

SKRIPSI

**PERBANDINGAN MODEL REGRESI LOGISTIK BINER
DAN METODE *CLASSIFICATION AND REGRESSION TREES*
(CART)**

**(STUDI KASUS: STATUS PENINGKATAN PENDAPATAN
USAHA ONLINE PADA TAHUN 2020 TERHADAP TAHUN 2019
DI DIY)**



FAIQOTUL MUNA
STATE ISLAMIC UNIVERSITY
21106010040
SUNAN KALIJAGA
YOGYAKARTA

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
YOGYAKARTA**

2025

**PERBANDINGAN MODEL REGRESI LOGISTIK BINER
DAN METODE *CLASSIFICATION AND REGRESSION TREES*
(CART)**

(STUDI KASUS: STATUS PENINGKATAN PENDAPATAN
USAHA ONLINE PADA TAHUN 2020 TERHADAP TAHUN 2019
DI DIY)

Skripsi

Untuk memenuhi sebagian persyaratan
mencapai derajat Sarjana S-1
Program Studi Matematika



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

diajukan oleh
FAIQOTUL MUNA
21106010040
Kepada

PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
YOGYAKARTA

2025



SURAT PERSETUJUAN SKRIPSI/TUGAS AKHIR

Hal : Persetujuan Skripsi / Tugas Akhir
Lamp : -

Kepada
Yth. Dekan Fakultas Sains dan Teknologi
UIN Sunan Kalijaga Yogyakarta
di Yogyakarta

Assalamu'alaikum wr. wb.

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengadakan perbaikan seperlunya, maka kami selaku pembimbing berpendapat bahwa skripsi Saudara:

Nama : Faiqotul Muna
NIM : 21106010040

Judul Skripsi : Perbandingan Model Regresi Logistik Biner dan Metode Classification and Regression Tree (CART) (Studi Kasus: Status Peningkatan Pendapatan Usaha Online pada Tahun 2020 terhadap Tahun 2019 di DIY)

sudah dapat diajukan kembali kepada Program Studi Matematika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata Satu dalam Program Studi Matematika.

Dengan ini kami mengharap agar skripsi/tugas akhir Saudara tersebut di atas dapat segera dimunaqasyahkan. Atas perhatiannya kami ucapkan terima kasih.

Wassalamu'alaikum wr. wb.

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

Yogyakarta, 6 Mei 2025

Pembimbing

Sri Utami Zuliana, S.Si., M.Sc., Ph.D.

NIP. 19741003 200003 2 002



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
FAKULTAS SAINS DAN TEKNOLOGI
Jl. Marsda Adisucipto Telp. (0274) 540971 Fax. (0274) 519739 Yogyakarta 55281

PENGESAHAN TUGAS AKHIR

Nomor : B-1091/Un.02/DST/PP.00.9/06/2025

Tugas Akhir dengan judul : PERBANDINGAN MODEL REGRESI LOGISTIK BINER DAN METODE CLASSIFICATION AND REGRESSION TREES (CART) (STUDI KASUS: STATUS PENINGKATAN PENDAPATAN USAHA ONLINE PADA TAHUN 2020 TERHADAP TAHUN 2019 DI DIY)

yang dipersiapkan dan disusun oleh:

Nama : FAIQOTUL MUNA
Nomor Induk Mahasiswa : 21106010040
Telah diujikan pada : Rabu, 21 Mei 2025
Nilai ujian Tugas Akhir : A

dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

TIM UJIAN TUGAS AKHIR



Ketua Sidang

Sri Utami Zuliana, S.Si., M.Sc., Ph.D.
SIGNED

Valid ID: 6849302790614



Penguji I

Lilih Deva Martias, M.Sc.
SIGNED

Valid ID: 6847aebff3a33



Penguji II

Aulia Khifah Futhona, M.Sc.
SIGNED

Valid ID: 6847b83f0742b



Yogyakarta, 21 Mei 2025
UIN Sunan Kalijaga
Dekan Fakultas Sains dan Teknologi

Prof. Dr. Dra. Hj. Khurul Wardati, M.Si.
SIGNED

Valid ID: 684f8931bb27d

SURAT PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini:

Nama : Faiqotul Muna
NIM : 21106010040
Program Studi : Matematika
Fakultas : Sains dan Teknologi

Dengan ini menyatakan bahwa isi skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar sarjana di suatu Perguruan Tinggi dan sesungguhnya skripsi ini merupakan hasil pekerjaan penulis sendiri sepanjang pengetahuan penulis, bukan duplikasi atau saduran dari karya orang lain kecuali bagian tertentu yang penulis ambil sebagai bahan acuan. Apabila terbukti pernyataan ini tidak benar, sepenuhnya menjadi tanggung jawab penulis.

Yogyakarta, 6 Mei 2025



Faiqotul Muna

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

HALAMAN PERSEMBAHAN

Dengan penuh rasa syukur, karya sederhana ini penulis persembahkan kepada:

Kedua orang tua penulis, kakak dan adik penulis, keluarga dekat, sahabat, serta pengasuh Pondok Pesantren Inayatullah yang memberikan semangat yang tak ternilai.

Terima kasih atas cinta, kesabaran, dan dukungannya selama ini.

Dengan segala hormat, karya ini penulis persembahkan pula kepada Dosen Pembimbing yang telah sabar membimbing dan memberikan ilmu yang sangat bermanfaat.

Tak lupa kepada Almamater kampus tercinta Universitas Islam Negeri Sunan Kalijaga Yogyakarta.

"Matematika angkatan 2021 tercinta"
STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

MOTTO

”dan hanya kepada Tuhanmulah hendaknya kamu berharap.”

(QS. Al Insyirah:8)

”Jika memulai karena Allah maka jangan menyerah karena manusia.”

(Gus Baha)

”Barang siapa menginginkan kebaikan di dunia ini, hendaklah ia mencapainya dengan ilmu. Barang siapa menginginkan kebaikan di akhirat, maka ia harus mencapainya dengan ilmu. Dan barang siapa menginginkan keduanya, hendaklah mencari ilmu.”

(HR. Thabrani)

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

PRAKATA

Puji syukur kehadirat Allah SWT, Tuhan semesta alam yang telah melimpahkan rahmat, hidayah, serta karunia-Nya sehingga penulis dapat menyelesaikan skripsi ini dengan baik. Shalawat serta salam semoga senantiasa tercurah kepada junjungan kita Nabi Muhammad SAW, beserta keluarga, sahabat, dan seluruh umatnya.

