

Nama : Yuniar Ayu Rachmadini
NRP : 2043201103

Analisis Regresi Logistik Biner

PENERAPAN REGRESI LOGISTIK BINER FAKTOR-FAKTOR YANG MEMPENGARUHI BERAT BADAN BAYI USIA 6 BULAN

A. Studi Kasus

Masa bayi hingga balita merupakan periode penting dalam proses tumbuh kembang manusia. Perkembangan dan pertumbuhan di masa itu menjadi penentu keberhasilan pertumbuhan dan perkembangan anak di periode selanjutnya. Masa tumbuh kembang di usia ini merupakan masa yang berlangsung cepat dan tidak akan pernah terulang, karena itu seringkali disebut golden age atau masa keemasan. Balita merupakan istilah umum bagi anak usia 1-3 tahun (batita) dan anak prasekolah (3-5 tahun). Saat usia balita, anak masih tergantung penuh kepada orang tua untuk melakukan kegiatan penting, seperti mandi, buang air dan makan. Perkembangan berbicara dan berjalan sudah bertambah baik, namun kemampuan lain masih terbatas. Dalam proses tumbuh kembang, berat badan menjadi salah satu faktor penting yang harus diperhatikan. Berat badan adalah suatu ukuran untuk menilai keadaan gizi seseorang. Berat badan ideal adalah bobot optimal dari tubuh untuk menjaga kesehatan dan kebugaran. Rentang dari berat badan ideal seseorang dapat diperhitungkan berdasarkan berbagai macam faktor, di antaranya ras, jenis kelamin, usia, serta tinggi badan.

Pada penelitian ini, akan dianalisis faktor penyebab bayi memiliki berat badan ideal pada umur 6 bulan. Menggunakan data sekunder yang diperoleh melalui Tugas Akhir, berupa data berat badan bayi usia 6 bulan di Posyandu Kartini dan Posyandu Mawar sebanyak 107 sampel. Sehingga, dengan diketahuinya faktor – faktor yang berpengaruh terhadap berat badan ideal bayi, serta diharapkan dapat dilakukan upaya untuk mencapai berat badan ideal bayi dan dapat menjadi referensi bagi pemerintah untuk menentukan program selanjutnya dan memberi informasi untuk penyuluhan kesehatan bayi.

B. Metode Analisis

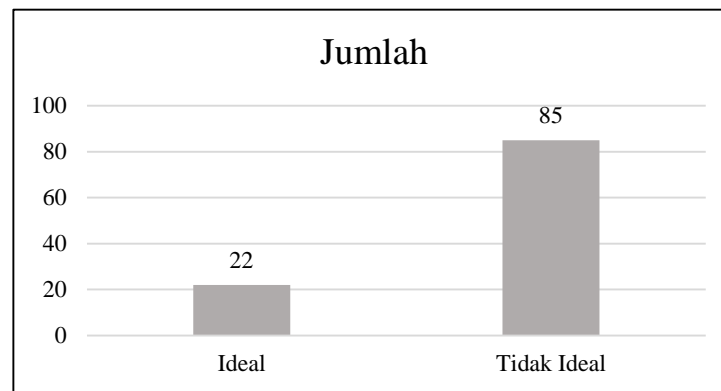
Pada penelitian ini, analisis data akan dilakukan dengan menggunakan konsep regresi logistik biner yang merupakan analisis data yang mendeskripsikan antara sebuah variabel respon dan satu atau lebih variabel penjelas atau prediktor dengan variasi variabel respon berskala nominal yang berupa dua kategori “sukses” atau “gagal”, sedangkan variabel prediktor dapat berupa data berskala ordinal atau data berskala rasio.

Menurut Hosmer & Lemeshow (2000), pada regresi linier variabel respon diasumsikan berdistribusi normal, sedangkan variabel respon pada regresi logistik biner mengikuti distribusi Bernoulli dengan fungsi probabilitas. Tujuan regresi logistik biner adalah mencari pola hubungan antara prediktor (X) dengan $\pi(X_i)$ dimana $\pi(X_i)$ adalah probabilitas kejadian yang diakibatkan variabel X . Sehingga hasil fungsi logistik kemungkinan bernilai 0 atau 1.

C. Analisis dan Pembahasan

Bagian ini berisi hasil analisis regresi logistik biner pada faktor- faktor penyebab bayi memiliki berat badan ideal pada umur 6 bulan. Variabel prediktor yang digunakan adalah jenis kelamin bayi, pemberian ASI, BMI bayi, profesi ibu, dan kesejahteraan keluarga bayi.

Analisis statistika deskriptif secara umum menjelaskan karakteristik dari data yaitu berat badan bayi pada usia 6 bulan di Posyandu Kartini dan Posyandu Mawar Kecamatan Tandes Surabaya berdasarkan BMI (Body Mass Index) untuk menentukan kecenderungan berat badan bayi yang ideal dan tidak ideal.



Gambar 1 Berat Badan Bayi Sebagai Variabel Respon (Y)

Deskripsi berat badan meliputi sebanyak 22 bayi yang memiliki berat badan ideal (21%) sedangkan bayi yang memiliki berat badan tidak ideal sebanyak 85 bayi (79%), hal ini menunjukkan bahwa mayoritas bayi yang melakukan pemeriksaan di Posyandu Kartini dan Posyandu Mawar dalam keadaan memiliki berat badan tidak Ideal.

Tabel 1 Berat Badan Bayi Usia 6 Bulan Menurut Profesi Ibu

BMI	Profesi Ibu		Total
	Bekerja (0)	Tidak Bekerja (1)	
Ideal (1)	12 (11,2%)	10 (9,3%)	22 (20,5%)
Tidak Ideal (0)	45 (42,1%)	40 (37,4%)	85 (79,5%)
Total	57 (53,3%)	50 (46,7%)	107(100%)

Tabel 1 menunjukkan bahwa dari total 107 bayi, profesi ibu bekerja yang memiliki bayi dengan berat badan ideal sebanyak 12 bayi (11,2%) dan bayi yang memiliki berat badan tidak ideal sebanyak 45 bayi (42,1%). Sedangkan profesi ibu tidak bekerja yang memiliki bayi dengan berat badan ideal sebanyak 10 bayi (9,3%) dan bayi yang memiliki berat badan tidak ideal sebanyak 40 bayi (37,4%).

