi.

3. Hasil dan Pembahasan

3.1 Statistik Deskriptif Data

Analisis deskriptif dimaksudkan untuk melihat karakteristik data, yaitu nilai minimum, nilai maksimum, nilai rata-rata (*mean*), dan standar deviasi dari masing-masing variabel. Statistik deskriptif data disajikan pada tabel berikut.

Tabel 1. Statistik Deskriptif Data

Variabel	Minimum (%)	Maksimum (%)	Mean (%)	Standar Deviasi (%)
CAR	12,89	60,54	20,32	8,07
EAQ	77,79	99,94	97,84	2,78
NPL	0,03	5,45	1,5	1,4
ROA	-4,96	7,66	1,83	1,71
NIM	0,24	11,41	5,48	2,29
STM	90,53	171,81	115,83	13,89
LDR	51,97	169,76	91,25	18,8

3.2 Analisis Diskriminan

Banyaknya fungsi diskriminan yang terbentuk secara umum tergantung dari $\min(p, m - 1)$, dengan p adalah banyaknya variabel prediktor dan m adalah banyaknya kelompok yang telah ditetapkan. Dalam penelitian ini terdapat 7 variabel prediktor dan 3 kategori (kelompok) variabel respon (Sangat Sehat, Sehat, dan Cukup Sehat), sehingga akan terbentuk 2 fungsi diskriminan.

Dari *software* SPSS 18 juga dihasilkan fungsi diskriminan tidak baku (bentuk umum fungsi diskriminan), sebagai berikut:

$$y_1 = -98,040 - 0,032 x_1 + 0,979 x_2 + 1,141 x_3 + 0,022 x_4 + 0,191 x_5 - 0,015 x_6 + 0,020 x_7,$$

$$y_2 = 23,174 + 0,009 x_1 - 0,289 x_2 - 0,136 x_3 + 0,586 x_4 + 0,234 x_5 + 0,035 x_6 - 0,014 x_7.$$

Kemudian dilakukan uji signifikansi:

Tabel 2. Uji Signifikansi Fungsi Diskriminan

Test of Function(s)	Wilk's Lambda	Chi square	df	Sig.
1 through 2	0,117	158,615	14	0,000
2	0,563	42,491	6	0,000

Berdasarkan tabel tersebut, diperoleh uji signifikansi masing-masing fungsi diskriminan, yaitu:

Fungsi diskriminan 1 s/d 2:

$$\chi^2_{hitung} = - \left[(n - 1) - \frac{(p + m)}{2} \right] \ln \Delta = - \left[(80 - 1) - \frac{(7 + 3)}{2} \right] \ln(0,117)$$

$$= 158,77$$

$$\chi^2_{tabel(0,05;14)} = 23,68$$

Berdasarkan perhitungan statistik uji dan hasil *software* SPSS 18, nilai statistik uji $\chi^2_{hitung} > \chi^2_{tabel(0,05;14)}$ dan nilai *p-value* (*Sig.*) $\leq \alpha$ ($= 0,05$), maka H_0 ditolak. Dengan resiko berbuat kesalahan 5%, dapat dikatakan bahwa fungsi diskriminan 1 s/d 2 signifikan dalam menentukan kelompok kondisi bank.

Fungsi diskriminan 2 saja:

$$\begin{aligned} \chi^2_{hitung} &= - \left[(n - 1) - \frac{(p + m)}{2} \right] \ln \Delta = - \left[(80 - 1) - \frac{(7 + 3)}{2} \right] \ln(0,563) \\ &= 42,51 \\ \chi^2_{tabel(0,05;6)} &= 12,59 \end{aligned}$$

Berdasarkan perhitungan statistik uji dan hasil *software* SPSS 18, nilai statistik uji $\chi^2_{hitung} > \chi^2_{tabel(0,05;6)}$ dan nilai *p-value* (*Sig.*) $\leq \alpha$ ($= 0,05$), maka H_0 ditolak. Dengan resiko berbuat kesalahan 5%, dapat dikatakan bahwa fungsi diskriminan 2 saja juga signifikan. Dengan demikian kedua fungsi diskriminan tersebut signifikan dan akan digunakan dalam analisis diskriminan.