Tanpa mengurangi rasa hormat, penulis ingin menyampaikan ucapan terima kasih dan apresiasi yang sebesar-besarnya kepada:

1. Bapak Prof. Noorhaidi, MA., M.Phil., Ph.D. selaku Rektor UIN Sunan Kalijaga Yogyakarta.
2. Ibu Prof. Dr. Dra. Hj. Khurul Wardati, M.Si. selaku Dekan Fakultas Sains dan Teknologi, Universitas Islam Negeri Sunan Kalijaga Yogyakarta.
3. Ibu Dr. Epha Diana Supandi, S.Si., M.Sc., selaku Ketua Program Studi Matematika dan Dosen Pembimbing Akademik yang telah banyak memberikan pengarahan kepada penulis dari awal perkuliahan hingga akhir.
4. Ibu Sri Utami Zuliana, S.Si., M.Sc., Ph.D., selaku Dosen Pembimbing Skripsi atas setiap arahan dan bimbingan yang diberikan. Pertemuan rutin mingguan yang dilakukan secara konsisten telah memberikan kesempatan kepada penulis untuk memonitor perkembangan penulisan skripsi secara berkala. Bimbingan dan masukan yang diberikan telah mendorong penulis untuk senantiasa berupaya memberikan yang terbaik dan menghindari jalan pintas.

5. Seluruh Dosen Program Studi Matematika dan staf Fakultas Sains dan Teknologi yang senantiasa memberikan ilmu dan layanan terbaik kepada penulis dari awal hingga akhir perkuliahan.
6. Kedua orang tua penulis, kakak-adik penulis, dan keluarga besar yang telah memberikan banyak nasehat, kasih sayang, doa dan dukungannya sehingga penulis dapat bertahan hingga titik ini.
7. Keluarga Matematika angkatan 2021 yang telah kebersamai selama 4 tahun dibangu perkuliahan, saling mendukung, berbagi ilmu dan banyak kenangan manis yang selalu dikenang.
8. Lidyia dan teman-teman seperbimbingan Sindi, Laili, Zaza, dan Citra yang telah memberikan dukungan moral dan semangat. Dukungan yang kalian berikan telah menjadi motivasi tersendiri bagi penulis dalam menyelesaikan penulisan skripsi ini. Diskusi-diskusi yang kita lakukan bersama sangat bermanfaat dalam memahami konsep-konsep yang sulit.
9. Pengasuh Pondok Pesantren Inayatullah, Abah Yai Chamdani Yusuf dan Umi Nyai Nailil Izzah yang tanpa lelah memberikan didikan, nasihat, serta menjadi orang tua sambung bagi penulis selama penulis tinggal di Yogyakarta.
10. Keluarga besar Pondok Pesantren Inayatullah yang telah memberikan wara di setiap lembaran cerita penulis, menjadi energi positif dan keluarga di perantauan ini.
11. Laila, Dewi, Siska, mbak Naili, Najwa, mbak Reza, mbak Ais dan teman-teman dekat lainnya yang selalu siap mendengarkan segala keluh kesah penulis.

12. Teman-teman KKN 272 Wonosari yang sangat humoris dan supportif, terima kasih cerita 49 hari nya selama KKN di Lereng Gunung Kawi.
13. Teman-teman satu jiwa yang selalu kebersamai dan meramaikan suasana di kamar.
14. Teman-teman mahasiswa pendamping (MP) PPKK Saintek dan adik-adik maba Saintek 2022-2024 yang penulis dampingi.
15. Terakhir kepada semua pihak yang tidak dapat penulis sebut satu-persatu.

Penulis menyadari bahwa penyusunan skripsi ini masih jauh dari sempurna. Oleh karena itu, segala bentuk kritik dan saran yang bersifat membangun sangat diharapkan untuk perbaikan di masa mendatang. Penulis berharap karya tulis ini dapat memberikan kontribusi positif bagi perkembangan ilmu pengetahuan, khususnya dalam bidang statistika matematika. Semoga Allah SWT senantiasa melimpahkan rahmat dan hidayah-Nya kepada kita semua, *Aamiin*.

Yogyakarta, 22 April 2025

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA
Faiqotul Muna

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PERSEMBAHAN	v
HALAMAN MOTTO	vi
PRAKATA	vii
DAFTAR ISI	x
DAFTAR TABEL	xiv
DAFTAR GAMBAR	xvi
DAFTAR LAMBANG	xvii
INTISARI	xx
ABSTRACT	xxi
I PENDAHULUAN	1
1.1. Latar Belakang Masalah	1
1.2. Rumusan Masalah	4
1.3. Batasan Masalah	4
1.4. Tujuan Penelitian	5
1.5. Manfaat Penelitian	5
1.6. Sistematika Penulisan	6
1.7. Tinjauan Pustaka	7
II DASAR TEORI	10
2.1. Model Linier	10
2.2. <i>Generalized Linear Model</i>	12
2.3. Distribusi Binomial	16

2.4. Regresi Logistik	17
2.4.1. Pengertian Regresi Logistik	17
2.4.2. Hubungan Regresi Linier dan Regresi Logistik	17
2.4.3. Model Regresi Logistik	18
2.4.4. Asumsi dalam Regresi Logistik	18
2.5. Regresi Logistik Biner	19
2.6. Normalisasi Data	21
2.7. Data Training dan Data Testing	23
2.8. Uji Multikolinieritas	24
2.8.1. Pengertian Multikolinieritas	24
2.8.2. Metode Deteksi Multikolinieritas	25
2.9. Estimasi Parameter	26
2.10. Uji Signifikansi Parameter	29
2.11. Uji Kecocokan Model	31
2.12. Interpretasi Koefisien Regresi	33
2.13. <i>Classification and Regression Trees</i> (CART)	35
2.14. Cara Kerja Algoritma <i>Classification and Regression Trees</i> (CART)	38
2.14.1. Konstruksi Pohon Maksimum	38
2.14.2. Penentuan Ukuran Pohon Optimal (<i>Pruning</i>)	42
2.14.3. Klasifikasi Data Baru	43
2.15. Ketepatan Klasifikasi	44
III METODE PENELITIAN	46
3.1. Pendekatan Penelitian	46
3.2. Jenis dan Sumber Data	46
3.3. Variabel Penelitian	47
3.4. Langkah Analisis Data	47

