

Analisis Faktor yang Mempengaruhi Risiko Gagal Bayar Debitur pada Lembaga Keuangan Mikro Menggunakan Regresi Logistik dan *Ant Colony Optimization* (ACO)

Ratih Hadiani^{1, a)} dan Ayu Nike Retnowati^{1, b)}

¹*Prodi Manajemen, Universitas Informatika dan Bisnis Indonesia (UNIBI), Indonesia*

^{a)}*email: ratih.hadiani@unibi.ac.id*

^{b)}*email: anr3515@unibi.ac.id*

Abstrak

Paper ini bertujuan untuk melakukan analisis faktor-faktor yang mempengaruhi risiko gagal bayar dari calon debitur. Metode yang digunakan adalah regresi logistik dan *Ant Colony Optimization* (ACO). Terdapat beberapa tahap dalam penelitian ini: (1) melakukan standarisasi data pada data faktor risiko calon debitur, (2) menetapkan asumsi model regresi logistik, (3) melakukan estimasi parameter model regresi logistik menggunakan algoritma ACO, dan (4) melakukan uji signifikansi setiap variabel. Dalam paper ini, data yang digunakan adalah data historis debitur sebuah Bank pada periode 2001-2011 pada Lembaga Keuangan Mikro (LKM) di Bandung, Indonesia. Hasilnya signifikansi koefisien regresi menunjukkan bahwa lima faktor yang dianalisis berpengaruh signifikan terhadap risiko gagal bayar, yaitu: usia, jumlah tanggungan keluarga, nilai jaminan, besarnya kredit yang diajukan, dan jangka waktu pengembalian kredit, dengan kekuatan korelasi sebesar 93.5%. Probabilitas risiko gagal bayar ditentukan oleh kelima faktor yang mempengaruhi. Mengetahui nilai probabilitas sangat berguna bagi LKM guna menentukan klasifikasi faktor kelayakan pemberian kredit berdasarkan predikat risiko calon debitur. Demikian sehingga, LKM dapat mengetahui faktor-faktor risiko gagal bayar dan mengambil keputusan pemberian kredit yang layak atau tidak layak.

Kata kunci: Ant Colony Optimization (ACO), faktor risiko debitur, Lembaga Keuangan Mikro (LKM), model regresi logistik, risiko gagal bayar

Abstract

This article is aimed to analyze the factors that influence the risk of default from prospective debtors. The methods that are used are logistic regression and *Ant Colony Optimization* (ACO). There are some steps in this research, such as: (1) standardizing the data of prospective debtor risk factor; (2) defining the assumptions of logistic regression model; (3) estimating the parameters of logistic regression model by using ACO algorithm; and (4) significance testing for each variable. In this article, the data used is debtor's historical banking data for 2001-2011 period, in microfinance institution Bandung, Indonesia. As a result, the factors are regression coefficient significance shows that five analyzed factors give impact to the risk of default. They are age, number of family dependents, value of collateral, amount of credit applied for, and credit repayment period. The correlation is 93.5%. Probability of risk of default is determined from the five factors. Finding out the value of probability is very useful for microfinance institutions to determine the classification of

credit eligibility factors based on the risk predicate of the prospective debtor. In this way, microfinance institutions can identify risk factors for default and make decisions about granting credit that is appropriate or not.

Keywords: Ant Colony Optimization (ACO), logistic regression method, microfinance institution, risk factor of debtor, risk of default

Pendahuluan

Dewasa ini mekanisme keuangan dan kredit mikro merupakan strategi penting dalam penanggulangan kemiskinan. Buktinya adalah pemerintah pusat maupun daerah menyalurkan dana bergulir kepada kelompok masyarakat melalui Lembaga Keuangan Mikro (LKM). Harapan dari mekanisme tersebut adalah membantu ekonomi masyarakat dengan menyalurkan dana kepada usaha mikro yang terus berkembang [1]. Penyaluran dana dari LKM kepada usaha-usaha mikro, salah satunya dengan transaksi kredit. Transaksi kredit ini dapat terjadi jika terdapat suatu keinginan, khususnya para pengusaha yang kekurangan modal untuk memperlancar usahanya. Pada proses penyaluran kredit, LKM sering dihadapkan pada suatu risiko yang dikenal dengan risiko gagal bayar debitur [2], [3]. Oleh karena itu, diperlukan analisis faktor-faktor yang mempengaruhi risiko gagal bayar untuk mengantisipasi terjadinya kesalahan LKM dalam menganalisis kredit terhadap debiturnya. Adapun hasil analisis faktor-faktor risiko berupa klasifikasi predikat kredit untuk debitur dalam kategori baik atau buruk sebagai acuan untuk pemberian kredit [4],[5].

Beberapa kajian terkait dalamantisipasi gagal bayar kredit dengan berbagai teknik klasifikasi calon debitur yang telah diusulkan diantaranya: pengklasifikasian menggunakan *Neural Networks* dan *SVM* digambarkan sebagai fungsi matematika yang kompleks [6]. Mempelajari pembentukan model penilaian kredit untuk lembaga keuangan di Bank Jerman menggunakan data pendekatan penambangan untuk analisis [7]. Penerapan *Ant Colony Optimasi (ACO)* pada bidang penambangan data untuk mengekstraksi pengklasifikasi berbasis aturan [8],[9]. Penelitian tentang performansi model *credit scoring* menggunakan metode regresi logistik biner dan teknik data mining seperti *Classification and Regression Tree (CART)*, *Chi-squared Automatic Interaction Detection (CHAID)*, *Neural Network*, serta *Multivariate Adaptive Regression Spline (MARS)* [10]. Penelitian risiko kredit di bank komersial Pakistan dilakukan dengan model penilaian kredit dengan model penilaian kredit untuk individu, regresi logistik, dan analisis diskriminan [11]. Penelitian lain tentang penilaian kredit didasarkan pada analisis diskriminan, yaitu metode statistik yang dirancang untuk mengklasifikasikan debitur ke dalam kelompok yang dapat dibedakan dengan jelas secara optimal baik atau buruk yang dilakukan pada penilaian risiko kredit Usaha Kecil dan Menengah (UKM) [12].