Setelah diketahui fungsi diskriminan yang signifikan, akan dilakukan uji signifikansi untuk mengetahui variabel mana yang signifikan dalam fungsi diskriminan tersebut. Dari output *software* SPSS 18, diperoleh hasil uji signifikansi untuk masing-masing variabel sebagai berikut.

Tabel 3. Uji Signifikansi Fungsi Diskriminan

	<i>Wilk's Lambda</i>	<i>F</i>	<i>df1</i>	<i>df2</i>	<i>Sig.</i>
CAR	0,973	1,050	2	77	0,355
EAQ	0,474	42,725	2	77	0,000
NPL	0,934	2,706	2	77	0,073
ROA	0,602	25,403	2	77	0,000
NIM	0,692	17,132	2	77	0,000
STM	0,883	5,085	2	77	0,008
LDR	0,918	3,451	2	77	0,037

Berdasarkan tabel tersebut, dapat diketahui bahwa variabel CAR dan NPL tidak signifikan terhadap pembentukan kelompok pada variabel respon.

Korelasi Kanonik yang diperoleh adalah sebagai berikut:

Tabel 4. Nilai Koefisien Kanonik

<i>Function</i>	<i>Eigen value</i>	<i>% of Variance</i>	<i>Cumulative %</i>	<i>Canonical Correlation</i>
1	3,803	83,1	83,1	0,890
2	0,776	16,9	100	0,661

Berdasarkan tabel tersebut, dapat dilihat pada fungsi diskriminan 1, kontribusi pengaruh CAR, EAQ, NPL, ROA, NIM, STM, dan LDR terhadap kondisi kesehatan bank adalah sebesar $(0,890)^2 = 0,7921$ atau 79,21%. Sedangkan, pada fungsi diskriminan 2, kontribusi pengaruh CAR, EAQ, NPL, ROA, NIM, STM, dan LDR terhadap kondisi kesehatan bank adalah sebesar $(0,661)^2 = 0,4369$ atau 43,69%.

Pada penelitian ini akan digunakan 2 fungsi diskriminan, karena fungsi diskriminan 1 dan fungsi diskriminan 2 terbukti signifikan. Selain itu, berdasarkan kontribusi variabel bebas, fungsi diskriminan 1 memiliki nilai korelasi kanonik sebesar 89% dan fungsi diskriminan 2 sebesar 66,1%, kedua fungsi tersebut memiliki nilai korelasi kanonik lebih dari 50%.

Pengelompokan dilakukan terhadap data pengamatan (80 bank) dengan menggunakan cara pengelompokan Fisher, yaitu:

Data pengamatan ke-1 (Bank AGRIS)

$$x_1 = 17,58; x_2 = 99,56; x_3 = 0,66; x_4 = 0,29; x_5 = 2,78; x_6 = 104,39; x_7 = 70,02,$$

diperoleh skor diskriminan fungsi diskriminan 1:

$$\hat{y}_{11} = -98,040 - 0,032 (17,58) + 0,979 (99,56) + 1,141 (0,66) + 0,022 (0,29) + 0,191 (2,78) + 0,015 (104,39) + 0,020 (70,02)$$

$$\hat{y}_{11} = -0,008$$

dan diperoleh juga skor diskriminan fungsi diskriminan 2:

$$\hat{y}_{21} = 23,174 + 0,009 (17,58) - 0,289(99,56) - 0,136(0,66) + 0,586(0,29) + 0,234 (2,78) + 0,035 (104,39) - 0,014 (70,02)$$

$$\hat{y}_{21} = -2,037$$

Setelah diperoleh semua skor diskriminan fungsi diskriminan 1 dan fungsi diskriminan 2 dari pengamatan ke-1 hingga ke-80, kemudian akan dihitung rata-rata skor tiap kelompok (*centroid*) dari fungsi diskriminan 1 dan fungsi diskriminan 2, yaitu:

$$\bar{y}_{11} = \frac{\text{Jumlah skor diskriminan fungsi diskriminan 1 kelompok 1}}{\text{Banyak pengamatan kelompok 1}}$$