3.5. Diagram Analisis Data (<i>Flowchart</i>)	49
IV HASIL DAN PEMBAHASAN	50
4.1. Analisis Statistik Deskriptif Awal	50
4.1.1. Matriks Korelasi	54
4.1.2. Dampak Korelasi terhadap Regresi Logistik	55
4.1.3. Dampak Korelasi terhadap Model CART	56
4.2. Data <i>Training</i> dan Data <i>Testing</i>	56
4.3. Pemodelan Regresi Logistik Biner	57
4.3.1. Uji Multikolinieritas	57
4.3.2. Estimasi Parameter	58
4.3.3. Uji Signifikansi Parameter	60
4.3.4. Model Regresi Logistik Biner	68
4.3.5. Uji Kecocokan Model	70
4.3.6. Interpretasi Koefisien Regresi	71
4.3.7. Ketepatan Klasifikasi Regresi Logistik Biner	74
4.4. Analisis <i>Classification and Regression Trees</i> (CART)	75
4.4.1. Konstruksi Pohon Maksimum	75
4.4.2. Penentuan Ukuran Pohon Optimal	83
4.4.3. Klasifikasi Data Baru	86
4.4.4. Ketepatan Klasifikasi <i>Classification and Regression Trees</i> (CART)	90
4.4.5. Pemilihan Model terbaik	91
V PENUTUP	94
5.1. Kesimpulan	94
5.2. Saran	97
DAFTAR PUSTAKA	99

LAMPIRAN	104
A Deskripsi Variabel	104
B <i>Source Code</i> Program R	111
C Tinjauan Pustaka	126
D Biodata Penulis	129



DAFTAR TABEL

2.1	Beberapa Fungsi Link yang Umum Digunakan dalam GLM	15
2.2	Tabel Ketepatan Klasifikasi	44
4.1	Statistika Deskriptif	51
4.2	Nilai VIF	57
4.3	Estimasi Parameter	59
4.4	Uji G	61
4.5	Uji Wald Pertama	62
4.6	Uji Wald Kedua	64
4.7	Uji Wald Ketiga	65
4.8	Koefisien Regresi	68
4.9	Hasil Uji Statistik Devians untuk Model Regresi Logistik Biner . . .	70
4.10	Odds Ratio	71
4.11	Kesalahan klasifikasi	74
4.12	Kemungkinan pemisah Setiap Variabel Prediktor	75
4.13	Variabel Penting Pembentukan Pohon Maksimum	77
4.14	Keterangan Variabel	80
4.15	<i>Important Splits</i> dan <i>Improve Values</i>	81
4.16	Tabel <i>Complexity Parameter</i>	83
4.17	Variabel Penting Pohon Optimal	86
4.18	Keterangan Variabel	87
4.19	Tabel Kesalahan klasifikasi	90
4.20	Perbandingan Variabel	91

4.21 Perbandingan Hasil Ketepatan Klasifikasi 92



DAFTAR GAMBAR

3.1	Flowchart Analisis Data	49
4.1	Status Peningkatan Pendapatan	50
4.2	Boxplot Variabel Numerik	52
4.3	Diagram Pie untuk Variabel X_4	53
4.4	Matriks Korelasi Pearson antar Variabel Prediktor	54
4.5	Visualisasi Pohon Keputusan	79
4.6	Plot Kompleksitas Pohon Keputusan	84
4.7	Visualisasi Pohon Setelah Dipangkas	87
5.1	Pohon Klasifikasi Optimal	95

DAFTAR LAMBANG

Y	= Variabel respon
x_j	= Nilai dari variabel prediktor ke- j
β_j	= Koefisien regresi ke- j
X	= Matriks prediktor (termasuk kolom konstanta)
β	= Vektor koefisien regresi
Y	= Vektor respon
ϵ	= Vektor error
β_0	= Intersep (nilai Y ketika semua X bernilai 0)
$\beta_1, \beta_2, \dots, \beta_k$	= Koefisien regresi
x_1, x_2, \dots, x_k	= Variabel prediktor
ϵ	= <i>Error</i> atau residual
η	= Nilai prediktor linier untuk suatu pengamatan
$g(\mu)$	= <i>Link function</i> GLM
μ	= Nilai ekspektasi (rata-rata) dari variabel respon
$P(x)$	= Probabilitas suatu hasil tertentu dalam distribusi probabilitas
π	= Probabilitas keberhasilan dalam distribusi binomial
$1 - \pi$	= Probabilitas kegagalan dalam distribusi binomial
π_i	= Probabilitas hasil kategori ke- i
X^T	= Matriks hasil transpose
j	= Indeks untuk kategori ke- j
n	= Jumlah total percobaan

x	= Banyaknya keberhasilan
$x!$	= Faktorial dari x , yaitu $x \cdot (x - 1) \cdots 1$
$0 \leq \pi \leq 1$	= Batasan probabilitas dalam distribusi binomial
λ	= Rata-rata jumlah kejadian dalam interval yang diamati
e	= Bilangan euler ($e \approx 2,71828$)
$E(Y)$	= Nilai ekspektasi (rata-rata) dari variabel respon
σ	= Standar deviasi
$\exp(x)$	= Fungsi eksponensial
\log	= Fungsi logaritma natural (basis e)
R_j^2	= Koefisien determinasi dari regresi variabel x_j
$L(\beta)$	= Fungsi <i>log-likelihood</i>
Σ	= Lambang penjumlahan
Π	= Simbol produk yang berarti perkalian dari semua elemen dalam suatu rentang indeks
D	= Statistik devians
χ^2	= Distribusi chi-kuadrat
$\hat{\beta}_j$	= Estimasi dari parameter β_j
$SE(\hat{\beta}_j)$	= Standar error dari estimasi parameter β_j
\mathbf{X}'	= vektor berisi nilai fitur setelah dinormalisasi
\mathbf{X}	= vektor berisi nilai asli
X_{\min}	= nilai minimum dari variabel prediktor
X_{\max}	= nilai maksimum dari variabel prediktor
$R(T)$	= Kesalahan klasifikasi pohon T
(\tilde{T})	= Kompleksitas berdasarkan jumlah simpul terminal
ψ_j	= <i>Odds ratio</i>

L_0	= Likelihood tanpa peubah prediktor
L_k	= Likelihood dengan k peubah prediktor
z^2	= Statistik uji-Wald
G	= Statistik uji-G
$i(t)$	= Fungsi impuritas Gini pada simpul t
$\Delta i(t)$	= Perubahan ukuran impuritas setelah pemisahan
$P(j t)$	= Probabilitas kelas j dalam simpul t
n_l dan n_r	= Jumlah sampel pada node anak kiri dan kanan

INTISARI

Perbandingan Model Regresi Logistik Biner dan Metode *Classification and*

Regression Trees (CART)

(Studi kasus: Status Peningkatan Pendapatan Usaha Online pada Tahun 2020 terhadap Tahun 2019 di DIY)