Tabel 2 Berat Badan Bayi Usia 6 Bulan Menurut Pemberian ASI Eksklusif

BMI	Asi Eksklusif		Total
	Tidak (0)	Ya (1)	
Ideal (1)	3(2,8%)	19(17,7%)	22(20,5%)
Tidak Ideal (0)	75(70,1%)	10(9,4%)	85(79,5%)
Total	78(72,9%)	29(27,1%)	107(100%)

Tabel 2 menunjukkan bahwa dari total 107 bayi, terdapat 75 bayi (70,1%) yang tidak mendapatkan ASI Eksklusif memiliki BMI yang tidak ideal dan 3 bayi(2,8%) lainnya memiliki BMI ideal. Sedangkan 19 bayi (17,7%) yang mendapatkan ASI Eksklusif memiliki BMI yang ideal dan 10 bayi (9,4%) lainnya memiliki BMI tidak ideal. Hal tersebut mengindikasikan bahwa pemberian ASI Eksklusif pada 6 bulan pertama kelahiran memberikan pengaruh terhadap BMI bayi.

Tabel 3 Berat Badan Bayi Usia 6 Bulan Menurut Kesejahteraan Keluarga

BMI	Kesejahteraan		Total
	Gakin (0)	Non Gakin (1)	
Ideal (1)	1(0,09%)	21(19,6%)	22(20,5%)
Tidak Ideal (0)	22(20,6%)	63(58,9%)	85(79,5%)
Total	23(21,5%)	84(78,5%)	107(100%)

Tabel 3 menunjukkan bahwa dari total 107 bayi, Gakin dengan bayi yang memiliki berat badan ideal sebanyak 1 bayi (0,09%) dan bayi yang memiliki berat badan tidak ideal sebanyak 22 bayi (20,6%). Sedangkan tidak gakin dengan bayi yang memiliki berat badan ideal sebanyak 21 bayi (19,6%) dan bayi yang memiliki berat badan tidak ideal sebanyak 63 bayi (58,9%).

Tabel 4 Uji Independensi Berat Badan Bayi Usia 6 Bulan

Variabel	Chi Square	df	Chi Square	Keputusan	Keterangan
			Tabel		
Profesi Ibu (X ₁)	0,18	1	3,841	Gagal Tolak H ₀	Tidak Ada Hubungan
Paritas (X ₂)	2,161	1	3,841	Gagal Tolak H ₀	Tidak Ada Hubungan
ASI Eksklusif (X ₃)	49,226	1	3,841	Tolak H ₀	Ada Hubungan
Tingkat Kesejahteraan Keluarga (X ₄)	4,715	1	3,841	Tolak H ₀	Ada Hubungan

Uji independensi digunakan untuk mengetahui hubungan antara dua variabel. Tabel 4 menunjukkan bahwa variabel yang memiliki hubungan dengan berat badan bayi pada usia 6 bulan adalah ASI eksklusif dan tingkat Kesejahteraan keluarga. Hal tersebut didasarkan pada *chi square* lebih besar daripada *chi square*.

Tabel 5 Estimasi Parameter

Variabel	β	Wald	Chi-Square Tabel	Odds Ratio
Profesi Ibu X1 ₍₁₎	-0,467	0,480	9,488	0,627
Paritas X2 ₍₁₎	0,185	0,055	9,488	1,204
Pemberian ASI Eksklusif X3 ₍₁₎	3,795	25,280	9,488	44,488
Tingkat Kesejahteraan Keluarga X4 ₍₁₎	0,568	0,212	9,488	1,764

Nilai Chi-Square yang dihasilkan dari pengujian tersebut sebesar 46,823. Nilai tersebut lebih besar dari Chi-Square tabel sebesar 9,488. Dari hasil tersebut disimpulkan bahwa $G > \chi^2$ ($46,823 > 9,488$) diartikan memiliki satu atau lebih dari satu variabel yang berpengaruh signifikan dan diperoleh model estimasi parameter sebagai berikut :

$$\pi(x) = \frac{e^{-3,499-0,467+0,185+3,795+0,568}}{1 + e^{-3,499-0,467+0,185+3,795+0,568}}$$

Kemudian, dilanjutkan pengujian parsial. Hasil pengujian parsial dengan semua variabel yang diikutsertakan dalam pemodelan akan menghasilkan beberapa variabel yang signifikan adalah sebagai berikut.

Tabel 6 Uji Parsial

Keterangan Variabel	B	Wald	P-Value	Exp(B)
X3(1) Pemberian ASI Eksklusif (1)	3,861	29,852	0,000	47,500
Constant	-3,219	29,888	0,000	0,040

Menunjukkan bahwa variabel yang memiliki hubungan dengan berat badan bayi pada usia 6 bulan adalah ASI eksklusif (X3). Berdasarkan hasil dari pengujian regresi logistik biner secara serentak pada Tabel 4.7 maka diperoleh model regresi logistik adalah sebagai berikut.

$$\hat{g}(x) = -3,219 + 3,861X_{3(1)}$$

$$\hat{\pi}(x) = \frac{e^{g(x)}}{1 + e^{g(x)}} = \frac{\exp(-3,219 + 3,861x_{3(1)})}{1 + \exp(-3,219 + 3,861x_{3(1)})} = 0,655$$

Berdasarkan hasil perhitungan model diperoleh kesimpulan bahwa peluang pemberian ASI eksklusif pada bayi akan berdampak sebesar 0,655 untuk memiliki berat badan ideal. Artinya jika terdapat 100 bayi yang diberikan ASI eksklusif maka 65 bayi akan memiliki berat badan ideal. Sedangkan peluang tidak pemberian ASI eksklusif akan berdampak sebesar 0,35 untuk memiliki berat badan ideal. Artinya jika terdapat 100 bayi yang tidak diberikan ASI eksklusif maka 35 bayi akan memiliki berat badan ideal.

Tabel 7 Odds Ratio Model Regresi Logistik Biner

Variabel	Odds Ratio
ASI Eksklusif	47,50

Tabel 7 menunjukkan bahwa nilai odds ratio antara variabel ASI Eksklusif dengan berat badan bayi sebesar 47,5, yang artinya bayi yang diberi ASI eksklusif memiliki kecenderungan 47,500 kali lebih besar memiliki berat badan ideal dibandingkan dengan bayi yang tidak diberi ASI eksklusif.