$$\begin{aligned}
 &= \frac{14,8159}{48} = 0,309 \\
 \bar{Y}_{12} &= \frac{\text{Jumlah skor diskriminan fungsi diskriminan 1 kelompok 2}}{\text{Banyak pengamatan kelompok 2}} \\
 &= \frac{7,4354}{30} = 0,248 \\
 \bar{Y}_{13} &= \frac{\text{Jumlah skor diskriminan fungsi diskriminan 1 kelompok 3}}{\text{Banyak pengamatan kelompok 3}} \\
 &= \frac{-23,9072}{2} = -11,954 \\
 \bar{Y}_{21} &= \frac{\text{Jumlah skor diskriminan fungsi diskriminan 2 kelompok 1}}{\text{Banyak pengamatan kelompok 1}} \\
 &= \frac{33,5756}{48} = 0,699 \\
 \bar{Y}_{22} &= \frac{\text{Jumlah skor diskriminan fungsi diskriminan 2 kelompok 2}}{\text{Banyak pengamatan kelompok 2}} \\
 &= \frac{-32,9866}{30} = -1,099 \\
 \bar{Y}_{23} &= \frac{\text{Jumlah skor diskriminan fungsi diskriminan 2 kelompok 3}}{\text{Banyak pengamatan kelompok 3}} \\
 &= \frac{0,2052}{2} = 0,103
 \end{aligned}$$

rata-rata skor tiap kelompok (*centroid*) fungsi diskriminan 1 dan fungsi diskriminan 2:

$$\begin{aligned}
 \bar{y}_{11} &= 0,309; & \bar{y}_{21} &= 0,699, \\
 \bar{y}_{12} &= 0,248; & \bar{y}_{22} &= -1,099, \\
 \bar{y}_{13} &= -11,954; & \bar{y}_{23} &= 0,103.
 \end{aligned}$$

klasifikasi Fisher:

$$\begin{aligned}
 \text{Kel. 1} &\rightarrow \sum_{g=1}^q (\hat{y}_{g1} - \bar{y}_{g1})^2 = (\hat{y}_{11} - \bar{y}_{11})^2 + (\hat{y}_{21} - \bar{y}_{21})^2 \\
 &= (-0,008 - 0,309)^2 + (-2,037 - 0,699)^2 \\
 &= 7,586 \\
 \text{Kel. 2} &\rightarrow \sum_{g=1}^q (\hat{y}_{g1} - \bar{y}_{g2})^2 = (\hat{y}_{11} - \bar{y}_{12})^2 + (\hat{y}_{21} - \bar{y}_{22})^2 \\
 &= (-0,008 - 0,248)^2 + (-2,037 - (-1,099))^2 \\
 &= 0,945 \\
 \text{Kel 3.} &\rightarrow \sum_{g=1}^q (\hat{y}_{g1} - \bar{y}_{g3})^2 = (\hat{y}_{11} - \bar{y}_{13})^2 + (\hat{y}_{21} - \bar{y}_{23})^2 \\
 &= (-0,008 - (-11,954))^2 + (-2,037 - 0,103)^2 \\
 &= 147,286
 \end{aligned}$$

Dari nilai tersebut yang terkecil adalah kelompok 2. Sehingga Bank AGRIS dengan CAR = 17,58; EAQ = 99,56; NPL = 0,66; ROA = 0,29; NIM = 2,78; STM = 104,39; LDR = 70,02, dikelompokkan ke kategori kondisi bank Sehat (sesuai dengan kondisi sebenarnya). Kemudian dilakukan prediksi dengan model analisis diskriminan tersebut pada seluruh data *training* sejumlah 80 Bank dan pada data *testing* sejumlah 20 Bank.. Hasil prediksi klasifikasi kondisi bank dengan menggunakan model analisis diskriminan pada data *testing* dapat dilihat pada Tabel 5 berikut.

Tabel 5. Hasil Prediksi Klasifikasi Kondisi Bank dengan Model Analisis Diskriminan

Nama Bank	Kondisi Bank (Y) Aktual	Kondisi Bank (Y) Prediksi (An Diskriminan)	Ketepatan Klasifikasi
Bank Andara	2	2	Benar
Bank Anglomas Internasional	1	1	Benar
Bank Commonwealth	1	2	Salah
Bank Dinar Indonesia	2	2	Benar
Bank GANESHA	3	2	Salah
Bank Mayapada International Tbk	2	2	Benar
Bank Mizuho Indonesia	2	1	Salah
Bank Nusantara Parahyangan Tbk	1	2	Salah
Bank Pundi Indonesia Tbk	2	2	Benar
Bank Resona Perdania	1	1	Benar
Bank Sumitomo Mitsui Indonesia	2	1	Salah
BPD Riau	1	1	Benar
BPD Sulawesi Selatan	1	1	Benar
BPD Sulawesi Tengah	1	1	Benar
BPD Sulawesi Tenggara	1	1	Benar
BPD Sulawesi Utara	1	1	Benar
BPD Sumatera Barat	1	1	Benar
BPD Sumatera Selatan	1	1	Benar
BPD Sumatera Utara	1	1	Benar
The Bank Of Tokyo Mitsubishi	2	2	Benar