Penggunaan metode pendekatan algoritma *ACO* pada model regresi logistik untuk menganalisis faktor-faktor risiko gagal bayar debitur belum pernah dilakukan. Algoritma *ACO* ini menggunakan lima faktor yang berpengaruh signifikan terhadap risiko gagal bayar, yaitu usia, jumlah tanggungan keluarga, nilai jaminan, besarnya kredit yang diajukan, dan jangka waktu pengembalian kredit dengan tingkat korelasi sebesar 93,5%. Algoritma *ACO* yang dihasilkan lalu diterapkan untuk mengklasifikasikan debitur terhadap risiko gagal bayarnya. Tujuan dari paper ini adalah untuk mendapat klasifikasi subjek calon debitur menjadi beberapa interval tingkat risiko gagal bayar untuk yang baik atau buruk. Penelitian ini melakukan analisis faktor-faktor risiko gagal bayar untuk mengklasifikasikan predikat risiko debitur [13],[14]. Proses analisis dilakukan dengan regresi logistik berdasarkan faktor-faktor risiko debitur, dimana untuk menaksir parameter dilakukan menggunakan algoritma *ACO*. Selanjutnya, estimator model regresi logistik digunakan untuk menentukan

probabilitas gagal bayar dari calon debitur. Taksiran probabilitas gagal bayar kemudian dicocokkan dengan interval kelayakan debitur, sehingga setiap debitur menyandang predikat kredit yang menggambarkan klasifikasi tingkat risiko gagal bayar. Berdasarkan klasifikasi tersebut, LKM mendapat analisis faktor risiko untuk mengambil sebuah keputusan pemberian kredit kepada debitur.

Metode

1. Metode Penelitian

Dalam penelitian ini secara keseluruhan akan dilakukan analisis secara kuantitatif dengan memanfaatkan matematika dalam penyelesaian perhitungan dan studi pustaka. Studi kepustakaan adalah teknik pengumpulan data dengan mengadakan studi penelaahan terhadap buku-buku, literatur-literatur yang berhubungan dengan masalah yang akan diselesaikan dalam penelitian ini. Masalah yang diteliti yaitu analisis faktor-faktor yang mempengaruhi risiko gagal bayar debitur pada LKM. Hasil dari penelitian ini adalah predikat klasifikasi risiko kredit berdasarkan faktor risiko gagal bayar debitur sebagai acuan pemberian kredit yang tepat sasaran untuk LKM.

2. Data Penelitian

Data yang digunakan pada penelitian ini diperoleh dari data kredit konsumtif (non ritel) debitur Bank Z periode 2001-2011. Data terdiri dari 100 sampel yang terbagi dalam 2 kategori, yaitu kategori 0 sebagai kredit tidak bermasalah atau dikatakan layak berjumlah $n_0 = 84$ dan kategori 1 yang dikatakan kredit bermasalah atau tidak layak berjumlah $n_1 = 16$. Variabel-variabel yang berpengaruh pada analisis penilaian kredit yang dilakukan dalam penelitian ini, yaitu terdiri dari lima faktor (variabel bebas), antara lain usia calon debitur (X_1), jumlah tanggungan keluarga (X_2), nilai jaminan (X_3), besarnya kredit yang diajukan (X_4), dan jangka waktu pengembalian kredit (X_5).

Lima faktor sebagai parameter analisis pada penelitian ini berdasarkan kriteria variabel debitur, sebagai berikut: (1) Usia (X_1): Bank Indonesia membatasi usia 21 tahun hingga 65 tahun yang dapat memiliki akses kredit, aturan pembatasan usia ini masuk dalam revisi peraturan Bank Indonesia Nomor 11 Tahun 2009; (2) Jumlah Tanggungan Keluarga (X_2): Jumlah anggota keluarga yang ditanggung oleh pihak debitur merupakan jumlah tanggungan keluarga. Semakin sedikit jumlah tanggungan keluarga, semakin kecil kemungkinan terjadinya gagal bayar; (3) Nilai Jaminan (X_3): Nilai jaminan merupakan sejumlah nilai dari barang yang dijamin debitur sebagai syarat pengajuan kredit, yang diukur dalam satuan rupiah. Nilai jaminan yang tinggi menjadi bahan pertimbangan yang baik untuk memberikan pinjaman kredit kepada debitur; (4) Besarnya Kredit yang Diajukan (X_4): Besarnya kredit yang diajukan debitur akan menjadi bahan pertimbangan dalam mengambil keputusan. Lembaga keuangan akan menyeleksi calon debitur berdasarkan riwayat finansial dan aset yang dimiliki sekarang. Hal ini dilakukan untuk meminimalisir terjadinya gagal bayar; (5) Jangka Waktu Pengembalian Kredit (X_5): Jangka waktu pengembalian kredit bergantung pada besarnya pinjaman dan kemampuan bayar debitur. Jangka waktu ini dinyatakan dalam satuan tahun atau bulan. Pada penelitian ini analisis faktor risiko yang digunakan yaitu analisis regresi logistik biner.

3. Model Regresi Logistik Biner

Analisis regresi logistik merupakan alat analisis data yang digunakan pada penelitian dengan tujuan untuk mengidentifikasi pengaruh variabel bebas (X) terhadap variabel terikat (Y), dimana variabel bebas dalam penelitian bersifat kategorik. Prinsip-prinsip yang digunakan dalam analisis regresi logistik, pada dasarnya sama dengan prinsip dalam analisis regresi linier secara umum.

Perbedaannya hanya dalam hal skala pengukuran dari variabel bebas (Y), sehingga teknik-teknik yang digunakan dalam analisis regresi linier juga dapat digunakan dalam analisis regresi logistik [15],[16].