Faiqotul Muna

21106010040

Regresi logistik biner merupakan teknik analisis untuk melihat hubungan antara variabel prediktor dan variabel respon. Adapun, *Classification and Regression Trees (CART)* adalah metode pohon keputusan yang fleksibel, mampu menangani hubungan non-linear dan interaksi antar variabel secara otomatis. Pada tahun 2020, telah terjadi pandemi COVID-19 yang berdampak pada perekonomian global, termasuk UMKM di Indonesia, di mana beberapa UMKM berhasil meningkatkan pendapatan dengan memanfaatkan platform digital. Penelitian ini membandingkan performa regresi logistik biner dan *Classification and Regression Trees (CART)* dalam mengklasifikasikan faktor yang memengaruhi peningkatan pendapatan usaha online menggunakan data SUSENAS dari BPS tahun 2021. Metode penelitian yang digunakan adalah analisis kuantitatif dengan pendekatan statistik. Hasil menunjukkan regresi logistik biner memiliki ketepatan klasifikasi 76,29% dengan APER 23,71%, sedangkan *Classification and Regression Trees (CART)* memiliki ketepatan klasifikasi 72,16% dengan APER 27,84%. Kesimpulannya, regresi logistik biner menunjukkan hasil yang lebih baik dalam mengklasifikasikan faktor peningkatan pendapatan usaha online, terutama jika pola hubungan dalam data relatif linier dan asumsi model terpenuhi.

Kata Kunci: Regresi Logistik Biner, *Classification and Regression Trees (CART)*, Peningkatan Pendapatan, Klasifikasi Data.

ABSTRACT

Comparison of the Binary Logistic Regression Model and Classification and Regression Trees (CART) Method

(Case study: Status of Online Business Income Increase in 2020 Compared to 2019 in the Special Region of Yogyakarta)

Faiqotul Muna

21106010040

Binary logistic regression is an analytical technique to see the relationship between predictor variables and response variables. Meanwhile, Classification and Regression Trees (CART) is a flexible decision tree method, capable of handling non-linear relationships and interactions between variables automatically. In 2020, there has been a COVID-19 pandemic that has an impact on the global economy, including MSMEs in Indonesia, where some MSMEs have succeeded in increasing revenue by utilizing digital platforms. This study compares the performance of binary logistic regression and Classification and Regression Trees (CART) in classifying factors that influence the increase in online business income using SUSENAS data from BPS in 2021. The research method used is quantitative analysis with a statistical approach. The results show that binary logistic regression has a classification accuracy of 76.29% with an APER of 23.71%, while the Classification and Regression Trees (CART) have a classification accuracy of 72.16% with an APER of 27.84%. In conclusion, binary logistic regression shows better results in classifying factors that increase online business income, especially if the relationship pattern in the data is relatively linear and the model assumptions are met.

Keyword: Binary Logistic Regression, Classification and Regression Trees (CART), Income Increase, Data Classification.

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Pandemi COVID-19 telah memberikan dampak signifikan terhadap perekonomian global, termasuk sektor Usaha Mikro, Kecil, dan Menengah (UMKM) di Indonesia. Di Daerah Istimewa Yogyakarta (DIY), banyak UMKM yang mengalami penurunan pendapatan akibat pembatasan sosial dan perubahan perilaku konsumen (Prabandari & Atmojo, 2022). Meskipun demikian, tidak sedikit UMKM yang berhasil beradaptasi melalui pemanfaatan teknologi digital, seperti platform *e-commerce* dan media sosial, untuk mempertahankan bahkan meningkatkan pendapatan selama masa pandemi (Arianty, 2024).

Salah satu contoh strategi digitalisasi adalah program “Si Bakul Jogja” yang diinisiasi oleh Dinas Koperasi dan UKM DIY. Program ini memfasilitasi UMKM untuk memasarkan produk secara daring, serta memberikan insentif seperti gratis ongkos kirim guna menarik minat konsumen (Prabandari & Atmojo, 2022). Inovasi ini tidak hanya membantu UMKM bertahan, tetapi juga membuat produk mereka bisa dikenal oleh lebih banyak orang meskipun adanya keterbatasan interaksi fisik.

Perspektif keislaman juga menekankan pentingnya usaha, inovasi, dan adaptasi terhadap perubahan. Allah SWT berfirman dalam Surah Ar-Ra'd ayat 11:

لَهُ مَعْقَبَاتٌ مِّنْ بَيْنِ يَدَيْهِ وَمِنْ خَلْفِهِ يَحْفَظُونَهُ مِنْ أَمْرِ اللَّهِ إِنَّ اللَّهَ لَا يُغَيِّرُ مَا بِقَوْمٍ حَتَّىٰ يُغَيِّرُوا مَا بِأَنْفُسِهِمْ
وَإِذَا أَرَادَ اللَّهُ بِقَوْمٍ سُوءًا فَلَا مَرَدَّ لَهُ وَمَا لَهُمْ مِنْ دُونِهِ مِنْ وَالٍ

Dalam tafsir Ibnu Katsir terhadap QS. Ar-Ra'd ayat 11, dijelaskan bahwa perubahan keadaan suatu kaum (baik menuju kebaikan atau keburukan) terjadi karena perubahan sikap dan amal mereka sendiri. Jika mereka meninggalkan ketaatan dan melakukan kemaksiatan, maka nikmat akan diganti dengan azab. Sebaliknya, jika mereka kembali kepada ketaatan, maka azab akan diganti dengan rahmat Allah. Penafsiran ini ditegaskan pula melalui kisah seorang nabi Bani Israil yang diperintahkan menyampaikan bahwa setiap perubahan dalam kehidupan sosial adalah akibat langsung dari perubahan moral dan spiritual masyarakat tersebut. (Ibnu Katsir, 2000)

Penafsiran ini sejalan dengan prinsip bahwa perubahan dan perbaikan hanya akan terjadi apabila manusia berusaha dan beradaptasi terhadap kondisi yang ada. Di mana perubahan dan perbaikan dalam kehidupan tidak terjadi secara tiba-tiba atau semata karena kehendak takdir, tetapi sangat bergantung pada usaha manusia dalam mengubah keadaan dirinya sendiri. Dalam konteks penelitian ini, hal tersebut tercermin dalam upaya pelaku UMKM untuk bertransformasi secara digital dan menyesuaikan strategi bisnis mereka di tengah tantangan pandemi. Upaya ini penting dalam menghadapi tantangan selama pandemi, karena kemampuan beradaptasi menjadi kunci untuk mempertahankan dan meningkatkan pendapatan.