3.3 Analisis Regresi Logistik Ordinal

Model yang dapat dipakai untuk regresi logistik ordinal adalah model logit kumulatif. Pada model ini terdapat sifat ordinal dari respon Y yang dituangkan dalam peluang kumulatif, sehingga model logit kumulatif merupakan model yang didapatkan dengan cara membandingkan peluang kumulatif yaitu peluang kurang dari atau sama dengan ketegori respon ke- k pada \mathbf{p} variabel prediktor yang dinyatakan dalam vektor

x_i , $P(Y \leq k | x_i)$ dengan peluang lebih besar daripada kategori respon ke- k , $P(Y > k | x_i)$ [2]. Dari output *software* SPSS 18, diperoleh konstanta dan koefisien variabel prediktor sebagai berikut:

Tabel 6. Nilai Konstanta Dan Koefisien Variabel Prediktor

	<i>Parameter</i>	<i>Estimate</i>
<i>Threshold</i>	[Y = 1]	-97,302
	[Y = 2]	-89,278
<i>Location</i>	X1	-0,017
	X2	-0,803
	X3	-1,58
	X4	-1,883
	X5	-0,942
	X6	-0,126
	X7	0,061

Berdasarkan hasil output *software* SPSS 18 tersebut, maka diperoleh 2 model analisis regresi logistik ordinal, yaitu:

$$\pi_1(x) = \frac{e^{(-97,302 - 0,017x_1 - 0,803x_2 - 1,58x_3 - 1,883x_4 - 0,942x_5 - 0,126x_6 + 0,061x_7)}}{1 + e^{(-97,302 - 0,017x_1 - 0,803x_2 - 1,58x_3 - 1,883x_4 - 0,942x_5 - 0,126x_6 + 0,061x_7)}}$$

$$\pi_2(x) = \frac{e^{(-89,278 - 0,017x_1 - 0,803x_2 - 1,58x_3 - 1,883x_4 - 0,942x_5 - 0,126x_6 + 0,061x_7)}}{1 + e^{(-89,278 - 0,017x_1 - 0,803x_2 - 1,58x_3 - 1,883x_4 - 0,942x_5 - 0,126x_6 + 0,061x_7)}}$$

Jika model analisis regresi logistik ordinal tersebut ditransformasi menjadi bentuk logit, maka model berikut akan diperoleh, yaitu:

$$\begin{aligned} \text{logit}(\pi_1(x)) &= \ln\left(\frac{\pi_1(x)}{1 - \pi_1(x)}\right) \\ &= -97,302 - 0,017x_1 - 0,803x_2 - 1,58x_3 - 1,883x_4 - 0,942x_5 - 0,126x_6 + 0,061x_7. \end{aligned}$$

$$\begin{aligned} \text{logit}(\pi_2(x)) &= \ln\left(\frac{\pi_2(x)}{1 - \pi_2(x)}\right) \\ &= -89,278 - 0,017x_1 - 0,803x_2 - 1,58x_3 - 1,883x_4 - 0,942x_5 - 0,126x_6 + 0,061x_7. \end{aligned}$$

Kemudian dilakukan uji kesesuaian model:

Tabel 7. Nilai Statistik Uji *Deviance*

	<i>Chi-Square</i>	<i>df</i>	<i>Sig.</i>
<i>Pearson</i>	64,010	151	1,000
<i>Deviance</i>	53,807	151	1,000

Berdasarkan hasil *software* SPSS 18, nilai statistik uji $\chi^2_{hitung} (= 53,807) \leq \chi^2_{tabel(0,05;151)} (= 180,68)$ dan nilai *p-value* ($=$ mendekati 1) $> \alpha (= 0,05)$, maka H_0 : model logit layak digunakan, diterima. Dengan resiko berbuat kesalahan 5%, dapat dikatakan bahwa model logit layak untuk digunakan. Setelah itu, dilihat variabel mana yang signifikan terhadap pembentukan kelompok:

Tabel 8. Nilai Statistik Uji Wald

	<i>Parameter</i>	<i>Wald</i>	<i>df</i>	<i>Sig.</i>
<i>Threshold</i>	[Y = 1]	6,336	1	0,012
	[Y = 2]	5,796	1	0,016
<i>Location</i>	X1	0,145	1	0,704
	X2	4,766	1	0,029
	X3	7,726	1	0,005
	X4	9,896	1	0,002
	X5	8,766	1	0,003
	X6	5,607	1	0,018
	X7	3,707	1	0,054

Berdasarkan tabel tersebut, dapat dilihat bahwa variabel X1 (CAR) dan X7 (LDR) tidak signifikan terhadap penentuan kelompok pada variabel respon

Berdasarkan nilai koefisien determinasi Nagelkerke diperoleh nilai sebesar 73,6%. Hal ini menandakan bahwa kontribusi pengaruh CAR, EAQ, NPL, ROA, NIM, STM, dan LDR terhadap kondisi kesehatan bank adalah sebesar 0,736 atau 73,6%, sedangkan 26,4% keragaman sisanya dijelaskan oleh variabel lain yang tidak diikuti dalam model.

Kemudian akan dihitung nilai peluang kumulatif logit yang didapat dari model regresi logistik ordinal yang terbentuk dan dikelompokkan ke kategori variabel respon tertentu, sebagai berikut:

Data pengamatan ke-1 (Bank AGRIS)

Untuk mengetahui pengelompokan variabel respon dari model analisis regresi logistik ordinal ini, dihitung dulu nilai y^*_i dengan cara sebagai berikut:

$$y^*_i = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p.$$

Untuk data pengamatan ke-1: $x_1 = 17,58$; $x_2 = 99,56$; $x_3 = 0,66$; $x_4 = 0,29$; $x_5 = 2,78$; $x_6 = 104,39$; $x_7 = 70,02$,

diperoleh nilai y^*_i :

$$y^*_1 = -0,017(17,58) - 0,803(99,56) - 1,58(0,66) - 1,883(0,29) - 0,942(2,78) - 0,126(104,39) + 0,061(70,02)$$

$$y^*_1 = -93,335$$

konstanta dari model logit:

$$\alpha_1 = -97,302$$

$$\alpha_2 = -89,278$$

klasifikasi kelompok

$$\begin{aligned} \text{Kel. 1} \rightarrow \rho(y_k = 1) &= \frac{1}{1+e^{(y^*_1-\alpha_1)}} = \frac{1}{1+e^{(-93,335-(-97,302))}} \\ &= 0,019 \end{aligned}$$

$$\begin{aligned} \text{Kel. 2} \rightarrow \rho(y_k = 2) &= \frac{1}{1+e^{(y^*_1-\alpha_2)}} - \frac{1}{1+e^{(y^*_1-\alpha_1)}} \\ &= \frac{1}{1+e^{(-93,335-(-89,278))}} - \frac{1}{1+e^{(-93,335-(-97,302))}} \\ &= 0,964 \end{aligned}$$

$$\begin{aligned} \text{Kel. 3} \rightarrow \rho(y_k = 3) &= 1 - \frac{1}{1+e^{(y^*_1-\alpha_2)}} = 1 - \frac{1}{1+e^{(-93,335-(-89,278))}} \\ &= 0,017 \end{aligned}$$

Dari nilai peluang kumulatif logit tersebut yang terbesar adalah kelompok 2. Sehingga Bank AGRIS dengan CAR = 17,58; EAQ = 99,56; NPL = 0,66; ROA = 0,29; NIM = 2,78; STM = 104,39; LDR = 70,02, dikelompokkan ke kategori kondisi bank Sehat (sesuai dengan kondisi sebenarnya). Kemudian dilakukan prediksi dengan model analisis regresi logistik ordinal tersebut pada seluruh data *training* sejumlah 80 Bank dan pada data *testing* sejumlah 20 Bank. Hasil prediksi klasifikasi kondisi bank dengan menggunakan model analisis regresi logistik ordinal pada data *testing* dapat dilihat pada Tabel 9.