Analisis regresi logistik biner digunakan untuk menaksir pengaruh dari beberapa variabel penjelas (X), terhadap variabel respon (Y), yang bersifat biner atau dikotomi. Variabel ini dikatakan biner atau dikotomi karena memiliki dua nilai kemungkinan, yaitu 0 dikatakan berhasil dan 1 dikatakan gagal. Bentuk persamaan regresi logistik biner yang digunakan adalah:

$$\pi(x_i) = \frac{\exp(\beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki})}{1 + \exp(\beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki})} \quad i = 1, 2, \dots, N \quad (1)$$

Fungsi $\pi(X)$ berbentuk non linier sehingga perlu ditransformasikan dengan menggunakan transformasi logit, sehingga diperoleh fungsi $g(X)$ yang linier dalam parameter-parameternya [16]. Fungsi logitnya adalah:

$$g(X) = \log \left[\frac{\pi(X)}{1 - \pi(X)} \right] = \sum_{k=0}^K x_{ik} \beta_k \quad (2)$$

4. Penaksiran Parameter Model Regresi Logistik

Tujuan dari penaksiran model regresi logistik adalah untuk menaksir parameter-parameter β_k , yang berkontribusi pada persamaan (2). Andaikan terdapat k variabel bebas X_1, X_2, \dots, X_k , merujuk fungsi densitas peluang bersyarat Y terhadap β mengikuti distribusi bernoulli adalah [17]:

$$f(y|\beta) = \prod_{i=1}^N \pi_i^{y_i} (1 - \pi_i)^{1-y_i} \quad y_i = 0, 1 \quad (3)$$

Y diberi kode dengan 0 dan 1 untuk setiap pasangan x_i, y_i . Jika $y_i = 1$, maka kontribusi untuk fungsi Likelihood adalah $\pi(x_i)$, dan jika $y_i = 0$ maka kontribusi untuk fungsi Likelihood adalah $1 - \pi(x_i)$, di mana $\pi(x_i)$ menuliskan nilai dari $\pi(x)$ pada x_i [18]. Demikian sehingga kontribusi untuk fungsi likelihood dari pasangan (x_i, y_i) adalah:

$$l(\beta) = \sum_{i=1}^N \left\{ y_i \sum_{j=0}^J x_{ij} \beta_j - \ln \left(1 + e^{\sum_{j=0}^J x_{ij} \beta_j} \right) \right\} \quad (4)$$

5. Ant Colony Optimization (ACO)

Algoritma *Ant Colony Optimization* (ACO) diusulkan oleh Marc Dorigo pada tahun 1992, sebagai algoritma pencarian metaheuristik yang terinspirasi dari kelompok semut yang biasa dilakukan untuk mencari jalur terpendek dalam mencari makanan. Semut menggunakan komunikasi tidak langsung yang terjadi di antara semut ketika melakukan perjalanan ke sumber makanan, dan kemudian membawa kembali ke sarang mereka. Algoritma ACO adalah sebuah model yang dikembangkan dengan melihat semut sebagai objek utama pembentuk algoritmanya [18],[19].

Sesuai dengan algoritma ACO yang terinspirasi oleh perilaku koloni semut mencari makanan. Semut menemukan jarak terpendek antara sarang semut dan sumber makanan. Untuk menandai jalan yang mereka lalui ditandai dengan feromon. Feromon adalah bahan kimia yang berasal dari kelenjar endokrin. Semut dapat mencium bau feromon, dan cenderung memilih jalur yang telah ditandai dengan feromon. Jika semut telah menemukan jalur terpendek, semut melanjutkan perjalanannya melalui jalur itu. Jalur terpendek memiliki ketebalan feromon dengan probabilitas

tinggi [20],[21],[22],[23]. dalam algoritma ACO, semut dapat berpindah dari vertex i ke vertex j dengan probabilitas yang dihitung sebagai berikut [24]:

Pemilihan Jalur: Semut berjalan dari titik i ke titik j dengan probabilitas seperti diberikan dengan persamaan berikut:

$$p_{i,j} = \frac{(\tau_{i,j}^\alpha)(\eta_{i,j}^\beta)}{\sum (\tau_{i,j}^\alpha)(\eta_{i,j}^\beta)} \quad (5)$$

di mana $\tau_{i,j}$: jumlah feromon di samping i,j ; α : parameter pengontrol pengaruh $\tau_{i,j}$; $\eta_{i,j}$: sisi diinginkan i,j (umumnya $1/d_{i,j}$, di mana d adalah jarak); β : parameter pengontrol pengaruh $\eta_{i,j}$.

Penambahan dan penguapan feromon, secara matematis diberikan sebagai persamaan berikut ini:

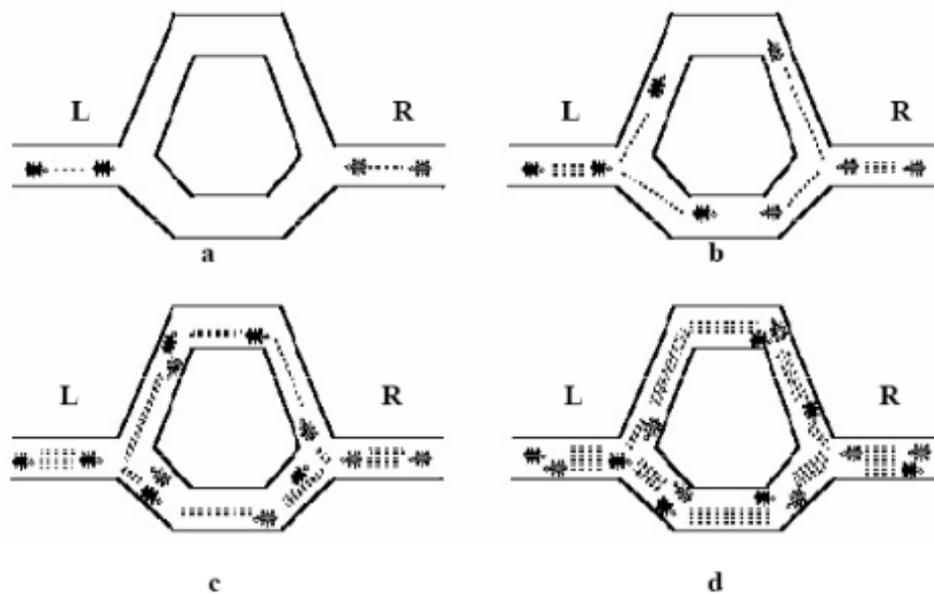
$$\tau_{i,j} = (1 - \rho)\tau_{i,j} + \Delta\tau_{i,j} \quad (6)$$

dimana ρ : tingkat penguapan feromon; dan $\Delta\tau_{i,j}$: jumlah feromon yang dihasilkan.