Tabel 8 Odds Ratio Model Regresi Logistik Biner

Chi-Square	Df	Keputusan
2,173	2	Gagal Tolak Ho

Tabel 8 menunjukkan bahwa nilai $C > \chi^2_{(db,\alpha)}$ yaitu $2,173 < 5,991$, maka keputusan yang diambil adalah gagal tolak H_0 . Sehingga dapat diketahui bahwa model telah sesuai atau tidak terdapat perbedaan yang signifikan antara hasil pengamatan dan kemungkinan hasil prediksi model.

D. Kesimpulan

Faktor – Faktor yang mempengaruhi secara signifikan terhadap berat badan bayi pada usia 6 bulan yaitu pemberian ASI eksklusif, berdasarkan hasil perhitungan model diperoleh kesimpulan bahwa peluang pemberian ASI eksklusif pada bayi akan berdampak sebesar 0,655 untuk memiliki berat badan ideal . artinya jika terdapat 100 bayi yang diberikan ASI eksklusif maka 65 bayi akan memiliki berat badan ideal. Sedangkan peluang tidak pemberian ASI eksklusif akan berdampak sebesar 0,35 untuk memiliki berat badan ideal. Artinya jika terdapat 100 bayi yang tidak diberikan ASI eksklusif maka 35 bayi akan memiliki berat badan ideal.

Analisis Regresi Logistik Multinomial

PEMODELAN FAKTOR-FAKTOR YANG MEMPENGARUHI PENYAKIT PERNAPASAN MENGGUNAKAN REGRESI LOGISTIK MULTINOMIAL

A. Studi Kasus

Pernafasan atau respirasi adalah suatu proses mulai dari pengambilan oksigen, pengeluaran karbohidrat hingga penggunaan energi di dalam tubuh. Manusia dalam bernapas menghirup oksigen dalam udara bebas dan membuang karbon dioksida ke lingkungan. Saluran pernapasan adalah bagian tubuh manusia yang berfungsi sebagai tempat pertukaran gas yang diperlukan untuk proses pernapasan. Saluran ini berpangkal pada hidung, tekak (faring), tenggorokan (trakea), cabang tenggorokan (bronkus), bronkiolus, alveolus, dan berakhir pada paru-paru. Namun, dalam organ-organ tersebut dapat mengalami gangguan. Gangguan ini biasanya berupa kelainan, penyakit, atau karena ulah manusia itu sendiri (seperti merokok). Penyakit atau gangguan yang menyerang sistem pernapasan ini dapat menyebabkan terganggunya proses pernapasan. Adapun faktor-faktor pendukung yang dapat menyebabkan munculnya penyakit pernapasan seperti jenis kelamin, usia, riwayat merokok, dan lain sebagainya.

Pada penelitian ini, akan dilakukan analisis lebih lanjut mengenai hubungan jenis kelamin, umur, dan riwayat merokok terhadap penyakit pernapasan. Menggunakan data sekunder yang diperoleh dari website UCL Machine Learning yang berjudul “Exasens Data Set” sebanyak 399 data. Sehingga dapat diketahui apakah faktor-faktor tersebut berpengaruh secara signifikan menyebabkan seseorang terjangkit penyakit pernapasan. Serta diharapkan dapat bermanfaat bagi pembaca dan instansi terkait.

B. Metode Analisis

Penelitian ini dilakukan dengan menggunakan metode regresi logistik multinomial yang merupakan regresi logistik yang digunakan saat variabel dependen mempunyai skala yang bersifat *polichotomous* atau multinomial. Atau dengan arti lain, variabel respon yang digunakan pada regresi logistik multinomial mengikuti distribusi multinomial dimana distribusi multinomial merupakan generalisasi dari distribusi binomial dengan menggunakan lebih dari dua kategori.

C. Analisis dan Pembahasan

Bagian ini berisi hasil analisis regresi logistik biner pada faktor-faktor penyebab bayi memiliki berat badan ideal pada umur 6 bulan. Variabel prediktor yang digunakan adalah jenis kelamin, usia, dan riwayat merokok.

Tabel 1 Tabel Kontingensi Jenis Kelamin

Diagnosis Penyakit Pernapasan	Jenis Kelamin		Total
	Perempuan	Laki-laki	
COPD	23	56	79
HC	104	56	160
Asthma	55	25	80

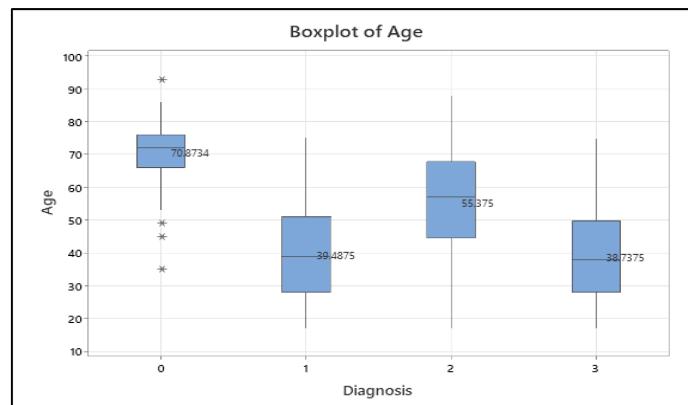
Infected	58	22	80
Total	240	159	399

Tabel 1 menunjukkan bahwa orang dengan jenis kelamin perempuan memiliki kecenderungan terdiagnosis penyakit pernapasan jenis HC yang diketahui dari 160 orang pengidap HC dengan frekuensi sebanyak 104 orang, sedangkan orang dengan jenis kelamin laki-laki memiliki kecenderungan terdiagnosis penyakit pernapasan jenis *Infected* yang diketahui dari 80 orang pengidap *Infected* dengan frekuensi sebanyak 58 orang. Serta dapat diketahui, mayoritas sampel yang diamati terdiagnosis mengidap penyakit pernapasan jenis HC yang diketahui dari 399 orang terdiagnosis penyakit pernapasan dengan frekuensi sebanyak 160.

Tabel 2 Tabel Kontingensi Status Merokok

Diagnosis Penyakit Pernapasan	Status Merokok			Total
	Tidak Pernah	Pernah	Aktif	
COPD	6	63	10	79
HC	93	35	32	160
Asthma	36	35	9	80
Infected	44	17	19	80
Total	179	150	70	399

Tabel 2 menunjukkan bahwa orang yang tidak pernah merokok memiliki kecenderungan terdiagnosis mengidap penyakit pernapasan jenis HC yang diketahui dari 160 orang pengidap HC dengan frekuensi sebanyak 93 orang dan tidak memiliki kecenderungan terdiagnosis mengidap penyakit pernapasan jenis COPD yang diketahui dari 79 orang pengidap COPD dengan frekuensi sebanyak 6 orang, sedangkan orang yang pernah merokok memiliki kecenderungan terdiagnosis mengidap penyakit pernapasan jenis COPD yang diketahui dari 79 orang pengidap COPD dengan frekuensi sebanyak 63 orang.



