

Pemodelan regresi *logistik firth penalized maximum likelihood estimation* pada kepemilikan tempat usaha

¹Siti Maslihah, ²Eva Khoirun Nisa

^{1,2}Fakultas Sains dan Teknologi, UIN Walisongo Semarang
Email korespondensi: evakn@walisongo.ac.id

Abstrak

Pada penelitian sosial sering memperoleh data kategorik dengan proporsi kategori yang diinginkan (sukses) jauh lebih sedikit dibandingkan dengan proporsi kategori yang tidak diinginkan (gagal). Data tersebut dikatakan sebagai separasi sempurna atau rare events. Data dengan rare events bila digunakan dalam regresi termasuk regresi logistik akan menghasilkan estimator parameter yang bias dalam proses Maximum Likelihood Estimation (MLE). Untuk mengatasi kebiasaan karena data rare events atau separasi sempurna maka Firth menambahkan unsur penalti matriks informasi Fisher fungsi \ln likelihood MLE. Metode tersebut dinamakan Firth Penalized Maximum Likelihood Estimation. Penelitian ini bertujuan untuk memodelkan regresi logistik dengan Firth Penalized Maximum Likelihood Estimation pada data kepemilikan tempat usaha. Model regresi logistik yang diperoleh sudah sesuai dengan data dan menghasilkan bahwa bidang usaha, umur usaha, skala UMKM, dan lokasi usaha berpengaruh terhadap kepemilikan tempat usaha. Dengan demikian, para pelaku UMKM perlu mempertimbangkan dengan bijak kepemilikan tempat usaha.

Kata kunci: Regresi Logistik; Firth Penalized Maximum Likelihood Estimation; Maximum Likelihood Estimation; Kepemilikan Tempat Usaha

Abstract

In social research, categorical data is often obtained with the proportion of desirable categories (success) smaller than that of undesirable categories (failure). This data is called the perfect separation or rare events data. Rare events in regression including logistic regression, will produce biased estimators, especially the Maximum Likelihood Estimation (MLE) estimators. To overcome the bias problem of rare events, Firth added a penalty element in MLE. The penalty is the Fisher information matrix. This method is called Firth Penalized Maximum Likelihood Estimation. This research aims to logistic regression modeling with Firth Penalized Maximum Likelihood Estimation on business place ownership data. The logistic regression model obtained is under the data and produces that business field, business age, MSME scale, and business location influence business premises ownership. Therefore, MSME need to wisely consider the ownership of business places.

Keywords: *Logistic Regression; Firth Penalized Maximum Likelihood Estimation; Maximum Likelihood Estimation; Business Place Ownership*

A. Pendahuluan

Dalam riset sosial seperti bidang ekonomi, politik, kedokteran sering kali bertemu dengan data kategorik (Cameron and Trivedi, 1998) yang tentunya tidak dapat diterapkan dengan metode statistika parametrik. Regresi logistik menjadi satu dari beberapa metode statistika yang dapat menjelaskan pengaruh beberapa variabel independen terhadap variabel dependen yang berbentuk kategorik (Agresti, 2002). Dapat juga dikatakan sebagai model yang selain dapat menyatakan hubungan antar variabel prediktor dengan variabel respon kategorik dikotomi juga menghasilkan peluang tiap kategorinya (Nisa and Maslihah, 2024)

Menurut Klieštik et al. (2015), regresi logistik fungsinya lebih sederhana dibandingkan regresi probit sehingga mudah diterapkan pada data kategorik. Selain karena bentuk sederhana, model ini juga terkenal dengan keistimewaannya yaitu hasilnya signifikan, kuat (Antipov, 2019) serta mempunyai akurasi tinggi (Jain, Khunteta, and Srivastava, 2020) Menurut Nisa and Muanalifah (2021) mengatakan bahwa keuntungan praktis yang teridentifikasi dari model logit dibandingkan model probit adalah kesederhanaannya bentuk struktural dan interpretasi hasil. Namun, di lapangan banyak terjadi kategori yang diharapkan justru persentase atau proporsi kategorinya jauh lebih sedikit daripada proporsi kategori yang tidak diharapkan (Firth, 1993).