Selanjutnya, memperbarui feromon secara algoritma matematis diberikan seperti pada persamaan sebagai berikut:

$$\tau(t, v) \leftarrow (1 - \alpha) \cdot \tau(t, v) + \alpha \cdot \Delta\tau(t, v) \quad (7)$$

Pola rute perjalanan semut-semut dari sarang menuju sumber makanan, dan sebaliknya dari sumber makanan kembali ke sarangnya, dapat diilustrasikan seperti pada Gambar 1.



Gambar 1. Perjalanan semut dari sarang ke sumber makanan

Memperhatikan Gambar 1, dapat diceritakan secara singkat sebagai berikut: (a) Pada awalnya semut-semut melakukan perjalanan secara acak dari sarang ke sumber makanan, ada yang melalui jalur panjang dan melalui jalur pendek. (b) Baik semut-semut yang melalui jalur panjang dan jalur pendek ditandai menggunakan feromon. (c) Semut-semut telah menemukan jalur terpendek, dan memperbarui feromon. (d) Akhirnya semua semut-semut berlalu-lalang dari sarang ke sumber makanan, dan sebaliknya melalui jalur terpendek.

Prinsip perjalanan semut menemukan jalur terpendek inilah, selanjutnya digunakan untuk estimasi parameter model regresi logistik, dengan menggunakan suatu algoritma yang dinamakan Ant Colony Optimization (ACO).

1) Uji Signifikansi Parameter

Uji signifikansi parameter di sini, dimaksudkan untuk menguji tingkat signifikansi dari estimator parameter-parameter yang telah diperoleh dari estimasi yang dilakukan menggunakan ACO. Pada uji signifikansi parameter logistik dilakukan dengan beberapa pengujian diantaranya yaitu, uji Likelihood ratio, uji Wald, uji Hosmer dan Lemeshow, dan R^2 dengan $\alpha = 5\%$.

- Uji Likelihood Ratio

Pada bagian ini dibahas metode pengujian serempak menggunakan statistik uji Likelihood ratio G . langkah-langkah uji Likelihood ratio yaitu perumusan hipotesis: $H_0 : \beta_1 = \beta_2 = \dots \beta_k = 0$, artinya secara simultan variabel bebas tidak mempunyai pengaruh terhadap $\pi(X)$; melawan alternatif $H_1 : \beta_1 \neq \beta_2 \neq \dots \beta_k \neq 0$, artinya secara simultan variabel bebas mempunyai pengaruh terhadap $\pi(X)$ [15]. Selanjutnya, menentukan statistik Likelihood ratio G hasil perhitungan berdasarkan persamaan:

$$G = -2 \frac{\left(\frac{n_1}{n}\right)^{n_0}}{\prod \pi_i^{y_i} (1 - \pi)^{1-y_i}}; i = 1, 2, \dots, k, \quad (8)$$

dengan $i = 1, 2, \dots, k$ dengan k banyaknya variabel bebas. Menggunakan tingkat signifikansi α , nilai kritis statistik $\chi^2_{(1-\alpha)(1)}$ diperoleh dari Tabel χ^2 (Chi-Square). Kemudian membandingkan statistik G dan $\chi^2_{(1-\alpha)(1)}$, dengan kriteria sebagai berikut: Jika $\hat{G} < \chi^2_{(1-\alpha)(1)}$ maka H_0 diterima dan H_1 ditolak, dan Jika $\hat{G} \geq \chi^2_{(1-\alpha)(1)}$ maka H_0 ditolak dan H_1 diterima.

- Uji Wald

Merujuk Hosmer et al. (1989), langkah-langkah uji Wald yaitu perumusan hipotesis: $H_0 : \beta_i = 0$, ($i = 1, 2, \dots, k$ dengan k banyaknya variabel bebas), artinya variabel X tidak mempunyai pengaruh terhadap $\pi(X)$; melawan alternatif $H_1 : \beta_i \neq 0$, artinya variabel X mempunyai pengaruh terhadap $\pi(X)$ [15]. Selanjutnya, menentukan statistik Z hitung, dan nilai kritis statistik $Z_{(1-\frac{1}{2}\alpha)}$ diperoleh dari tabel distribusi normal standar. Kemudian membandingkan statistik Z dan $Z_{(1-\frac{1}{2}\alpha)}$, dengan kriteria: Jika $-Z_{\frac{1}{2}\alpha} < Z < Z_{(1-\frac{1}{2}\alpha)}$ maka H_0 diterima dan H_1 ditolak. Jika $Z > Z_{(1-\frac{1}{2}\alpha)}$ atau $Z < -Z_{\frac{1}{2}\alpha}$ maka H_0 ditolak dan H_1 diterima.