Gambar 1 Boxplot Usia

Gambar 1 menunjukkan bahwa seseorang yang terdiagnosis mengidap COPD memiliki rata-rata usia 70.87, terdiagnosis mengidap HC memiliki rata-rata usia 39.49, terdiagnosis mengidap Asthma memiliki rata-rata usia 55.34, dan terdiagnosis mengidap Infected memiliki rata-rata usia 38.74. Artinya, seseorang yang terdiagnosis penyakit pernapasan jenis COPD cenderung berusia lanjut.

Tabel 3 Statistik Uji *Goodness of Fit*

Variabel	df	G^2_{hitung}	$\chi^2_{0,05}$	P-Value
Jenis Kelamin*Diagnosis Penyakit Pernapasan	3	40,609	7,81	0,00
Status Merokok*Diagnosis Penyakit Pernapasan	6	98,447	12,59	0,00

Tabel 3 menunjukkan bahwa dari hasil uji *goodness of fit* dari jenis kelamin dan diagnosis penyakit pernapasan diperoleh G^2_{hitung} sebesar 40,609 yang lebih besar dari $\chi^2_{0,05(3)}$ sebesar 7,81 serta diperkuat dengan P-Value sebesar 0,00 yang lebih besar dari nilai α sebesar 0,05 sehingga diputuskan tolak H_0 , artinya terdapat hubungan antara jenis kelamin dan diagnosis penyakit pernapasan (dependen). Sedangkan hasil uji *goodness of fit* dari status merokok dan diagnosis penyakit pernapasan diperoleh G^2_{hitung} sebesar 98,447 yang lebih besar dari $\chi^2_{0,05(6)}$ sebesar 12,59 serta diperkuat dengan P-Value sebesar 0,00 yang lebih kecil dari nilai α sebesar 0,05 sehingga diputuskan tolak H_0 , artinya terdapat hubungan antara status merokok dan diagnosis penyakit pernapasan (dependen).

Model logit yang terbentuk dari variabel-variabel yang berpengaruh adalah sebagai berikut.

$$g_1(x) = 10,164 - 0,183X_1 + 0,779X_{2(0)} + 2,934X_{3(0)} + 0,054X_{3(1)} \\ = 10,164 - 0,183(30) + 0,779(1) + 2,934(1) + 0,054(0) = 8,39$$

$$g_2(x) = 5,740 - 0,116X_1 + 1,137X_{2(0)} + 3,058X_{3(0)} + 0,748X_{3(1)} \\ = 5,740 - 0,116(30) + 1,137(1) + 3,058(1) + 0,748(0) = 6,47$$

$$g_3(x) = 9,520 - 0,186X_1 + 1,124X_{2(0)} + 2,709X_{3(0)} - 0,111X_{3(1)} \\ = 9,520 - 0,186(30) + 1,124(1) + 2,709(1) - 0,111(0) = 7,77$$

Berdasarkan model logit di atas dapat dibentuk model probit sebagai berikut.

$$\pi_1(x) = \frac{\exp(g_1(x))}{1 + \exp(g_1(x)) + \exp(g_2(x)) + \exp(g_3(x))}$$

$$\pi_1(x) = \frac{\exp(8,39)}{1 + \exp(8,39) + \exp(6,47) + \exp(7,77)}$$

$$\pi_1(x) = \frac{4412,81}{1 + 4412,81 + 645,39 + 2376,23} = 0,59$$

Model probit untuk $g_1(x)$ menunjukkan bahwa peluang seorang perempuan berumur 30 tahun yang mengidap penyakit jenis HC adalah sebesar 0,59.

$$\pi_2(x) = \frac{\exp(g_2(x))}{1 + \exp(g_1(x)) + \exp(g_2(x)) + \exp(g_3(x))}$$

$$\pi_2(x) = \frac{\exp(6,47)}{1 + \exp(8,39) + \exp(6,47) + \exp(7,77)}$$

$$\pi_2(x) = \frac{645,39}{1 + 4412,81 + 645,39 + 2376,23} = 0,087$$

Model probit untuk $g_2(x)$ menunjukkan bahwa peluang seorang perempuan berumur 30 tahun yang mengidap penyakit jenis Asthma adalah sebesar 0,087.

$$\pi_3(x) = \frac{\exp(g_3(x))}{1 + \exp(g_1(x)) + \exp(g_2(x)) + \exp(g_3(x))}$$

$$\pi_3(x) = \frac{\exp(7,77)}{1 + \exp(8,39) + \exp(6,47) + \exp(7,77)}$$

$$\pi_3(x) = \frac{2376,23}{1 + 4412,81 + 645,39 + 2376,23} = 0,319$$

Model probit untuk $g_3(x)$ menunjukkan bahwa peluang seorang perempuan berumur 30 tahun yang mengidap penyakit jenis Infected adalah sebesar 0,319.

Tabel 4 Statistik Uji Serentak

χ^2	DB	$\chi^2_{(0,05,12)}$	P-VALUE
291,91	12	21,03	0,00

Tabel 4 menunjukkan bahwa uji serentak dari usia, jenis kelamin, dan status pernikahan menghasilkan χ^2 sebesar 291,91 yang lebih besar dari $\chi^2_{(0,05,12)}$ sebesar 21,03 serta diperkuat dengan nilai P-Value sebesar 0,00 yang kurang dari α sebesar 0,05 sehingga diputuskan tolak H_0 , artinya minimal ada satu variabel prediktor yang berpengaruh signifikan terhadap diagnosis penyakit pernapasan.