Tabel 9. Hasil Prediksi Klasifikasi Kondisi Bank dengan Model Analisis Regresi Logistik Ordinal

Nama Bank	Kondisi Bank (Y) Aktual	Kondisi Bank (Y) Prdiksi (Logistik Ordinal)	Ketepatan Klasifikasi
Bank Andara	2	2	Benar
Bank Anglomas Internasional	1	1	Benar
Bank Commonwealth	1	2	Salah
Bank Dinar Indonesia	2	2	Benar
Bank GANESHA	3	1	Salah
Bank Mayapada International Tbk	2	1	Salah
Bank Mizuho Indonesia	2	1	Salah
Bank Nusantara Parahyangan Tbk	1	2	Salah
Bank Pundi Indonesia Tbk	2	2	Benar
Bank Resona Perdana	1	1	Benar
Bank Sumitomo Mitsui Indonesia	2	1	Salah
BPD Riau	1	1	Benar
BPD Sulawesi Selatan	1	1	Benar
BPD Sulawesi Tengah	1	1	Benar
BPD Sulawesi Tenggara	1	1	Benar
BPD Sulawesi Utara	1	1	Benar
BPD Sumatera Barat	1	1	Benar
BPD Sumatera Selatan	1	1	Benar
BPD Sumatera Utara	1	1	Benar
The Bank Of Tokyo Mitsubishi	2	3	Salah

3.4 Perbandingan Analisis Diskriminan dan Analisis Regresi Logistik Ordinal

Dari penghitungan prediksi klasifikasi kondisi bank pada data *training* dan data *testing* (data klasifikasi kondisi bank di Indonesia tahun 2014) dengan menggunakan analisis diskriminan dan analisis regresi logistik ordinal diperoleh prediksi klasifikasi kondisi bank. Prediksi klasifikasi kondisi bank tersebut akan dibandingkan dengan klasifikasi kondisi bank aktualnya, kemudian akan dilihat akurasi dari prediksi masing-masing analisis dengan menggunakan ukuran akurasi APER .

Dari perbandingan prediksi klasifikasi kondisi bank masing-masing analisis dengan klasifikasi kondisi bank aktualnya, diperoleh nilai APER untuk analisis diskriminan pada data training sebesar 0,15 dan pada data testing sebesar 0,25. Sedangkan nilai APER untuk analisis regresi logistik ordinal pada data training sebesar

0,1625 dan pada data testing sebesar 0,35. Nilai APER disajikan secara ringkas pada Tabel 10.

Tabel 10. Nilai APER Analisis Diskriminan dan Regresi Logistik Ordinal

Model	Nilai APER		
	Data <i>Training</i>	Data <i>Testing</i>	Rata-Rata
Analisis Diskriminan	0,15 (15%)	0,25 (25%)	0,20 (20%)
Analisis Regresi Logistik Ordinal	0,1625 (16,25%)	0,35 (35%)	0,2562 (25,62%)

Berdasarkan nilai APER tersebut dapat dikatakan analisis diskriminan lebih baik dalam memprediksi klasifikasi kondisi bank, hal ini dikarenakan analisis diskriminan memiliki nilai APER yang lebih kecil daripada analisis regresi logistik ordinal. Analisis diskriminan secara rata-rata memiliki nilai ketepatan prediksi klasifikasi sebesar 80%, sedangkan analisis regresi logistik ordinal hanya sebesar 74,38%.

4. Kesimpulan

Dengan menggunakan model analisis diskriminan pada data *training* dan data *testing*, diperoleh ketepatan prediksi sebesar 85% dan 75% (rata-rata 80%). Sedangkan, dengan menggunakan model analisis regresi logistik ordinal pada data *training* dan data *testing*, diperoleh ketepatan prediksi sebesar 83,75% dan 65% (rata-rata 74,38%). Sehingga dapat dikatakan bahwa dalam prediksi klasifikasi kondisi bank di Indonesia tahun 2014, analisis diskriminan lebih baik digunakan daripada analisis regresi logistik ordinal.

Daftar Pustaka

- [1] Johnson, R.A., and D.W. Wichern, 2007, *Applied Multivariate Statistical Analysis*, Prentice Hall, New Jersey.
- [2] Hosmer, D.W., and S. Lemeshow, 2000, *Applied Logistic Regression (Second ed.)*, John Wiley and Sons, New York.