- Uji Kecocokan Model Regresi Logistik

Merujuk Hosmer et al. (1989), langkah-langkah pengujian kecocokan model regresi logistik dapat dilakukan menggunakan uji Hosmer dan Lemeshow [15]. Perumusan hipotesis yang digunakan adalah H_0 : Tidak terdapat perbedaan antara hasil pengamatan dengan hasil penaksiran; melawan alternatif H_1 : Terdapat perbedaan antara hasil pengamatan dengan hasil penaksiran. Statistik uji yang digunakan adalah sebagai berikut:

$$C = \sum_{j=1}^n \frac{(o_k - n_k \pi_k)}{n_k \pi_k (1 - \pi_k)}, \quad (9)$$

dengan $o_k = \sum_{j=1}^{C_k} y_j$, $\pi_k = \sum_{j=1}^{C_k} \frac{m_j \hat{\pi}_j}{n_k}$, dan n_k banyaknya pengamatan grup k. Selanjutnya, membandingkan antara nilai statistik C dengan nilai kritis statistik Chi Squared $\chi^2_{(\alpha)(v-1)}$ yang diperoleh dari tabel χ^2 (Chi-Square), dimana α tingkat signifikansi dan v derajat kebebasan. Jika nilai $C > \chi^2_{(\alpha)(v-1)}$ atau $P - value < \alpha$ maka H_0 ditolak dan H_1 diterima.

- Menentukan nilai R-kuadrat (R^2)

Nilai R-kuadrat R^2 adalah digunakan untuk mengukur kekuatan korelasi antara variabel bebas dengan variabel tak bebas. Besarnya nilai statistik R^2 dapat ditentukan dengan menggunakan persamaan:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \pi_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (10)$$

dengan y_i adalah nilai biner hasil observasi ke- i (misalnya 1 atau 0), π_i adalah prediksi nilai probabilitas ke- i , $\bar{y} = \sum_{i=1}^n y_i / n$, dan n banyak data yang diobservasi. Nilai $0 \leq R^2 \leq 1$, dimana jika nilai R^2 mendekati 0 maka berarti korelasinya lemah, dan jika nilai R^2 mendekati 1 maka berarti korelasinya kuat.

2) Probabilitas Risiko Gagal Bayar

Probabilitas risiko gagal bayar (problem loans) adalah merupakan salah satu dampak negatif dari kegiatan pemberian kredit. Hubungan antara peluang gagal bayar dengan $\Lambda(score_i)$, yaitu: (a) $Prob(Default_i) = \Lambda(score_i)$ jika terjadi *problem loans*; (b) $Prob(No Default_i) = 1 - \Lambda(score_i)$ jika tidak terjadi *problem loans* [17], [25].

Peluang gagal bayar (problem loans) adalah $0 \leq \Lambda(score_i) \leq 1$. Debitur dengan $\Lambda(score_i)$ kredit yang rendah memiliki peluang gagal bayar (problem loans) rendah, sedangkan debitur $\Lambda(score_i)$ kredit yang tinggi memiliki peluang gagal bayar (problem loans) tinggi. Oleh karena itu, sebuah $\Lambda(score_i)$ kredit yang tinggi dapat mengakibatkan penolakan kredit.

Probabilitas risiko gagal bayar pada tahapan ini adalah digunakan untuk referensi pengambilan keputusan pemberian kredit. Keputusan ini diambil oleh pihak-pihak bank berdasarkan kriteria-kriteria tertentu, yaitu: (1) Jika diambil $0 \leq \pi(X) \leq 0.40$ maka permohonan kredit diterima; (2) Jika diambil $0.41 \leq \pi(X) \leq 0.70$ maka permohonan kredit dipertimbangkan; (3) Jika diambil $0.71 \leq \pi(X) \leq 1$ maka permohonan kredit ditolak.

Hasil dan Diskusi

Dalam bagian ini dilakukan pembahasan tentang standarisasi data, penaksiran parameter, pengujian signifikansi parameter, dan diskusi ilustrasi keputusan pemberian kredit calon debitur.

1. Standarisasi Data

Standarisasi data merupakan proses standarisasi terhadap data yang memiliki perbedaan nilai yang cukup besar. Standarisasi data dilakukan menggunakan Z-score, dan prosesnya dapat dilakukan dengan menggunakan software SPSS versi 23.00 agar lebih efisien. Sebagai ilustrasi, sebagai contoh digunakan sebagian data dari penelitian seperti diberikan pada Tabel 1.

Tabel 1. Data kredit

Usia Kreditor	Tanggung Keluarga	Jaminan Kredit	Kredit Diajukan	Waktu Kredit
30	3	5.000.000	6.000.000	2
32	4	6.400.000	5.000.000	2
41	4	7.560.000	9.000.000	3

42	5	71.625.000	80.000.000	7
----	---	------------	------------	---

Seperti dapat dilihat pada Tabel 1, di mana setiap variabel memiliki satuan yang sangat jauh berbeda. Bila data tersebut langsung dianalisis maka hasil output yang diperoleh tidak menggambarkan kekuatan pengaruh masing-masing variabel independen. Oleh karena itu, perlu dilakukan transformasi data menjadi Z-score. Hasil nilai Z-score menggunakan bantuan software SPSS versi 23.00, diperoleh hasil standarisasi seperti yang diberikan pada Tabel 2.

Tabel 2. Hasil standarisasi data

Usia Kreditor	Tanggungan Keluarga	Jaminan Kredit	Kredit Diajukan	Waktu Kredit
-0.625	-0.569	-0.317	-0.236	-0.792
-0.365	0.364	-0.210	-0.305	-0.792
0.804	0.364	-0.121	-0.030	-0.126
0.934	0.130	4.776	4.841	2.536

Seperti tampak pada Tabel 2, jika nilai standarisasi data dibatasi nilainya antara -3 sampai +3, pada hasil yang diperoleh terdapat nilai Z-score melebihi +3. Misalnya pada baris keempat pada kolom jaminan dan kredit yang diajukan, nilai Z-score yang diperoleh adalah 4.776 dan 4.841, sehingga data tersebut adalah outliers. Data outlier demikian dalam penelitian ini tidak diikutkan dalam analisis penaksiran parameter, yang dilakukan pada bagian berikut ini.

2. Penaksiran Parameter

Prosedur penaksiran parameter dilakukan dengan merujuk pada persamaan (4), dengan menggunakan algoritma ACO yang dibahas pada bagian 2.2.3, dan dilakukan menggunakan bantuan software Matlab R2015A [26]. Hasil taksiran parameter dan nilai ratio likelihood diberikan pada Tabel 3.