Tabel 5 Uji Parsial

Diagnosis Penyakit Pernapasan	Variabel	W	$\chi^2_{(0,05,1)}$	P-Value	Keputusan
HC	Usia	68,125	3,841	0,00	Tolak
	Jenis Kelamin (0)	3,212	3,841	0,073	Gagal Tolak
	Status Merokok (0)	15,023	3,841	0,00	Tolak
	Status Merokok (1)	0,009	3,841	0,926	Gagal Tolak
Asthma	Usia	30,187	3,841	0,00	Tolak
	Jenis Kelamin (0)	7,529	3,841	0,006	Tolak
	Status Merokok (0)	15,737	3,841	0,00	Tolak
	Status Merokok (1)	1,547	3,841	0,214	Gagal Tolak
Infected	Usia	65,438	3,841	0,00	Tolak
	Jenis Kelamin (0)	5,609	3,841	0,018	Tolak
	Status Merokok (0)	12,015	3,841	0,001	Tolak
	Status Merokok (1)	0,031	3,841	0,860	Gagal Tolak

Tabel 5 menunjukkan bahwa dari hasil uji parsial diketahui variabel usia dan status tidak pernah merokok didapatkan keputusan tolak H_0 karena nilai W berturut-turut sebesar 68,125 dan 15,023 lebih dari $\chi^2_{(0,05,1)}$ sebesar 3,841 dan diperkuat oleh nilai P-value sebesar 0,00 yang kurang dari nilai α sebesar 0,05 artinya variabel usia dan status merokok berpengaruh signifikan terhadap diagnosis penyakit pernapasan jenis HC. Variabel usia, jenis kelamin, dan status tidak pernah merokok didapatkan keputusan tolak H_0 karena nilai W berturut-turut sebesar 30,187; 7,529, dan 15,737 lebih dari $\chi^2_{(0,05,1)}$ sebesar 3,841 dan diperkuat oleh nilai P-value sebesar 0,00 yang kurang dari nilai α sebesar 0,05 artinya variabel usia, jenis kelamin, dan status merokok berpengaruh signifikan terhadap diagnosis penyakit pernapasan jenis Asthma. Variabel usia, jenis kelamin, dan status tidak pernah merokok didapatkan keputusan tolak H_0 karena nilai W berturut-turut sebesar 65,438; 5,609, dan 12,015 lebih dari $\chi^2_{(0,05,1)}$ sebesar 3,841 dan diperkuat oleh nilai P-value berturut-turut

sebesar 0,00; 0,018, dan 0,001 yang kurang dari nilai α sebesar 0,05 artinya variabel usia, jenis kelamin, dan status merokok berpengaruh signifikan terhadap diagnosis penyakit pernapasan jenis Infected.

Tabel 6 Uji Kesesuaian Model

χ^2	$\chi^2_{(0.05,549)}$	P-Value
502,19	604,62	0,924

Tabel 6 menunjukkan bahwa uji kesesuaian model dari usia, jenis kelamin, dan status merokok menghasilkan χ^2 sebesar 502,19 yang lebih kecil dari $\chi^2_{(0.05,549)}$ sebesar 604,62 serta diperkuat dengan nilai P-Value sebesar 0,924 yang lebih besar dari nilai α sebesar 0,05 sehingga diputuskan gagal tolak H_0 , artinya tidak ada perbedaan antara observasi dan prediksi (model sesuai).

Tabel 7 Odds Ratio

Diagnosis Penyakit Pernapasan	Variabel	Exp(β)
HC	Usia	0,833
	Jenis Kelamin (0)	2,18
	Status Merokok (0)	18,81
	Status Merokok (1)	1,055
Asthma	Usia	0,891
	Jenis Kelamin (0)	3,117
	Status Merokok (0)	21,28
	Status Merokok (1)	2,113
Infected	Usia	0,83
	Jenis Kelamin (0)	3,078
	Status Merokok (0)	15,021
	Status Merokok (1)	0,895

Tabel 7 menunjukkan bahwa seseorang yang mengidap penyakit HC dengan jenis kelamin perempuan memiliki kecenderungan 2,18 kali daripada seseorang berjenis kelamin laki-laki. Seseorang yang memiliki status tidak pernah merokok memiliki kecenderungan 18,81 kali mengidap penyakit HC daripada seseorang yang pernah dan aktif merokok. Serta seseorang yang memiliki status pernah merokok memiliki kecenderungan 1,055 kali mengidap penyakit HC daripada seseorang yang tidak pernah dan aktif merokok.

Seseorang yang mengidap penyakit Asthma dengan jenis kelamin perempuan memiliki kecenderungan 3,117 kali daripada seseorang berjenis kelamin laki-laki. Seseorang yang memiliki status tidak pernah merokok memiliki kecenderungan 21,28 kali mengidap penyakit Asthma daripada seseorang yang pernah dan aktif merokok. Serta seseorang yang memiliki status pernah merokok memiliki kecenderungan 2,113 kali mengidap penyakit Asthma daripada seseorang yang tidak pernah dan aktif merokok.

Seseorang yang mengidap penyakit Infected dengan jenis kelamin perempuan memiliki kecenderungan 3,078 kali daripada seseorang berjenis kelamin laki-laki. Seseorang yang memiliki status tidak pernah merokok memiliki kecenderungan 15,021 kali mengidap

penyakit Infected daripada seseorang yang pernah dan aktif merokok. Serta seseorang yang memiliki status pernah merokok memiliki kecenderungan 0,895 kali mengidap penyakit Infected daripada seseorang yang tidak pernah dan aktif merokok.

Tabel 8 Ketepatan Klasifikasi

Observasi		Prediksi				
		Diagnosis Penyakit Pernapasan				Persentase Ketepatan
		COPD	HC	Asthma	Infected	
Diagnosis Penyakit Pernapasan	COPD	64	4	11	0	81,0%
	HC	3	143	14	0	89,4%
	Asthma	21	37	22	0	27,5%
	Infected	0	74	6	0	0,0%
Total Persentase						57,4%

Tabel 8 menunjukkan bahwa seseorang yang terdiagnosis penyakit pernapasan jenis COPD memiliki persentase kebenaran sebesar 81% dengan rincian 64 orang diklasifikasikan dengan tepat dan 15 orang tidak tepat, seseorang yang terdiagnosis penyakit pernapasan jenis HC memiliki persentase kebenaran sebesar 89,4% dengan rincian 143 orang diklasifikasikan dengan tepat dan 17 orang tidak tepat, seseorang yang terdiagnosis penyakit pernapasan jenis Asthma memiliki persentase kebenaran sebesar 27,5% dengan rincian 22 orang diklasifikasikan dengan tepat dan 58 orang tidak tepat, serta seseorang yang terdiagnosis penyakit pernapasan jenis Infected memiliki persentase kebenaran sebesar 0,0% dengan rincian 80 orang diklasifikasikan tidak tepat. Diperoleh ketepatan klasifikasi sebesar 57,4%, sehingga nilai APER (*Apparent Error Rate*) adalah sebesar 42,6% yang dihitung dari 100% dikurangi dengan persentase ketepatan klasifikasi.