Ketika data terjadi separasi sempurna sehingga memunculkan *rare events* menyebabkan pada pendugaan parameter menghasilkan estimator yang bias (Suhas et al., 2023) juga terjadi *over fitting* model (Malau and Joseph, 2023). Firth (1993) memberikan inovasi pada proses pendugaan parameter melalui *Maximum Likelihood Estimation* (MLE) yaitu dengan menambahkan unsur penalti yang diperoleh dari determinan matriks informasi Fisher. MLE menjadi metode pendugaan parameter yang paling banyak digunakan (Nisa and Miasary, 2024) maka dengan penambahan unsur penalti ini akan menghasilkan estimator tak bias pada regresi termasuk regresi logistik.

Hutagalung and Sutarman (2024) telah membuktikan bahwa *Penalized Maximum Likelihood* lebih tinggi keakuratannya dibandingkan MLE. Terlebih pada data ukuran besar akurasi *Penalized Maximum Likelihood* signifikan meningkat. Suhas et al. (2023) menambahkan kelebihan dari *Penalized Maximum Likelihood* yaitu metode ini masih bisa mengakomodasi sampel kecil. Oleh karena itu, metode ini menjadi pilihan yang tepat ada permasalahan data rare events maupun sampel kecil yang menyebabkan

estimator dari MLE bias. Data *rare events* biasa terjadi pada data sosial ekonomi seperti kepemilikan tempat usaha.

Tempat usaha sangat penting dan mempengaruhi keberhasilan usaha (Astriyani, 2024). Pernyataan ini sesuai dengan kondisi UMKM yang ada di Kelurahan Ngaliyan, Kota Semarang. Hasil penelitian sebelumnya yang dilakukan peneliti pasca pandemi Covid-19 bahwa mayoritas UMKM yang ada di sana terutama yang berskala mikro mengalami banyak kerugian dan ketahanan usaha yang buruk (Nisa and Miasary, 2024). Dampaknya banyak pelaku usaha menutup usahanya. Namun, pada tahun 2024 ini melalui pra riset berupa wawancara pada 10 pemilik usaha mikro kondisi keuangan sudah membaik. Satu faktor yang ternyata mereka pilih dan dianggap mempengaruhi keberhasilan usaha adalah kepemilikan tempat usaha. Para responden menyatakan mereka lebih memilih sewa atau kontrak tempat usaha daripada harus membeli. Alasan memilih opsi tersebut karena dianggap lebih ringan dan fleksibel bila suatu saat terjadi peristiwa yang menyebabkan pemasukan sedikit. mempunyai tempat usaha sendiri melainkan sewa atau kontrak. Mereka juga menyatakan kondisi keuangan dengan tempat usaha menyewa tidak lebih buruk dari milik sendiri. Namun juga berpengaruh ketika usaha telah berumur puluhan tahun (Risnawati, 2018). Berdasarkan hasil pra riset yang dilakukan mayoritas UMKM di Ngaliyan menyewa tempat usaha daripada mengusahakan milik sendiri maka sesuai bila proses pemodelan regresi logistik nanti dibantu dengan *Penalized Maximum Likelihood*. Dengan demikian, penelitian ini bertujuan untuk mendeskripsikan pengaruh lokasi usaha, bidang usaha, umur usaha, dan skala UMKM terhadap pemilihan kepemilikan tempat usaha dengan regresi logistik.

B. Metode Penelitian

Penelitian ini menggunakan analisis secara deskriptif dan inferensial. Analisis secara deskriptif menggunakan tabel sedangkan metode regresi logistik *Firth Penalized Maximum Likelihood* digunakan untuk menganalisis secara inferensial.

Regresi logistik menjelaskan hubungan antara suatu variabel respon dengan satu atau lebih variabel prediktor dengan variabel responnya bersifat dikotomis (Hosmer and Lemeshow, 2000). Variabel respon atau terikat dinotasikan dengan Y dan variabel prediktor atau bebas dinotasikan dengan X . Dikotomis artinya variabel Y mempunyai dua kategori yaitu kategori 1 untuk kejadian sukses dan 0 untuk kejadian gagal sehingga variabel Y mengikuti distribusi Bernoulli (Agresti, 2002). Berbeda dengan model regresi linier, Regresi Logistik merupakan model regresi yang menggambarkan probabilitas atau resiko dari suatu objek. Dari model ini akan dapat dicari informasi peluang kejadian sukses dan gagal.