Berdasarkan taksiran parameter yang disajikan pada Tabel 3, merujuk pada persamaan (1) estimator model regresi logistik yang diperoleh adalah sebagai berikut:

$$\hat{\pi}(x) = \frac{\exp \exp (3.0010 + 0.8386X_1 + 0.3017X_2 + 0.3876X_3 - 0.3113X_4 + 0.3784X_5)}{1 + \exp \exp (3.0010 + 0.8386X_1 + 0.3017X_2 + 0.3876X_3 - 0.3113X_4 + 0.3784X_5)}. \quad (11)$$

Selanjutnya, dengan merujuk ke persamaan (2), diperoleh fungsi logit dari model regresi pada persamaan (8) yaitu:

$$\hat{g}(X) = 3.0010 + 0.8386X_1 + 0.3017X_2 + 0.3876X_3 - 0.3113X_4 + 0.3784X_5 \quad (12)$$

Tabel 3. Taksiran parameter model regresi logistik

Koefisien Parameter	Penaksir Parameter ACO (β)	$SE(\beta)$	$Z_{Wald} \beta = \frac{\beta}{SE(\beta)}$	P-Value	Keputusan
Usia Debitur X_1	0.8386	0.4818	1.7406	0.0000	Significance
Tanggungan Keluarga X_2	0.3017	0.5621	0.5367	0.0211	Significance
Nilai Jaminan X_3	0.3876	0.8190	0.4733	0.0013	Significance
Jumlah Kredit X_4	0.3113	0.2134	1,4588	0.0001	Significance
Jangka Waktu	0.3784	0.6542	0.5784	0.0015	Significance

Pinjaman X_5

Log likelihood ratio 3.0010

3. Pengujian Signifikansi Model

Pada bagian ini dilakukan pengujian signifikansi penaksir model yang diperoleh dari proses penaksiran seperti diberikan dengan persamaan (11). Pengujian signifikansi penaksir model meliputi: pengujian signifikansi serempak, pengujian signifikansi parsial, pengujian kecocokan model, dan penentuan nilai R-kuadrat.

1) Pengujian Serempak

Pada bagian ini dilakukan pengujian penaksir koefisien parameter secara serempak, dengan merujuk pembahasan pada bagian 2.2.4.(a). Hipotesis yang digunakan dalam pengujian ini adalah $H_0: \hat{\beta}_1 = \hat{\beta}_2 = \dots = \hat{\beta}_5 = 0$, artinya semua variabel bebas tidak mempunyai pengaruh terhadap $\pi(X)$; dan $H_1: \exists \hat{\beta}_1 \neq \hat{\beta}_2 \neq \dots \neq \hat{\beta}_5 \neq 0$, artinya terdapat variabel bebas mempunyai pengaruh terhadap $\pi(X)$. Berdasarkan hasil perhitungan yang diberikan pada Tabel 3, diperoleh nilai log likelihood $G = 3.000$, adalah nilai maksimum log likelihood yang diperoleh dengan algoritma ACO. Menggunakan tingkat signifikansi $\alpha = 0.05$, dari tabel distribusi χ^2 (Chi-Square) diperoleh nilai kritis statistik $\chi^2_{(0.05)(5)} = 1.145476226$. Oleh karena nilai statistik log likelihood ratio G lebih besar dari nilai kritis statistik $\chi^2_{(0.05)(5)}$, berarti hipotesis H_0 ditolak atau H_1 diterima. Demikian sehingga terdapat variabel bebas yang secara signifikan mempunyai pengaruh terhadap $\pi(X)$.

2) Pengujian Parsial

Pada bagian ini dilakukan uji signifikansi masing-masing penaksir koefisien parameter yang dihasilkan dari proses penaksiran, yang hasilnya diberikan pada persamaan (11) atau (12). Pengujian signifikansi dari masing-masing penaksir koefisien parameter dilakukan menggunakan statistik uji Wald, dengan merujuk pembahasan pada bagian 2.2.4.(b) adalah seperti sebagai berikut.

Untuk penaksir koefisien parameter Usia X_1 , yaitu $\hat{\beta}_1 = 0.08386$. Hipotesis yang digunakan adalah:

$$H_1: \hat{\beta}_1 = 0, \text{ artinya variabel } X_1 \text{ tidak mempunyai pengaruh terhadap } \pi(x)$$

$$H_1: \hat{\beta}_1 \neq 0, \text{ artinya variabel } X_1 \text{ mempunyai pengaruh terhadap } \pi(x)$$

Berdasarkan hasil yang diberikan pada Tabel 3, diperoleh bahwa untuk penaksir koefisien parameter $\hat{\beta}_1$ atau Usia Debitur X_1 nilai statistik $Z_{Wald} = 1.7406$, sedangkan untuk tingkat signifikansi $\alpha = 0.05$ dari tabel distribusi normal standar diperoleh nilai kritis $Z_{1-\frac{\alpha}{2}} = 1.959963985$. Jadi tampak jelas bahwa nilai $Z > Z_{1-0.05/2}$, yaitu diperoleh keputusan H_0 diterima atau H_1 ditolak, yang artinya bahwa penaksir koefisien parameter dari variabel X_1 secara parsial mempunyai pengaruh terhadap $\pi(x)$ dengan tingkat signifikansi $\alpha = 0.05$. Hal ini juga diperkuat dengan nilai dari P-Value = 0.0000 yang lebih kecil daripada tingkat signifikansi $\alpha = 0.05$.

Menggunakan cara sama, pengujian signifikansi juga dilakukan terhadap penaksir-penaksir koefisien parameter $\hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4$ dan $\hat{\beta}_5$. Hasil pengujian secara keseluruhan dapat dilihat pada Tabel 3.