Memahami faktor-faktor yang berkontribusi terhadap status peningkatan pendapatan UMKM selama pandemi menjadi penting untuk perencanaan strategi ekonomi ke depan. Beberapa variabel yang diduga berkontribusi terhadap status peningkatan pendapatan UMKM antara lain jenis usaha, penggunaan media promosi digital, kepemilikan izin usaha, dan jenis kepemilikan usaha. Variabel-variabel tersebut merepresentasikan aspek strategis yang berkaitan dengan adaptasi usaha terhadap kondisi krisis. Misalnya, penggunaan media sosial dapat meningkatkan

jangkauan pemasaran, sementara kepemilikan izin usaha dapat membuka akses ke bantuan formal dan dukungan pemerintah (Moehadi et al., 2025; Rizaldi & Djama-luddin, 2023).

Untuk menganalisis hubungan antara berbagai variabel dengan status peningkatan pendapatan, diperlukan pendekatan statistik yang tepat (Ghozali, 2016). Dalam penelitian ini, status peningkatan pendapatan UMKM online di DIY selama pandemi digunakan sebagai variabel respon yang bersifat biner (meningkat atau tidak). Sementara itu, variabel prediktor meliputi faktor-faktor seperti usia, lama usaha, kategori usaha, dan lainnya. Dua metode yang dapat digunakan untuk tujuan ini adalah regresi logistik biner dan *Classification and Regression Trees* (CART). regresi logistik biner merupakan metode yang umum digunakan untuk memodelkan hubungan antara variabel prediktor dengan variabel respon biner (Hosmer Jr et al., 2013). Adapun, CART adalah metode berbasis pohon keputusan yang lebih fleksibel karena mampu menangani hubungan non-linier dan interaksi antar variabel secara otomatis (Breiman, 2017).

Berbagai penelitian telah menunjukkan keunggulan masing-masing metode dalam berbagai bidang. Namun, penelitian yang secara langsung membandingkan performa regresi logistik biner dan CART dalam konteks status peningkatan pendapatan UMKM online selama pandemi COVID-19 di Yogyakarta masih jarang dilakukan. Oleh karena itu, penelitian ini bertujuan untuk membandingkan performa regresi logistik biner dan CART dalam mengklasifikasikan faktor-faktor yang memengaruhi status peningkatan pendapatan usaha online pada tahun 2020 terhadap tahun 2019 di DIY.

Hasil dari penelitian ini diharapkan dapat memberikan kontribusi terhadap pengembangan strategi berbasis data untuk pemberdayaan UMKM, serta memper-

kaya literatur mengenai pemilihan model klasifikasi yang sesuai dalam bidang ekonomi digital.

1.2. Rumusan Masalah

Mengacu pada permasalahan yang telah diuraikan sebelumnya, penelitian ini dirumuskan untuk menjawab pertanyaan:

1. Langkah-langkah apa saja yang perlu dilalui dalam melakukan analisis data menggunakan model regresi logistik biner dan metode *Classification and Regression Trees* (CART)?
2. Faktor-faktor apa yang memengaruhi status peningkatan pendapatan usaha online pada tahun 2020 terhadap tahun 2019 di DIY berdasarkan model regresi logistik biner dan pengklasifikasian dengan metode *Classification and Regression Trees* (CART)?
3. Metode manakah yang memberikan hasil lebih akurat antara model regresi logistik biner dan *Classification and Regression Trees* (CART) dalam memberikan hasil prediksi status peningkatan pendapatan usaha online pada tahun 2020 terhadap tahun 2019 di DIY?
4. Apakah terdapat kondisi atau jenis data tertentu di mana satu metode lebih unggul dibandingkan yang lain?

1.3. Batasan Masalah

Untuk menghindari terlalu luasnya cakupan penelitian, permasalahan yang akan diteliti difokuskan pada:

1. Analisis dalam penelitian ini akan dibatasi pada dua metode klasifikasi saja

yaitu regresi logistik biner dan *Classification and Regression Trees* (CART).

2. Asumsi linieritas pada regresi logistik dan pembagian data yang dilakukan oleh *Classification and Regression Trees* (CART) akan diikuti sesuai dengan teori dasar masing-masing metode.
3. Program komputasi statistika yang digunakan adalah *software R* versi 4.4.2

1.4. Tujuan Penelitian

1. Mengetahui langkah-langkah menganalisis dengan model regresi logistik biner dan pengklasifikasian dengan metode *Classification and Regression Trees* (CART).
2. Mengetahui faktor-faktor yang dianggap dapat memengaruhi status pendapatan usaha online selama tahun 2020 dibandingkan tahun 2019 di D.I. Yogyakarta, ditinjau berdasarkan model regresi logistik biner dan pengklasifikasian dengan metode *Classification and Regression Trees* (CART) dengan menggunakan data hasil kuesioner yang dikumpulkan oleh BPS dan diakses oleh peneliti melalui sumber resmi.
3. Membandingkan kinerja model regresi logistik biner dan metode *Classification and Regression Trees* (CART) dalam memprediksi status peningkatan pendapatan usaha online pada tahun 2020 terhadap tahun 2019 di DIY.
4. Menyelidiki apakah ada kondisi atau jenis data tertentu yang membuat salah satu metode lebih cocok digunakan dibandingkan metode lainnya.

1.5. Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat berupa:

1. Manfaat Teoritis

Penelitian ini memberikan kontribusi dalam memperdalam pemahaman mengenai model regresi logistik biner dan metode *Classification and Regression Trees* (CART) dalam menentukan faktor-faktor yang memiliki pengaruh signifikan terhadap status peningkatan pendapatan usaha online pada tahun 2020 terhadap tahun 2019 di DIY.

2. Manfaat Praktis

- Penelitian ini membantu mengidentifikasi faktor-faktor risiko yang paling signifikan dalam status peningkatan pendapatan usaha online, sehingga dapat digunakan sebagai dasar pengambilan keputusan strategis.
- Dengan mengetahui faktor-faktor yang memengaruhi keberhasilan usaha online, pelaku bisnis dapat mengembangkan strategi pencegahan dan perencanaan yang lebih efektif.
- Penelitian ini memberikan wawasan mengenai metode mana (regresi logistik biner atau *Classification and Regression Trees* (CART)) yang lebih akurat dalam memprediksi status peningkatan pendapatan usaha online, sehingga dapat digunakan oleh pengambil kebijakan atau pelaku usaha dalam menentukan metode analisis yang paling sesuai.

Penelitian ini diharapkan dapat membantu akademisi dan praktisi dalam memahami serta memilih metode yang paling optimal untuk analisis faktor yang memengaruhi pendapatan usaha online.