Bentuk umum model peluang regresi logistik dengan p banyaknya variabel prediktor, menurut Hosmer and Lemeshow (2000) diformulasikan dalam persamaan (1) berikut ini.

$$\pi(x) = \frac{e^{(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}}{1 + e^{(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}} \quad (1)$$

Fungsi $\pi(x)$ adalah peluang kejadian sukses dengan nilai probabilitas $0 \leq \pi(x) \leq 1$ sedangkan β_j adalah nilai parameter dengan $j = 1, 2, \dots, p$. $\pi(x)$ merupakan fungsi yang non linier, sehingga perlu dilakukan transformasi ke dalam bentuk logit untuk memperoleh fungsi yang linier. Persamaan (2) berikut merupakan transformasi dari logit $\pi(x)$.

$$g(x) = \ln\left(\frac{\pi(x)}{1 - \pi(x)}\right) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p \quad (2)$$

Untuk metode pendugaan parameter model regresi logistik dengan *Maximum Likelihood Estimation* tetapi karena data kepemilikan tempat usaha yang digunakan terjadi *rare events* maka dimodifikasi dengan *Firth Penalized Maximum Likelihood Estimation*. Data *rare events* merupakan data dengan proporsi sukses sangat kecil dibandingkan dengan proporsi gagal dilakukan (King et al., 2001). Adanya *rare events* rering kali menghasilkan estimator bias (Putri and Budyanra, 2021) dan kurang *powerful* untuk sampel kecil (Rahman and Sultana, 2017).

Firth Penalized Maximum Likelihood merupakan metode estimasi parameter dengan menambahkan unsur penalti dari matriks informasi Fisher $I(\beta)$ yang berasal dari *Jeffreys prior* (Firth, 1993). Persamaan (3) berikut ini merupakan unsur penalti yang sudah dilogaritma naturalkan (ln).

$$\text{Penalti} = \frac{1}{2} \ln |I(\beta)| \quad (3)$$

Dengan memasukkan unsur penalti ke dalam fungsi likelihood regresi logistik maka diperoleh fungsi *Firth Penalized Likelihood* seperti persamaan (4) ini.

$$\setminus L^*(\beta) = L(\beta) |I(\beta)|^{\frac{1}{2}} \quad (4)$$

dengan

$L^*(\beta)$: fungsi *Firth penalized likelihood*

$L(\beta)$: fungsi *likelihood*

$I(\beta)$: matriks informasi Fisher

Variabel yang digunakan pada penelitian ini adalah variabel respon berupa kepemilikan tempat usaha dan variabel prediktor berupa bidang usaha, umur usaha, skala UMKM, dan lokasi usaha. Untuk variabel kepemilikan tempat usaha mempunyai kategori 1 yaitu milik sendiri dan 0 yang berarti sewa/kontrak. Selanjutnya semua variabel prediktor mempunyai kategori sebagai berikut:

- a. Variabel bidang usaha dengan kategori 1: barang, 0: jasa
- b. Variabel umur usaha dengan kategori 1: < 5 tahun, 0: ≥ 5 tahun
- c. Variabel skala UMKM dengan kategori 1: usaha mikro, 2: usaha kecil, 3: usaha menengah
- d. Variabel lokasi usaha dengan kategori 1: pertokoan, 0: bukan pertokoan

Sementara untuk data penelitian ini merupakan data primer yang diperoleh langsung dari responden yang telah mengisi angket sebanyak 35 UMKM di Kelurahan Ngaliyan. Pengambilan sampel dilakukan dengan *purposive sampling* karena dengan teknik pengambilan sampel ini akan memperoleh responden atau informan yang berkualitas (Tongco, 2007). Instrumen penelitian dalam hal ini adalah angket telah melalui proses validasi yang meliputi pengujian validitas dan reliabilitas. Uji validitas dengan korelasi Pearson Product Momen (r) (Walpole et al, 2012) yang diformulasikan dengan persamaan (5).

$$r = \frac{n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{\sqrt{\left[n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right] \left[n \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i \right)^2 \right]}} \quad (5)$$