3) Menguji Kecocokan Model Regresi Logistik

Pada bagian ini dilakukan pengujian kecocokan penaksir model regresi logistik yang diberikan sebagai persamaan (8). Pengujian kecocokan ini dilakukan dengan merujuk pembahasan pada 2.2.4.(c), dimana hipotesis yang digunakan adalah H_0 : tidak ada perbedaan antara nilai observasi dengan nilai prediksi model; dan H_1 : ada perbedaan antara nilai observasi dengan nilai prediksi model. Hasil perhitungan yang dilakukan dengan merujuk persamaan (9) diperoleh nilai statistik $C=5.682$ dengan nilai P-Value = 0.392. Menggunakan tingkat signifikansi $\alpha=0.05$, dari tabel distribusi χ^2 (Chi-Square) diperoleh nilai kritis statistik $\chi^2_{(1-0.05)(4)}=9.487729037$. Oleh karena itu, nampak dengan jelas bahwa nilai statistik C lebih kecil dari nilai kritis $\chi^2_{(1-0.05)(4)}$ atau P-Value $> \alpha$, sehingga hipotesis H_0 diterima atau H_1 ditolak. Hal ini menunjukkan bahwa tidak ada perbedaan antara nilai observasi dengan nilai prediksi model.

4) Menentukan Nilai R^2

Pada bagian ini dilakukan pengukuran kekuatan korelasi antara variabel bebas dengan variabel tak bebas, dengan cara menentukan nilai R^2 merujuk pada persamaan (10). Berdasarkan hasil perhitungan diperoleh nilai $R^2= 0.935$, hal ini menunjukkan bahwa penaksir analisis regresi logistik memiliki hubungan yang kuat antara variabel bebas dan variabel tak bebas. Nilai $R^2 = 0.935$ memberikan arti bahwa variabel bebas (independen) yang terdiri dari usia yang berhutang, jumlah tanggungan keluarga, jumlah tabungan, besarnya jaminan, besarnya kredit, jangka waktu kredit, 93,5% dapat menjelaskan pada variabel peluang gagal bayar, dan hanya 6,5% saja yang dijelaskan oleh variabel lainnya.

4. Analisis Keputusan Pemberian Kredit

Pada bagian ini dilakukan pembahasan pengambilan keputusan pemberian kredit yang didasarkan nilai probabilitas gagal bayar. Nilai probabilitas gagal bayar ditentukan dengan menggunakan penaksir model regresi logistik yang diberikan pada persamaan (11). Berdasarkan taksiran probabilitas gagal bayar, tingkat kelayakan kredit calon debitur dalam hal ini dibagi atas tiga interval kategori kelayakan debitur, seperti diberikan dalam Tabel 4.

Tabel 4. Kelayakan kredit debitur

Probabilitas Gagal Bayar (Problem Loans)	Predikat	Kategori
$\pi(X) \leq 0.49$	A	Layak
$0.50 \leq \pi(X) \leq 0.69$	B	Cukup Layak
$0.70 \leq \pi(X) \leq 1$	C	Tidak Layak

Berdasarkan Tabel 4, kelayakan kredit debitur dibagi menjadi tiga kategori yaitu: jika debitur memperoleh nilai kelayakan kurang atau sama dengan 0.49, maka seorang debitur tersebut dikatakan layak; jika debitur memperoleh nilai di antara 0.50 dan kurang atau sama dengan 0.69, maka seorang debitur dikatakan cukup layak. Cukup layak di sini berarti, seorang debitur masih memungkinkan melakukan transaksi kredit dengan beberapa pertimbangan dari pihak kreditur, misalnya dengan menambah nilai agunan yang diajukan. Adapun jika debitur yang memperoleh nilai lebih dari atau sama dengan 0,70, maka debitur tersebut dikatakan tidak layak dan memiliki problem loans yang tinggi.

Sebagai ilustrasi, misalnya seseorang calon debitur: berusia (X_1) 40 tahun, dengan tanggungan keluarga (X_2) sebanyak 4 orang. Calon debitur ini memberikan jaminan (X_3) berupa BPKB sepeda motor, yang harganya ditaksir bernilai IDR.9.375.000,00. Calon debitur ini mengajukan pinjaman

sebesar (X_4) LKM sebesar IDR.11.000.000,00, untuk keperluan keluarga (konsumtif). Jangka waktu pengembalian kredit (X_5) calon debitur adalah selama 24 bulan (2 tahun). Berdasarkan informasi yang diperoleh, diprediksi nilai peluang gagal bayar (problem loans) calon debitur adalah sebesar $\hat{\pi}(X)=0.374$. Artinya peluang gagal bayar (problem loans) calon debitur sebesar 0.374 atau 37.4%. Berdasarkan tingkat kelayakan kredit debitur yang disajikan pada Tabel 4, seorang calon debitur memiliki predikat A, yaitu layak untuk memperoleh kredit sehingga penulis memberikan saran kepada pihak bank agar menerima usulan kredit tersebut. Demikian sehingga, dapat disimpulkan bahwa calon debitur akan mampu mengembalikan kredit sesuai dengan waktu yang ditentukan. Oleh karena itu, pemberian kredit dapat disetujui atau terealisasi untuk diberikan.

Kesimpulan

Berdasarkan hasil analisis dapat ditarik kesimpulan sebagai berikut: (1) Pada penelitian ini variabel-variabel yang signifikan yaitu usia, jumlah tanggungan keluarga, nilai jaminan, besarnya kredit yang diajukan, dan jangka waktu pengembalian kredit. (2) Analisis korelasi antara variabel bebas dengan variabel tak bebas diperoleh nilai R-Kuadrat sebesar 93,21%, yang menunjukkan memiliki korelasi yang kuat. Demikian sehingga estimator model regresi logistik ini dapat digunakan untuk memprediksi risiko gagal bayar calon debitur, berdasarkan lima faktor yang dianalisis disini.

Ucapan Terima Kasih

Penelitian ini didanai oleh Kementerian Riset Teknologi dan Pendidikan Tinggi Republik Indonesia dengan skema Penelitian Dosen Pemula (PDP) dengan nomor kontrak 025/SP2H/RDPKR-MONO/LL4/2021.