D. Kesimpulan

Seorang perempuan memiliki kecenderungan terdiagnosis penyakit pernapasan jenis HC, sedangkan seorang laki-laki memiliki kecenderungan terdiagnosis penyakit pernapasan jenis Infected. Berdasarkan uji independensi didapatkan hasil bahwa terdapat hubungan antara usia, jenis kelamin, status merokok, dan hasil diagnosis penyakit pernapasan. Hasil analisis regresi logistik multinomial menunjukkan bahwa variabel usia, jenis kelamin, dan status merokok berpengaruh signifikan terhadap hasil diagnosis penyakit pernapasan. Hasil uji kesesuaian model telah sesuai dengan kebaikan model serta persentase kebenaran dalam mengklasifikasikan orang dengan penyakit pernapasan yang cukup baik.

Analisis Regresi Logistik Ordinal

ANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI KELAS PERJALANAN PESAWAT MENGGUNAKAN REGRESI LOGISTIK ORDINAL

A. Studi Kasus

Di Indonesia, perkembangan industri penerbangan atau transportasi udara terjadi begitu pesat. Hal ini dapat dilihat dari banyaknya maskapai penerbangan yang sekarang telah melayani berbagai rute baik domestik maupun internasional. Bahkan persaingan dalam jasa penerbangan ini pun semakin ketat dengan munculnya maskapai pesawat berbiaya murah atau low cost carrier. Terdapat berbagai macam pilihan kelas kursi penerbangan yang telah disesuaikan dengan fasilitas yang memadai, yang mana dalam pengambilan keputusannya konsumen seringkali mempertimbangkan jarak perjalanan, tipe perjalanan, hingga kualitas pelayanan yang diberikan oleh maskapai penerbangan.

Berdasarkan permasalahan di atas, peneliti akan melakukan analisis mengenai hubungan jarak penerbangan, tipe perjalanan, dan kepuasan konsumen terhadap pemilihan kelas penerbangan dengan menggunakan analisis regresi logistik ordinal. Data yang digunakan adalah data sekunder yang diperoleh dari *website Kaggle* yaitu dataset yang berjudul “*Airlines Passenger Satisfaction*”. Sehingga dapat diperoleh hasil penelitian yang sesuai dan akurat. Serta diharapkan dapat bermanfaat bagi pembaca dan juga instansi terkait.

B. Metode Analisis

Penelitian ini dilakukan menggunakan metode regresi logistik ordinal yang merupakan suatu metode analisis yang digunakan untuk mengetahui hubungan antara variabel respon dengan variabel prediktor, dimana variabel respon berskala ordinal yang terdiri atas tiga kategori atau lebih dan skala pengukurannya bersifat tingkatan. Model yang dipakai untuk regresi logistik ordinal adalah model logit kumulatif. Model logit kumulatif merupakan model yang diperoleh dengan cara membandingkan peluang kumulatif yaitu peluang kurang dari atau sama dengan kategori respon ke- j $P(Y \leq j | x_i)$ dengan peluang lebih besar dari kategori respon ke- j $P(Y > j | x_i)$.

C. Analisis dan Pembahasan

Tipe perjalanan merupakan data kategorik yang terdiri dari dua kategori yaitu penerbangan pribadi dan bisnis, sedangkan pada kelas penerbangan merupakan data kategorik yang terdiri dari kelas *eco*, *eco plus*, dan *business*.

Tabel 1 Karakteristik Tipe perjalanan

Tipe perjalanan	Kelas Penerbangan			Total
	<i>Eco</i>	<i>Eco Plus</i>	<i>Business</i>	
Bisnis	32	27	19	78
Pribadi	2	26	17	45
Total	34	53	36	123

Tabel 1 menunjukkan bahwa terdapat 32 orang dengan tipe perjalanan bisnis cenderung memilih kelas penerbangan *eco*, terdapat 17 orang dengan tipe perjalanan pribadi

cenderung memilih kelas penerbangan *business*. Tipe perjalanan dari 123 orang menunjukkan bahwa terdapat 78 orang melakukan tipe perjalanan bisnis.

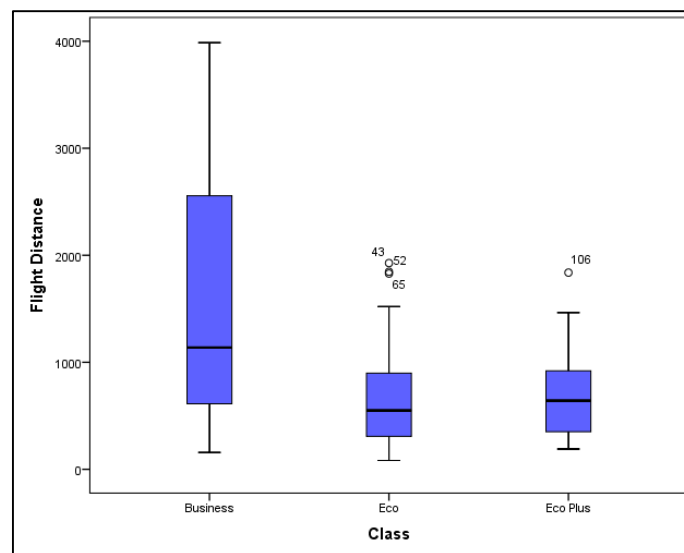
Kepuasan merupakan data kategorik yang terdiri dari dua kategori yaitu puas dan netral atau tidak puas, sedangkan pada kelas penerbangan merupakan data kategorik yang terdiri dari kelas *eco*, *eco plus*, dan *business*.

Tabel 2 Karakteristik Kepuasan

Kepuasan	Kelas Penerbangan			Total
	<i>Eco</i>	<i>Eco Plus</i>	<i>Business</i>	
Puas	25	13	10	48
Netral atau tidak puas	9	40	26	75
Total	34	53	36	123

Tabel 2 menunjukkan bahwa terdapat 40 orang dengan kepuasan netral atau tidak puas cenderung memilih kelas penerbangan *eco plus*, terdapat 10 orang dengan kepuasan puas cenderung memilih kelas penerbangan *business*. Kepuasan dari 123 orang menunjukkan bahwa terdapat 75 orang netral atau tidak puas.