Keterangan:

n : banyaknya responden

x : variabel kesatu

y : variabel kedua

Butir pertanyaan dikatakan valid jika $r_{hitung} > r_{\alpha, n-1}$. Selanjutnya untuk pengujian reliabilitas menggunakan Alpha Cronbach (Sudijono, 2015) dengan persamaan (6).

$$r_{11} = \left(\frac{n}{n-1} \right) \left(1 - \frac{\sum_i S_i^2}{S_t^2} \right) \quad (6)$$

Keterangan:

n : banyaknya butir pertanyaan

S_i^2 : varians tiap butir pertanyaan

S_t^2 : varians total

Instrumen pertanyaan reliabel jika $r_{11} > 0,7$ (Sudijono, 2015). Berikut hasil validitas dan reliabilitas instrumen yang digunakan.

Tabel 1 Pengujian Validitas dan Reliabilitas

Pertanyaan nomor	r_{hitung}	$r_{0,05;34}$	Keputusan Validitas	Alpha Cronbach	Keputusan Reliabilitas
1	0,647	0,334	Valid		
2	0,508	0,334	Valid		
3	0,462	0,334	Valid	0,728	Reliabel
4	0,738	0,334	Valid		
5	0,419	0,334	Valid		

Berdasarkan Tabel 1 diperoleh semua pertanyaan valid karena $r_{hitung} > r_{0,05;34}$ dan instrumen angket yang digunakan reliabel karena $0,728 > 0,7$.

Untuk tahapan analisis dalam penelitian ini diberikan sebagai berikut:

1. Pendeskripsian data

Data dideskripsikan dengan proporsi dari masing-masing kategori kepemilikan tempat usaha sebagai variabel respon.

$$p = \frac{x}{N}$$

dengan x : jumlah data pada kategori yang dituju

N : jumlah seluruh data

2. Pendugaan parameter regresi logistik dengan *Firth Penalized Maximum Likelihood*.

Pendugaan parameter regresi logistik dengan *Firth Penalized Maximum Likelihood* menggunakan metode Maximum Likelihood Estimation modifikasi penalti *Firth Penalized* yang telah dijelaskan sebelumnya.

3. Pemodelan regresi logistik pada data kepemilikan tempat usaha UMKM. Data diaplikasikan pada model regresi logistik *Firth Penalized Maximum Likelihood* berbantu *software* RStudio.

4. Pengujian parameter secara serentak dan parsial.

Pengujian parameter secara serentak menggunakan *Maximum Likelihood Ratio Test* (MLRT) dengan hipotesis sebagai berikut:

Hipotesis:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1 : \text{minimal ada satu } \beta_j \neq 0 ; j=1,2,\dots,k$$

Dengan k adalah banyaknya variabel prediktor.

$$\text{Statistik uji: } G^2 = -2 \ln \left(\frac{L(\hat{\omega})}{L(\hat{\Omega})} \right)$$

dengan

$L(\hat{\omega})$: fungsi likelihood di bawah H_0

$L(\hat{\Omega})$: fungsi likelihood di bawah H_1

Menurut Wilks (1938) $-2 \ln \Lambda$ merupakan *likelihood ratio* untuk n mendekati tak hingga yang asimtotik berdistribusi Chi-square dengan derajat bebas selisih jumlah parameter di bawah H_1 dan H_0 .

Kriteria penolakan:

Tolak H_0 jika $G^2_{\text{hitung}} > \chi^2_{k,\alpha}$ H_0 (Bourakadi, Mentagui, and Chakir, 2022).

Uji parameter secara parsial dilakukan untuk mengetahui pengaruh atau tidaknya parameter secara parsial.

Hipotesis:

$$H_0 : \beta_i = 0$$

$$H_1 : \text{terdapat salah satu } \beta_i \neq 0$$

Statistik Uji:

$$Z_{\text{hit}} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \text{ dengan } SE(\hat{\beta}_j) = \sqrt{\hat{\text{var}}(\hat{\beta}_j)}$$

Z berdistribusi hampiran (limit) normal standar $Z \underset{n \rightarrow \infty}{\square} N(0,1)$ dengan n besar mendekati tak hingga.