Referensi

- [1] M. A. Sila, "Lembaga keuangan mikro dan pengentasan kemiskinan: kasus lumbung pith nagari di Padang," *Masy. J. Sosiol.*, pp. 1–19, 2014.
- [2] S. Purnama and A. P. Kusumawardhani, "Deteksi Peluang Gagal Bayar Calon Debitur Menggunakan Algoritma Particle Swarm Optimization (PSO) untuk Meningkatkan Kinerja Manajemen Risiko pada Koperasi Simpan Pinjam ABC," *KUBIK J. Publ. Ilm. Mat.*, vol. 6, no. 2, pp. 71–84, 2021.
- [3] D. Damayanti, "Peran Lembaga Keuangan Mikro Terhadap Eksistensi Usaha Mikro (Studi Pada Nasabah PD BPR BANK Sleman)," *Mak. J. Manaj.*, vol. 5, no. 2, pp. 110–118, 2019.
- [4] R. F. Malik and H. Hermawan, "Credit scoring using classification and regression tree (CART) algorithm and binary particle swarm optimization," *Int. J. Electr. Comput. Eng.*, vol. 8, no. 6, p. 5425, 2018.
- [5] A. Blanco, R. Pino-Mejías, J. Lara, and S. Rayo, "Credit scoring models for the microfinance industry using neural networks: Evidence from Peru," *Expert Syst. Appl.*, vol. 40, no. 1, pp. 356–364, 2013.
- [6] S. Razzaki *et al.*, "A comparative study of artificial intelligence and human doctors for the purpose of triage and diagnosis," *arXiv Prepr. arXiv1806.10698*, 2018.
- [7] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad, "State-of-the-art in artificial neural network applications: A survey," *Heliyon*, vol. 4, no. 11, pp. 1-41, 2018.
- [8] B. O. Nyangena, "Consumer credit risk modelling using machine learning algorithms: a comparative approach." Strathmore University, 2019.
- [9] X. Ma and S. Lv, "Financial credit risk prediction in internet finance driven by machine learning," *Neural Comput. Appl.*, vol. 31, no. 12, pp. 8359–8367, 2019.
- [10] F. Shen, R. Wang, and Y. Shen, "A cost-sensitive logistic regression credit scoring model based

- on multi-objective optimization approach," *Technol. Econ. Dev. Econ.*, vol. 26, no. 2, pp. 405–429, 2020.
- [11] K. Ben Addi and N. Souissi, "An Ontology-Based Model for Credit Scoring Knowledge in Microfinance: Towards a Better Decision Making," in *2020 IEEE 10th International Conference on Intelligent Systems (IS)*, 2020, pp. 380–385.
- [12] K. O. Dao, T. Y. Nguyen, S. Hussain, and V. C. Nguyen, "Factors affecting non-performing loans of commercial banks: The role of bank performance and credit growth," Center for Open Science, 2020.
- [13] R. Panigrahi, S. Borah, N. Day, R. Babo, and A. S. Ashour, "Classification and analysis of facebook metrics dataset using supervised classifiers," *Soc. Netw. Anal. Comput. Res. Methods Tech.*, vol. 1, 2018.
- [14] M. R. C. Acosta, S. Ahmed, C. E. Garcia, and I. Koo, "Extremely randomized trees-based scheme for stealthy cyber-attack detection in smart grid networks," *IEEE access*, vol. 8, pp. 19921–19933, 2020.
- [15] D. W. Hosmer Jr, S. Lemeshow, and R. X. Sturdivant, *Applied logistic regression*, vol. 398. John Wiley & Sons, 2013.
- [16] P. D. Allison, *Logistic regression using SAS: Theory and application*. SAS institute, 2012.
- [17] A. S. Sukono, M. Mamat, and K. Prafidya, "Credit scoring for cooperative of financial services using logistic regression estimated by genetic algorithm," *Appl. Math. Sci.*, vol. 8, no. 1, pp. 45–57, 2014.
- [18] Q. D. Buchlak *et al.*, "Machine learning applications to clinical decision support in neurosurgery: an artificial intelligence augmented systematic review," *Neurosurg. Rev.*, vol. 43, no. 5, pp. 1235–1253, 2020.
- [19] E.-A. Costea, "Machine learning-based natural language processing algorithms and electronic health records data," *Linguist. Philos. Investig.*, vol. 19, pp. 93–99, 2020.
- [20] M. Dorigo and T. Stützle, "Ant colony optimization: overview and recent advances," *Handb. metaheuristics*, pp. 311–351, 2019.
- [21] J. Uthayakumar, N. Metawa, K. Shankar, and S. K. Lakshmanaprabu, "Financial crisis prediction model using ant colony optimization," *Int. J. Inf. Manage.*, vol. 50, pp. 538–556, 2020.
- [22] H. Peng, C. Ying, S. Tan, B. Hu, and Z. Sun, "An improved feature selection algorithm based on ant colony optimization," *IEEE Access*, vol. 6, pp. 69203–69209, 2018.
- [23] Y. Khourdifi and M. Bahaj, "Heart disease prediction and classification using machine learning algorithms optimized by particle swarm optimization and ant colony optimization," *Int. J. Intell. Eng. Syst.*, vol. 12, no. 1, pp. 242–252, 2019.
- [24] A. Runka, "Evolving an edge selection formula for ant colony optimization," in *Proceedings of the 11th Annual conference on Genetic and evolutionary computation*, 2009, pp. 1075–1082.
- [25] N. B. Nawai and M. N. B. M. Shariff, "Determinants of repayment performance in microfinance programs in Malaysia," *Labu. Bull. Int. Bus. Financ.*, pp. 14–29, 2013.
- [26] F. Barboza, H. Kimura, and E. Altman, "Machine learning models and bankruptcy prediction," *Expert Syst. Appl.*, vol. 83, pp. 405–417, 2017.