1.6. Sistematika Penulisan

Sistematika pada penelitian ini yaitu:

BAB I PENDAHULUAN

Pada bab ini menjabarkan latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, sistematika penulisan, dan tinjauan pustaka.

BAB II DASAR TEORI

Pada bab ini berisi landasan teori sebagai acuan dalam merancang penelitian.

BAB III METODE PENELITIAN

Pada bab ini menyajikan setiap langkah dalam penelitian yang akan digunakan dalam pemecahan masalah meliputi metode penelitian, sumber data, metode pengolahan, dan analisis data.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini memuat semua temuan ilmiah yang di peroleh dalam penelitian dan hasilnya akan di interpretasikan secara terperinci.

BAB V PENUTUP

Pada bab ini berisi kesimpulan dan saran dari penulis yang berasal dari penelitian yang telah di lakukan.

1.7. Tinjauan Pustaka

(Dethan et al., 2024) melakukan penelitian yang berfokus pada identifikasi faktor risiko Infeksi Saluran Pernapasan Akut (ISPA) pada balita di Kabupaten Kupang, Indonesia. Mereka menggunakan regresi logistik biner untuk menganalisis hubungan antara kejadian ISPA dan berbagai faktor risiko seperti status imunisasi, paparan asap rokok, paparan asap kayu, pemberian ASI eksklusif dan status gizi. Model logistik akhir menunjukkan hubungan yang signifikan antara kejadian ISPA dan faktor-faktor tersebut.

Penelitian yang dilakukan oleh (Rohmah et al., 2023) membahas hubungan antara beberapa variabel prediktor dan status migrasi dengan menggunakan regresi logistik berbasis Bayesian. Studi ini menyoroti metode estimasi parameter dengan

pendekatan Bayesian untuk meningkatkan validitas model. Mereka menggunakan metode *Markov Chain Monte Carlo* (MCMC) dalam pemeriksaan konvergensi model.

(Srimaneekarn et al., 2022) meneliti penerapan regresi logistik biner dalam bidang kedokteran gigi. Penelitian ini bertujuan untuk memperjelas konsep statistik dalam model *fitting*, uji kecocokan, dan validasi model dalam penelitian klinis. Hasil penelitian ini menunjukkan bahwa regresi logistik biner merupakan alat analisis yang penting dalam mengeksplorasi hubungan antara variabel prediktor dan hasil dikotomi dalam penelitian medis.

(Lestari et al., 2024) membandingkan algoritma C5.0 dan *Classification and Regression Trees* (CART) dalam mengklasifikasikan data *Stroke Prediction* yang berasal dari website Kaggle. Mereka menyoroti kesamaan kedua metode dalam hal kecepatan dan penanganan data, tetapi menunjukkan perbedaan dalam mekanisme klasifikasi. *Classification and Regression Trees* (CART) digunakan sebagai algoritma biner untuk mengklasifikasikan variabel respon numerik dan kategorikal, sementara C5.0 lebih fokus pada variabel respon kategorikal.

(Hairunnisa et al., 2023) meneliti penggunaan metode *Classification and Regression Trees* (CART) dalam klasifikasi penjurusan siswa baru. Penelitian ini menggunakan rasio data *training* dan *testing* sebesar 70:30 untuk memastikan validitas model. Mereka menyusun pohon keputusan berdasarkan berbagai indikator akademik yang relevan.

(Oyebanjo et al., 2023) menggunakan model *Classification and Regression Trees* (CART) untuk memprediksi bobot badan tiga ras kambing berdasarkan pengukuran morfologi. Penelitian ini mengidentifikasi prediktor paling signifikan menggunakan prosedur HPSPLIT di SAS. Metodologi yang mereka gunakan me-

nyoroti kompleksitas hubungan antara ciri morfologi dan bobot badan ternak.

Penelitian ini mengadaptasi pendekatan regresi logistik biner dan metode *Classification and Regression Trees* (CART) untuk menganalisis status peningkatan pendapatan usaha online di DIY. Berbeda dengan penelitian sebelumnya yang banyak berfokus pada bidang kesehatan dan klasifikasi data umum, penelitian ini menerapkan metode yang sama dalam konteks ekonomi. Dengan membandingkan dua metode tersebut, penelitian ini bertujuan untuk menentukan model prediksi terbaik dalam mengidentifikasi faktor-faktor yang memengaruhi status peningkatan pendapatan usaha online. Pendekatan ini memungkinkan pemanfaatan model statistik yang lebih optimal dalam pengambilan keputusan ekonomi berbasis data.



BAB V

PENUTUP

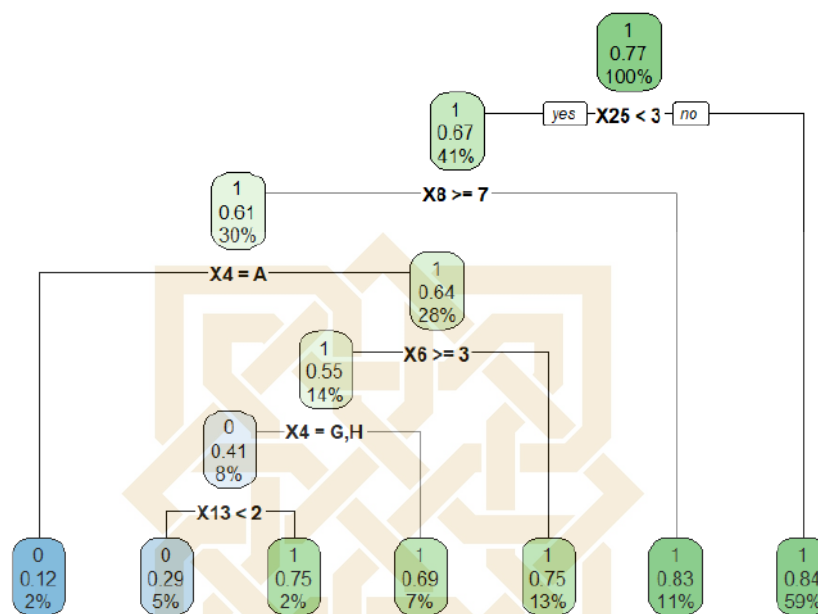
5.1. Kesimpulan

Berdasarkan hasil penelitian yang dilakukan, beberapa kesimpulan yang dapat diambil adalah sebagai berikut:

1. Model regresi logistik biner menunjukkan bahwa terdapat beberapa variabel yang berpengaruh signifikan terhadap status peningkatan pendapatan usaha online pada tahun 2020 terhadap tahun 2019 di DIY. Model regresi logistik biner yang terbentuk adalah sebagai berikut:

$$\begin{aligned} \log \left(\frac{\pi(x)}{1 - \pi(x)} \right) = & -2.597 + 1.554X_1 + 1.081X_{4(C)} + 1.239X_{4(G)} + 1.958X_{4(H)} \\ & + 1.305X_{4(I)} + 17.234X_{4(J)} + 16.972X_{4(M)} + 17.322X_{4(N)} \\ & + 1.210X_{4(P)} + 17.498X_{4(R)} + 1.666X_{4(S)} - 1.065X_8 \\ & + 1.671X_{18} + 1.348X_{25} \end{aligned}$$

Sementara, metode *Classification and Regression Trees* (CART) menghasilkan struktur pohon klasifikasi dengan 7 simpul terminal dan memiliki tingkat kedalaman sebanyak 6 tingkatan. Dimana variabel X_{25} (dampak pandemi terhadap distribusi barang) sebagai faktor utama (*root node*) dalam klasifikasi. Beberapa variabel penting lainnya yang muncul dalam pohon klasifikasi termasuk kategori usaha (X_4), durasi penjualan online (X_8), tahun operasional usaha (X_6) dan peran dalam e-commerce (X_{13}). Pohon yang terbentuk adalah:



Gambar 5.1 Pohon Klasifikasi Optimal

2. Berdasarkan hasil regresi logistik biner, variabel-variabel dengan pengaruh paling besar terhadap status peningkatan pendapatan usaha online pada tahun 2020 terhadap tahun 2019 di DIY dilihat dari nilai *odds ratio* adalah:

- Usia (X_1): 4,729
- Usaha industri pengolahan ($X_{4(C)}$): 2,948
- Usaha perdagangan besar dan eceran ($X_{4(G)}$): 3,453
- Usaha pengangkutan dan pergudangan ($X_{4(H)}$): 7,084
- Usaha akomodasi dan makan minum ($X_{4(I)}$): 3,687
- Usaha aktivitas jasa lainnya ($X_{4(S)}$): 5,294
- Jumlah bulan usaha melakukan penjualan online selama 2020 (X_8): 0,345
- Transaksi ke luar negeri (X_{18}): 5,319

- Dampak pandemi terhadap kelancaran distribusi barang selama pandemi (X_{25}): 3,851

Sedangkan, variabel yang berpengaruh besar dalam klasifikasi peningkatan pendapatan usaha online menggunakan metode *Classification and Regression Trees* (CART) adalah:

- Dampak pandemi COVID-19 terhadap kelancaran distribusi barang (X_{25})
- Jumlah bulan usaha melakukan penjualan online (X_8)
- Kategori usaha di bidang pertanian, kehutanan, dan perikanan ($X_{4(A)}$), bidang perdagangan besar dan eceran, reparasi dan perawatan mobil dan sepeda motor ($X_{4(G)}$), dan bidang pengangkutan dan pergudangan ($X_{4(H)}$)
- Tahun usaha mulai beroperasi (X_6)
- Peran dalam *e-commerce* (X_{13})

3. Berdasarkan hasil perbandingan ketepatan klasifikasi, regresi logistik biner menunjukkan performa yang lebih baik dibandingkan metode *Classification and Regression Trees* (CART). RLB memiliki ketepatan klasifikasi sebesar 76,29% dengan nilai *Apparent Error Rate* (APER) sebesar 23,71%, sedangkan *Classification and Regression Trees* (CART) memiliki ketepatan klasifikasi sebesar 72,16% dengan APER sebesar 27,84%. Perbedaan ini menunjukkan bahwa RLB memiliki keunggulan akurasi sebesar 4,13% dibandingkan dengan *Classification and Regression Trees* (CART)..
4. Keunggulan RLB dalam penelitian ini menunjukkan bahwa pola hubungan antara variabel prediktor dan status peningkatan pendapatan bersifat cukup linier, sehingga pendekatan parametrik seperti regresi logistik lebih sesuai.

Meski demikian, metode *Classification and Regression Trees* (CART) tetap memberikan insight tambahan melalui visualisasi pohon dan kemampuannya mendeteksi interaksi antar variabel tanpa asumsi distribusi.

Selisih akurasi yang relatif kecil antara kedua metode dapat disebabkan oleh kompleksitas data yang tidak terlalu tinggi, kemungkinan *overfitting*, serta pembagian data pelatihan dan pengujian. Namun demikian, berdasarkan hasil evaluasi, Regresi Logistik Biner merupakan metode yang lebih sesuai untuk data ini dalam mengklasifikasikan status peningkatan pendapatan usaha online pada tahun 2020 terhadap tahun 2019 di DIY.

5.2. Saran

Berdasarkan hasil penelitian, beberapa saran yang dapat diberikan untuk pengembangan lebih lanjut adalah sebagai berikut:

1. Penelitian ini hanya menggunakan data dari pelaku usaha online di D.I. Yogyakarta. Untuk penelitian selanjutnya, disarankan agar mencakup wilayah yang lebih luas agar hasil yang diperoleh lebih general dan dapat dibandingkan antar wilayah.
2. Variabel yang digunakan dalam penelitian ini terbatas pada variabel-variabel yang terdapat dalam data. Untuk penelitian selanjutnya, disarankan untuk menambahkan variabel lain seperti strategi pemasaran digital, jenis platform yang digunakan, dan tingkat kepuasan pelanggan guna memperluas wawasan tentang faktor-faktor yang memengaruhi peningkatan pendapatan usaha online.
3. Penelitian ini membandingkan dua metode klasifikasi, yaitu Regresi Logistik Biner dan *Classification and Regression Trees* (CART). Penelitian selanjutnya

jutnya dapat mencoba membandingkan regresi logistik biner dengan metode klasifikasi lain seperti Random Forest, Support Vector Machine (SVM), atau Neural Network untuk mengetahui metode mana yang memiliki kinerja klasifikasi terbaik dalam data status peningkatan pendapatan tersebut.

Dengan adanya penelitian ini, diharapkan hasil yang diperoleh dapat menjadi referensi bagi pelaku usaha online dalam menentukan strategi yang tepat guna meningkatkan pendapatan, serta membantu akademisi dalam memahami efektivitas metode klasifikasi dalam analisis data.