5. Pengujian kecocokan model

Model yang telah diperoleh diuji kecocokannya dengan data dengan uji Hosmer dan Lemeshow. Dinotasikan dengan \hat{C} yang dihitung berdasarkan pendugaan probabilitas (Hosmer and Lemeshow, 2000). Statistik Uji \hat{C} diformulasikan:

$$\hat{C} = \sum_{k=1}^g \frac{(o_k - n_k \pi_k)^2}{n_k \pi_k (1 - \pi_k)}$$

dimana:

o_k = jumlah respon kovariat c_k dengan formula:

$$o_k = \sum_{j=1}^{c_k} y_j$$

c_k = jumlah pola kovariat dalam desil k.

n_k' = jumlah total subyek dalam kelompok k ,

$\bar{\pi}_k$ = rata-rata estimasi probabilitas dengan formula:

$$\bar{\pi}_k = \sum_{j=1}^{c_k} \frac{m_j \hat{\pi}_j}{n_k'}$$

Dengan hipotesis H_0 : model yang dihipotesiskan fit dengan data dan H_1 : model yang dihipotesiskan tidak fit dengan data maka tolak H_0 jika $\hat{C} > \chi^2_{\alpha, g-2}$ dimana g adalah jumlah kelompok (Liu, Zhang, Hall, Liu, and Zhang, 2017).

6. Interpretasi model

Odds ratio digunakan untuk menginterpretasikan model regresi logistik yang diperoleh sehingga diketahui besar peluang masing-masing kategori kepemilikan tempat usaha. *Odds ratio* menyatakan risiko atau kecenderungan pengaruh observasi dengan $x_j = 1$ adalah berapa kali lipat jika dibandingkan dengan observasi dengan $x_j = 0$ (Agresti, 2002;

Holmberg and Andersen, 2020) dengan formula $\theta = \frac{\pi(1)/(1-\pi(1))}{\pi(0)/(1-\pi(0))}$.

C. Hasil dan Pembahasan

1. Deskripsi Kepemilikan Tempat Usaha

Kelurahan Ngaliyan termasuk salah satu kelurahan dengan perputaran uang tercepat karena banyaknya pelaku ekonomi atau lebih khusus disebut dengan UMKM yang berada di kelurahan tersebut. Berikut hasil proporsi kepemilikan tempat usaha berdasarkan variabel bidang usaha, umur usaha, skala UMKM, dan lokasi usaha.

Tabel 2 Proporsi Kepemilikan Tempat Usaha

Kategori	Kepemilikan tempat usaha		
	Milik Sendiri	Sewa/Kontrak	
Bidang usaha	Barang	0,142857143	0,857142857
	Jasa	0,085714286	0,914285714
Umur usaha	< 5 tahun	0,057142857	0,942857143
	>= 5 tahun	0,171428571	0,828571429
Skala UMKM	Mikro	0,114285714	0,885714286
	Kecil	0,114285714	0,885714286
	Menengah	0	1

	Kategori	Kepemilikan tempat usaha	
		Milik Sendiri	Sewa/Kontrak
Lokasi usaha	Pertokoan	0,057142857	0,942857143
	Bukan pertokoan	0,171428571	0,828571429

Berdasarkan Tabel 2 proporsi milik sendiri pada status kepemilikan tempat usaha sangat kecil dibandingkan dengan status kepemilikan tempat usaha sewa/kontrak. Hal ini terjadi pada semua variabel prediktor yaitu bidang usaha, umur usaha, jenis UMKM, dan lokasi usaha. Dari hasil wawancara lebih lanjut pada beberapa pemilik usaha menyatakan bahwa mereka tidak mampu bila harus membeli tempat usaha terlebih di kawasan pertokoan. Tentunya harganya sangat mahal. Harga sewa kios atau toko di Kelurahan Ngaliyan Kota Semarang juga lebih mahal dibandingkan dengan kelurahan lain karena kelurahan tersebut terdapat dua perguruan tinggi. Namun, bila dibandingkan dengan menyewa keuntungan yang diperoleh juga hampir sama.