Karakteristik jarak penerbangan berupa data numerik sedangkan pada sedangkan pada kelas penerbangan merupakan data kategorik yang terdiri dari kelas *eco*, *eco plus*, dan *business*.



Gambar 1 Boxplot Jarak Penerbangan Berdasarkan Kelas Penerbangan

Gambar 1 menunjukkan bahwa kelas penerbangan *business* cenderung memiliki rata-rata jarak penerbangan yang lebih jauh dibandingkan kelas *eco* dan *eco plus*. Kelas penerbangan *eco* dan *eco plus* cenderung memiliki rata-rata jarak penerbangan yang dekat dan varians yang kecil ditunjukkan pada panjang *boxplot* yang tidak terlalu panjang.

Uji independensi dilakukan untuk mengetahui ada atau tidak hubungan antara pemilihan kelas penerbangan dengan tipe perjalanan.

Tabel 3 Uji Independensi Pemilihan Kelas Penerbangan dengan Tipe perjalanan

G^2	$\chi^2_{(0,05;2)}$	$p\text{-value}$
23,088	5,991	0,0

Tabel 3 menunjukkan bahwa hasil uji independensi antara pemilihan kelas penerbangan dengan tipe perjalanan diperoleh nilai statistik uji *Likelihood Ratio* sebesar 23,088 lebih besar dari nilai $\chi^2_{(0,05;2)}$ sebesar 5,991 dan diperkuat dengan nilai $p\text{-value}$ sebesar 0 kurang dari 5% sehingga diputuskan tolak H_0 yang artinya terdapat hubungan antara pemilihan kelas penerbangan dengan tipe perjalanan.

Uji independensi dilakukan untuk mengetahui ada atau tidak hubungan antara pemilihan kelas penerbangan dengan kepuasan.

Tabel 4 Uji Independensi Kelas Penerbangan dengan Kepuasan

G^2	$\chi^2_{(0,05;2)}$	$p\text{-value}$
23,468	5,991	0,0

Berdasarkan Tabel 4 menunjukkan bahwa hasil uji independensi antara pemilihan kelas penerbangan dengan kepuasan diperoleh nilai statistik uji *Likelihood Ratio* sebesar 23,468 lebih besar dari nilai $\chi^2_{(0,05;2)}$ sebesar 5,991 dan diperkuat dengan nilai $p\text{-value}$ sebesar 0,0 kurang dari 5% sehingga diputuskan tolak H_0 yang artinya terdapat hubungan antara pemilihan kelas penerbangan dengan kepuasan.

Estimasi parameter merupakan langkah awal dalam analisis regresi logistik ordinal. Adapun model logit yang terbentuk dari variabel prediktor adalah sebagai berikut.

$$\hat{g}_1(x) = 2,288 + 0,909 X_{1(0)} - 0,634 X_{2(0)} + 0,001 X_3$$

$$\hat{g}_2(x) = 0,008 + 0,909 X_{1(0)} - 0,634 X_{2(0)} + 0,001 X_3$$

Berdasarkan model logit tersebut, maka dapat diperoleh model probit untuk jarak penerbangan sebesar 599 km masing-masing kategori respon sebagai berikut.

$$\pi_{1(x)} = \frac{\exp(g_1(x))}{1 + \exp(g_1(x))}$$

$$\pi_{1(x)} = \frac{23.62}{1 + 23.62}$$

$$\pi_{1(x)} = 0,21$$

Model probit untuk $\hat{g}_1(x)$ menunjukkan bahwa peluang seseorang dengan tipe perjalanan bisnis, kepuasan netral atau tidak puas dengan jarak penerbangan sebesar 599 memiliki peluang memilih kelas penerbangan *business* sebesar 0,21.

$$\pi_{2(x)} = \frac{\exp(g_1(x))}{1 + \exp(g_1(x))} - \frac{\exp(g_2(x))}{1 + \exp(g_2(x))}$$

$$\pi_{2(x)} = \frac{23.62}{1 + 23.62} - \frac{2.42}{1 + 2.42}$$

$$\pi_{2(x)} = 0,51$$

Model probit untuk $\hat{g}_2(x)$ menunjukkan bahwa peluang seseorang dengan tipe perjalanan bisnis, kepuasan netral atau tidak puas dengan jarak penerbangan sebesar 599 memiliki peluang memilih kelas penerbangan *eco* sebesar 0,51.

$$\pi_{0(x)} = 1 - \pi_{1(x)} - \pi_{2(x)}$$

$$\pi_{0(x)} = 1 - 0,21 - 0,51$$

$$\pi_{0(x)} = 0,28$$

Model probit untuk $\hat{g}_0(x)$ menunjukkan bahwa peluang seseorang dengan tipe perjalanan bisnis, kepuasan netral atau tidak puas dengan jarak penerbangan sebesar 599 memiliki peluang memilih kelas penerbangan *eco plus* sebesar 0,28. Berdasarkan nilai yang diperoleh menunjukkan bahwa seseorang dengan tipe perjalanan bisnis, kepuasan netral atau tidak puas dengan jarak penerbangan sebesar 599 km memiliki kemungkinan untuk lebih memilih kelas penerbangan *eco*.

Uji serentak dilakukan untuk mengetahui pengaruh variabel prediktor secara serentak atau bersama-sama dengan memasukkan semua variabel yang diduga berpengaruh.

Tabel 5 Uji Serentak

χ^2	$\chi^2_{(0,05;3)}$	$p\text{-value}$
37,283	7,815	0,0

Tabel 5 menunjukkan bahwa hasil uji serentak diperoleh nilai statistik uji sebesar 37,283 lebih besar dari nilai $\chi^2_{(0,05;3)}$ sebesar 7,815 dan diperkuat dengan nilai *p-value* sebesar 0,0 kurang dari 5% sehingga diputuskan tolak H_0 yang artinya minimal terdapat satu variabel prediktor yang berpengaruh signifikan terhadap pemilihan kelas penerbangan.

Uji parsial dilakukan untuk mengetahui signifikansi variabel tipe perjalanan, kepuasan, dan jarak penerbangan terhadap pemilihan kelas penerbangan.