DAFTAR PUSTAKA

- Agresti, A. (2018). *An Introduction to Categorical Data Analysis*. Wiley Series in Probability and Statistics. John Wiley Sons, Inc., ISBN: 9781119405283, <https://books.google.co.id/books?id=pHZyDwAAQBAJ>.
- Arianty, K. (2024). Analisis peran e-commerce terhadap pendapatan umkm emak cw di masa pandemi dan pasca pandemi covid-19. *Jurnal Visualika*, 10:30–44, DOI: 10.56459/jv.v10i1.105.
- Badan Pusat Statistik (2025). Silastik - sistem layanan statistik terpadu. <https://silastik.bps.go.id/v3/index.php/mikrodata/index/>. Accessed: 2025-01-21.
- Berkson, J. (1944). Application of the logistic function to bio-assay. *Journal of the American Statistical Association*, 39:357–365, <https://api.semanticscholar.org/CorpusID:122893121>.
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Information Science and Statistics. Springer, New York, NY, 1 edition, ISBN: 978-0-387-31073-2. Springer-Verlag New York.
- Breiman, L. (2017). *Classification and Regression Trees*. CRC Press, ISBN: 9781351460484, <https://books.google.co.id/books?id=gLs6DwAAQBAJ>.
- Casella, G. & Berger, R. L. (2002). *Statistical Inference*. Duxbury, Pacific Grove, CA, 2 edition.

- Cox, D. R. (2018). The regression analysis of binary sequences. *Journal of the Royal Statistical Society: Series B (Methodological)*, 20(2):215–232, ISSN: 0035-9246, DOI: 10.1111/j.2517-6161.1958.tb00292.x, <https://doi.org/10.1111/j.2517-6161.1958.tb00292.x>.
- Dethan, N., Luan, F., & Obe, L. (2024). Binary logistic regression analysis in determining risk factors of acute respiratory infection (ira) on under-five children. *Sainmatika: Jurnal Ilmiah Matematika Dan Ilmu Pengetahuan Alam*, 21(1):8–15.
- Geron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*. O'Reilly Media, Sebastopol, CA, 2 edition.
- Ghozali, I. (2016). *Aplikasi analisis multivariete dengan program IBM SPSS 23*. Univ. Diponegoro Press.
- Gujarati, D. N. & Porter, D. C. (2012). *Basic Econometrics*. McGraw-Hill Education, 5th edition.
- Hairunnisa, D. A., Suhery, C., & Hidayati, R. (2023). Penerapan metode cart dalam klasifikasi jurusan siswa baru. *Journal of Computer System and Informatics (JoSYC)*, 4(4):1036–1045, DOI: 10.47065/josyc.v4i4.3860.
- Han, J., Kamber, M., & Pei, J. (2011). *Data mining: Concepts and techniques*. Morgan Kaufmann, ISBN: 978-0123814791.
- Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression*. John Wiley & Sons.
- Ibnu Katsir, I. b. 2000). *Tafsir al-Qur'an al-'Aim*, volume 4. Dar al-Fikr. Tafsir Surah Ar-Ra'd ayat 11.

- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning: With Applications in R*. Springer, New York, ISBN: 978-1461471370.
- Johnson, R. A. & Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis*. Pearson Prentice Hall, Upper Saddle River, NJ, 6 edition, ISBN: 0-13-187715-1.
- Joseph, V. R. (2022). Optimal ratio for data splitting. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 15(4):531–538.
- Kelleher, J. D., Mac Namee, B., & D'Arcy, A. (2015). *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies*. MIT Press, Cambridge, MA.
- Lestari, I., Fitria, D., Syafriandi, S., & Salma, A. (2024). Comparison of the c5.0 algorithm and the cart algorithm in stroke classification. *UNP Journal of Statistics and Data Science*, 2(1):90–98, DOI: 10.24036/ujsds/vol2-iss1/144.
- Lewis, R. J. (2000). An introduction to classification and regression tree (cart) analysis. In *Annual meeting of the society for academic emergency medicine in San Francisco, California*, volume 14. Citeseer.
- Madsen, H. & Thyregod, P. (2011). *An Introduction to General and Generalized Linear Models*. Chapman & Hall/CRC, Boca Raton, 1st edition, ISBN: 9781439894682.
- McCullagh, P. (1992). *Introduction to Nelder and Wedderburn (1972) Generalized Linear Models*, pages 543–546. Springer New York, New York, NY, ISBN:

978-1-4612-4380-9, DOI: 10.1007/978-1-4612-4380-9_38,
https://doi.org/10.1007/978-1-4612-4380-9_38.

Moehadi, M., Astuti, H., Fadilah, M. F., & Utomo, I. A. (2025). Kolaborasi bisnis, human capital dan digital marketing terhadap pertumbuhan usaha mikro kecil. *DIMENSI*, 14(1):213–237.

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to Linear Regression Analysis*. John Wiley & Sons, Hoboken, NJ, 5th edition.

Nugraha, J. (2014). *Pengantar Analisis Data Kategorik Metode dan Aplikasi menggunakan Program R ii*. deepublish, Yogyakarta, ISBN: 978-602-280-857-2.

Oyebanjo, M., Coker, O., & Osaiyuwu, O. (2023). Predicting the body weight of indigenous goat breeds from morphological measurements using the classification and regression tree (cart) data mining algorithm. *Biotechnology in Animal Husbandry*, 39(1):33–49, DOI: 10.2298/bah2301033o.

Prabandari, D. & Atmojo, M. (2022). Implementasi platform sibakul jogja dalam pemberdayaan umkm terdampak pandemi covid-19 di daerah istimewa yogyakarta. *TheJournalish: Social and Government*, 3:170–181, DOI: 10.55314/tsg.v3i3.270.

Rizaldi, R. & Djamaluddin, D. (2023). Optimalisasi legalitas dan pemasaran dalam mendukung pengembangan umkm. *Jurnal Inovasi dan Usaha*, 10(2):45–56.

Rohmah, D. A. M., Astuti, A. B., & Efendi, A. (2023). A statistical analytics of migration using binary bayesian logistic regression. *BA-REKENG: Jurnal Ilmu Matematika Dan Terapan*, 17(3):1725–1738, DOI: 10.30598/barekengvol17iss3pp1725-1738.

Srimaneekarn, N., Hayter, A., Liu, W., & Tantipoj, C. (2022). Binary response analysis using logistic regression in dentistry. In *International Journal of Dentistry*. Hindawi Limited, DOI: 10.1155/2022/5358602.

Sugiyono (2013). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Alfabeta, Bandung.

Tan, P.-N., Steinbach, M., & Kumar, V. (2006). *Introduction to Data Mining*. Pearson, 1st edition.

Timofeev, R. (2004). Classification and regression trees (cart) theory and applications. *Humboldt University, Berlin*, 54:48.