2. *Firth Penalized Maximum Likelihood Estimation* pada Regresi Logistik

Berdasarkan persamaan (4) maka akan diperoleh fungsi ln likelihood *Firth Penalized Likelihood* sebagai berikut.

$$L^*(\beta) = \sum_{i=1}^n [y_i \ln(\pi(x_i)) + (1 - y_i) \ln(1 - \pi(x_i))] + \frac{1}{2} \ln |I(\beta)| \quad (7)$$

Selanjutnya memaksimalkan fungsi ln likelihood dilakukan dengan mencari turunan persamaan (7) terhadap β kemudian disamadengankan nol.

$$\frac{\partial \ln L(\beta)}{\partial \beta} = 0 \quad (8)$$

3. Regresi Logistik *Firth Penalized Maximum Likelihood* pada Kepemilikan Tempat Usaha

Pemodelan regresi logistik dengan metode estimasi *Firth Penalized Maximum Likelihood* pada kepemilikan tempat usaha diawali dengan estimasi parameter. Berikut hasil estimasi parameter yang diperoleh

Tabel 3 Estimator Parameter Regresi Logistik

β_0	-4,453
$\beta_{1(1)}$	0,361
$\beta_{2(1)}$	0,271
$\beta_{3(1)}$	1,822
$\beta_{3(2)}$	1,822
$\beta_{4(1)}$	2,733

Berdasarkan informasi estimator pada Tabel 3 maka diperoleh model regresi logistik sementara yaitu

$$g(x) = -4,453 + 0,361x_{1(1)} + 0,271x_{2(2)} + 1,822x_{3(1)} + 0,613x_{3(2)} + 2,733x_{4(1)}$$

Selanjutnya kita uji parameter secara serentak dan parsial. Dengan uji statistik *Maximum Likelihood Ratio Test* (MLRT) yang dinotasikan dengan G^2 diperoleh nilai $G^2 > \chi_{0,05;3}^2$ ($31,248 > 7,815$) maka H_0 ditolak sehingga semua parameter secara serentak signifikan dalam model. Dilanjutkan dengan uji parameter secara parsial.

Tabel 4 Uji Parameter secara Parsial

β	$ Z_{hitung} $	$Z_{0,025}$
-4,453	2,17	1,96
0,361	3,72	1,96
0,271	2,93	1,96
1,822	1,99	1,96
1,822	5,13	1,96
2,733	4,95	1,96

Berdasarkan Tabel 4 dapat dinyatakan semua parameter β signifikan karena $|Z| > Z_{0,025}$. Dengan demikian, model regresi logistik terbaik untuk data kepemilikan tempat usaha sebagai berikut:

$$\hat{g}(x) = -4,453 + 0,361x_{1(1)} + 0,271x_{2(2)} + 1,822x_{3(1)} + 0,613x_{3(2)} + 2,733x_{4(1)}$$

Dari model regresi logistic tersebut maka dapat dikatakan bidang usaha, umur usaha, skala UMKM, dan lokasi usaha berpengaruh terhadap pemilihan dalam kepemilikan tempat usaha.

Model yang diperoleh diuji kesesuaiannya agar diperoleh model regresi logistik terbaik untuk data kepemilikan tempat usaha ini. Dari hasil perhitungan diperoleh statistik $\hat{C} < \chi^2_{0,05;2}$ ($3,25 < 5,99$) maka H_0 gagal ditolak sehingga model regresi logistik *Firth Penalized Maximum Likelihood* yang terbentuk sesuai dengan model. Selanjutnya dari model regresi logistik maka dapat dibentuk fungsi peluang untuk kepemilikan tempat usaha sebagai berikut:

$$\pi(x) = \frac{e^{-4,453+0,361x_{1(1)}+0,271x_{2(2)}+1,822x_{3(1)}+0,613x_{3(2)}+2,733x_{4(1)}}}{1 + e^{-4,453+0,361x_{1(1)}+0,271x_{2(2)}+1,822x_{3(1)}+0,613x_{3(2)}+2,733x_{4(1)}}$$

Dari fungsi peluang tersebut dapat dicari nilai odds ratio untuk melihat kecenderungan dari masing-masing variabel yang mempengaruhi kepemilikan tempat usaha.