Tabel 6 Uji Parsial

Variabel	Wald	$\chi^2_{(0,05;1)}$	<i>P-value</i>	Keputusan
Konstanta (<i>Business</i>)	13,762	3,841	0,0	Tolak H_0
Konstanta (<i>Eco</i>)	0,0		0,989	Gagal Tolak H_0
Jarak Penerbangan	13,03		0,0	Tolak H_0
Tipe perjalanan (Bisnis)	4,642		0,031	Tolak H_0
Kepuasan (Netral atau tidak puas)	2,178		0,14	Gagal Tolak H_0

Tabel 6 menunjukkan bahwa nilai statistik uji pada konstanta model pemilihan kelas penerbangan *business*, jarak penerbangan, dan tipe perjalanan kategori bisnis diperoleh nilai *Wald* berturut-turut sebesar 13,762, 13,03, 4,642 dan nilai *p-value* kurang dari α sebesar 5% sehingga diputuskan tolak H_0 yang artinya konstanta model pemilihan kelas penerbangan *business*, jarak penerbangan, dan tipe perjalanan kategori bisnis berpengaruh signifikan terhadap pemilihan kelas penerbangan.

Uji kecocokan model dilakukan untuk menguji cocok atau tidak data dilakukan analisis menggunakan regresi logistik ordinal.

Tabel 7 Uji Kecocokan Model

χ^2	$\chi^2_{(0,05;3)}$	<i>P-value</i>
3,391	7,815	0,335

Tabel 7 menunjukkan bahwa uji kecocokan model dari tipe perjalanan, kepuasan, dan jarak penerbangan menghasilkan χ^2 sebesar 3,391 yang lebih kecil dari $\chi^2_{(0,05;3)}$ sebesar 7,815 serta diperkuat dengan nilai *P-Value* sebesar 0,335 yang lebih besar dari nilai α sebesar 0,05 sehingga diputuskan gagal tolak H_0 , artinya *slope* untuk masing-masing kategori variabel respon adalah sama. Serta, model ini lebih sesuai menggunakan regresi logistik ordinal dari pada regresi logistik multinomial.

Uji kesesuaian model pada tipe perjalanan, kepuasan dan jarak penerbangan terhadap kelas penerbangan adalah sebagai berikut.

Tabel 8 Uji Kesesuaian Model

χ^2	$\chi^2_{(0,05;3)}$	<i>P-value</i>
37,283	7,815	0,00

Tabel 8 menunjukkan bahwa uji kesesuaian model dari tipe perjalanan, kepuasan, dan jarak penerbangan menghasilkan χ^2 sebesar 37,283 yang lebih besar dari $\chi^2_{(0,05;3)}$ sebesar 7,815 serta diperkuat dengan nilai *p-value* sebesar 0,00 yang lebih besar dari nilai α sebesar 0,05 sehingga diputuskan gagal tolak H_0 , artinya tidak ada perbedaan antara observasi dan prediksi (model sesuai).

Data pengaruh tipe perjalanan, kepuasan, dan jarak penerbangan terhadap pemilihan kelas penerbangan, diperoleh nilai *R-square* sebesar 0,296. Artinya, model yang didapatkan bisa dikatakan sebagai model yang tidak baik karena hanya 29,6% varians dalam variabel respon berupa pemilihan kelas penerbangan dapat diprediksi dari variabel prediktor berupa

tipe perjalanan, kepuasan, dan jarak penerbangan serta sisanya sebesar 70,4% merupakan error yang dijelaskan oleh variabel lain diluar model.

Hasil analisis nilai *odds ratio* digunakan untuk mengukur risiko atau kecenderungan antara tipe perjalanan, kepuasan, dan jarak penerbangan adalah sebagai berikut.

Tabel 9 Odds Ratio

Variabel	Exp(β)
Tipe perjalanan (Bisnis)	2,482
Kepuasan (Netral atau Tidak Puas)	0,530
Jarak Penerbangan	1,001

Tabel 9 menunjukkan bahwa peluang seseorang melakukan pemilihan kelas penerbangan dengan tipe perjalanan bisnis memiliki kecenderungan 2,482 kali daripada tipe perjalanan pribadi, peluang seseorang melakukan pemilihan kelas penerbangan dengan status kepuasan netral atau tidak puas memiliki kecenderungan 0,530 kali dari pada status kepuasan puas. Serta jika jarak penerbangan bertambah satu satu, maka 1,001 kali lebih baik daripada tidak bertambah satu satuan.

Analisis ketepatan klasifikasi berguna untuk mengetahui proporsi kasus yang tepat diklasifikasikan melalui model regresi logistik ordinal.

Tabel 10 Ketepatan Klasifikasi

Kelas Penerbangan	Kelas Penerbangan			Total
	<i>Business</i>	<i>Eco</i>	<i>Eco Plus</i>	
<i>Business</i>	17	3	0	16.3%
<i>Eco</i>	17	36	28	65.9%
<i>Eco Plus</i>	0	14	8	17.9%
Total	27.6%	43.1%	29.3%	49,6%

Tabel 10 menunjukkan bahwa seseorang yang memilih kelas penerbangan *Business* memiliki persentase kebenaran sebesar 16,3% dengan rincian 17 orang diklasifikasikan dengan tepat dan 3 orang tidak tepat, seseorang yang memilih kelas penerbangan *Eco* memiliki persentase kebenaran sebesar 65,9% dengan rincian 36 orang diklasifikasikan dengan tepat dan 45 orang tidak tepat, serta seseorang yang memilih kelas penerbangan *Eco Plus* memiliki persentase kebenaran sebesar 17,9% dengan rincian 8 orang diklasifikasikan dengan tepat dan 14 orang tidak tepat. Sehingga diperoleh ketepatan klasifikasi sebesar 49,6%, sehingga nilai APER (*Apparent Error Rate*) adalah sebesar 50,4% yang dihitung dari 100% dikurangi dengan persentase ketepatan klasifikasi.

D. Kesimpulan

Karakteristik kelas penerbangan berdasarkan tipe perjalanan dan kepuasan yang paling banyak dipilih adalah kelas *eco plus*. Karakteristik kelas penerbangan berdasarkan jarak penerbangan pada jarak jauh cenderung memilih kelas *business*. Hasil uji independensi menunjukkan bahwa terdapat hubungan antara kelas penerbangana dengan tipe perjalanan dan kepuasan. Hasil analisis regresi logistik ordinal menunjukkan bahwa variabel jenis tipe perjalanan dan jarak penerbangan berpengaruh signifikan terhadap kelas penerbangan. Data cocok dilakukan analisis menggunakan regresi logistik ordinal. Model sesuai antara prediksi dan observasi, namun kebaikan model belum termasuk model yang baik. Ketepatan klasifikasi yang diperoleh masih rendah.