Tabel 5 *Odds Ratio*

Variabel prediktor	<i>Odds ratio</i>
Bidang usaha	1,47
Umur usaha	0,64
Skala UMKM (1)	1,73
Skala UMKM (2)	0,71
Lokasi usaha	0,45

Dari Tabel 5 dapat diinterpretasikan bahwa

- Kecenderungan UMKM bidang usaha barang mempunyai tempat usaha milik sendiri sebesar 1,47 kali dibandingkan dengan UMKM bidang usaha jasa. Informasi dari beberapa responden menyatakan usaha jasa seperti agen ekspedisi pengiriman barang atau laundry tidak membutuhkan tempat yang khusus seperti di pertokoan ataupun membutuhkan temoat yang luas sehingga mayoritas UMKM terlebih yang berskala mikro lebih memilih menyewa tempat usaha.
- Kecenderungan UMKM dengan umur usaha kurang dari lima tahun mempunyai tempat usaha milik sendiri sebesar 0,64 kali lebih rendah dibandingkan dengan UMKM yang telah berumur lebih dari lima tahun usahanya. Tentunya bila usaha semakin lama maka pastinya pemilk usaha akan berpikir untuk mempunyai tempat usaha yang permanen. Terlebih dari enam responden menyatakan bahwa pemilik kios atau tempat usaha yang disewa kadang-kadang menaikkan harga sewa. Terlebih bila usaha yang kiosnya disewa laris, pasti akan dinaikkan harga sewa kedepannya.
- Kecenderungan UMKM skala mikro mempunyai tempat usaha milik sendiri sebesar 1,73 kali dibandingkan dengan UMKM skala kecil dan menengah. Hal ini sesuai dengan data dari BPS Kecamatan Ngaliyan

- [5] bahwa hampir 90% usaha mikro yang ada di Ngaliyan tidak mempunyai tempat usaha milik sendiri.
- d. Kecenderungan UMKM skala kecil mempunyai tempat usaha milik sendiri sebesar 0,71 kali dibandingkan dengan UMKM skala mikro dan menengah. Hasil ini dikarenakan UMKM skala kecil dan menengah jauh lebih sedikit daripada skala mikro.
 - e. Kecenderungan UMKM dengan lokasi usaha di pertokoan mempunyai tempat usaha milik sendiri sebesar 0,45 kali lebih rendah dibandingkan dengan UMKM dengan lokasi usaha tidak di pertokoan. Hal ini sesuai dengan pernyataan responden bahwa biasanya pelaku usaha mikro mempunyai tempat milik sendiri bila berjualan di rumah sendiri. Bila ingin mencari tempat yang strategis seperti di pertokoan lebih murah menyewa daripada beli kios di pertokoan.

D. Simpulan

Regresi logistik dengan metode estimasi parameter *Firth Penalized Maximum Likelihood* tepat digunakan untuk data separasi sempurna atau *rare events* seperti pada data kepemilikan tempat usaha ini. Bidang usaha, umur usaha, skala UMKM, dan lokasi usaha mempengaruhi pemilihan kepemilikan usaha milik sendiri atau sewa.

E. Daftar Pustaka

- Agresti, A. (2002). *An introduction to categorical data analysis* (Second Edi). New Jersey: John Wiley & Sons, Inc Publication.
- Antipov, E. (2019). *Applying CHAID for logistic regression diagnostics and classification accuracy improvement*. (January). <https://doi.org/10.2139/ssrn.1412208>
- Astriyani, R. (2024). Pengaruh lokasi, modal dan kemampuan usaha terhadap keberhasilan usaha. *Jurnal Kewirausahaan dan Multi Talenta*, 2. <https://doi.org/10.38035/jkmt.v2i2>
- Bourakadi, A., Mentagui, D., and Chakir, B. A. (2022). Hypothesis test of sample mean of random intervals and comparing between methods based on Hausdorff distance and the maximum likelihood ratio. *Italian Journal of pure and applied mathematics*, 48, 384–398.
- BPS. (2021). *Kecamatan Ngaliyan dalam Angka 2021*. Semarang: Badan Pusat Statistik Kota Semarang.
- Cameron, A.C. and Trivedi, P. K. (1998). *Regression analysis of count data*. Cambridge: Cambridge University Press. <https://doi.org/https://doi.org/10.1017/CBO9780511814365>
- Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika*, 80(1), 27–38. <https://doi.org/https://doi.org/10.2307/2336755>

- Holmberg, M. J., and Andersen, L. W. (2020). Estimating risk ratios and risk differences. *JAMA*, 324(11), 1098. <https://doi.org/10.1001/jama.2020.12698>
- Hosmer, D., and Lemeshow, S. (2000). *Applied logistic regression*. Canada: A John Wiley & Sons, Inc Publication.
- Hutagalung, M. A. I., and Sutarman. (2024). Penalized maximum likelihood estimation dengan algoritma gradient descent pada model regresi logistik multinomial. *IJM: Indonesian Journal of Multidisciplinary*, 2. Retrieved from <https://journal.csspublishing/index.php/ijm>
- Jain, H., Khunteta, A., and Srivastava, S. (2020). Churn prediction in telecommunication using logistic regression and logit Boost. *Procedia Computer Science*, 167, 101–112. <https://doi.org/10.1016/j.procs.2020.03.187>
- King, G., Langche Zeng, G. H. E., Alt, J., Freeman, J., Gleditsch, K., Imbens, G., ... Zeng, L. (2001). Logistic regression in rare events data. In *the Society for Political Methodology*. Retrieved from <http://GKing.Harvard.Edu>.
- Klieštík, T., Kočíšová, K., and Mišanková, M. (2015). Logit and probit model used for prediction of financial health of company. *Procedia Economics and Finance*, 23, 850–855. [https://doi.org/10.1016/S2212-5671\(15\)00485-2](https://doi.org/10.1016/S2212-5671(15)00485-2)
- Liu, H., Zhang, Z., Hall, H., Liu, H., and Zhang, &. (2017). Logistic regression with misclassification in binary outcome variables: A method and software. In *Behaviormetrika* (Vol. 44).
- Malau, T., and Joseph, T. (2023). Analisis metode logistik regresi ensemble untuk klasifikasi dengan pra-pemrosesan menggunakan principal component analysis. In *IJM: Indonesian Journal of Multidisciplinary* (Vol. 1). Retrieved from <https://journal.csspublishing/index.php/ijm>
- Nisa, E. K., and Maslihah, S. (2024). *Regresi logistik dan penerapannya*. PPSM.
- Nisa, E. K., and Miasary, S. D. (2024). *Parameter estimation and application of inverse gaussian regression*. 020008. <https://doi.org/10.1063/5.0194666>
- Nisa, E. K., and Muanalifah, A. (2021). *The comparison results of logit and probit regression on factors of woman criminal*. 5(2), 287–293. <https://doi.org/10.31764/jtam.v5i2.4150>
- Putri, S. S., and Budyanra. (2021). Penerapan metode firth's penalized maximum likelihood estimation pada analisis determinan penyalahgunaan narkoba remaja di Indonesia tahun 2017. *Statistika*, 21(2), 117–124. <https://doi.org/https://doi.org/10.29313/statistika.v21i2.308>
- Rahman, M. S., and Sultana, M. (2017). Performance of firth-and logF-type penalized methods in risk prediction for small or sparse binary data. *BMC Medical Research Methodology*, 17(1), 33–48. <https://doi.org/10.1186/s12874-017-0313-9>

- Risnawati, N. (2018). Profil UMK, permasalahan dan upaya pemberdayaannya. *Coopetition: Jurnal Ilmiah Manajemen*, IX(2).
- Sudijono, A. (2015). *Pengantar Evaluasi Pendidikan*. Jakarta: PT Raja Grafindo Persada.
- Suhas, S., Manjunatha, N., Kumar, C. N., Benegal, V., Rao, G. N., Varghese, M., and Gururaj, G. (2023). Firth's penalized logistic regression: A superior approach for analysis of data from India's National Mental Health Survey, 2016. *Indian Journal of Psychiatry*, 65(12), 1208–1213. https://doi.org/10.4103/indianjpsychiatry.indianjpsychiatry_827_23
- Tongco, Ma. D. C. (2007). Purposive sampling as a tool for informant selection. *Ethnobotany Research & Application*, 5(2), 147–158.
- Walpole, R. E., Myers, R. H., and Myers, S. L. (2012). *Probability & Statistics for Engineers & Scientists*. Boston: Pearson Prentice Hall.
- Wilks. (1938). The large sample distribution of the likelihood ratio for testing composite hypotheses. *American Mathematical Society*, 231. Retrieved from <https://www.jstor.org/stable/i348755>