

**PEMODELAN REGRESI LOGISTIK MULTINOMIAL UNTUK  
KETEPATAN WAKTU KELULUSAN MAHASISWA  
(Studi Kasus: Program Studi Matematika UIN  
Maulana Malik Ibrahim Malang)**

**SKRIPSI**

**OLEH  
MUHAMAD BAYU RAMDHANI  
NIM. 210601110027**



**PROGRAM STUDI MATEMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2025**

**PEMODELAN REGRESI LOGISTIK MULTINOMIAL UNTUK  
KETEPATAN WAKTU KELULUSAN MAHASISWA  
(Studi Kasus: Program Studi Matematika UIN  
Maulana Malik Ibrahim Malang)**

**SKRIPSI**

**Diajukan Kepada  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri Maulana Malik Ibrahim Malang  
untuk Memenuhi Salah Satu Persyaratan dalam  
Memperoleh Gelar Sarjana Matematika (S.Mat)**

**Oleh  
Muhamad Bayu Ramdhani  
NIM. 210601110027**

**PROGRAM STUDI MATEMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2025**

**PEMODELAN REGRESI LOGISTIK MULTINOMIAL UNTUK  
KETEPATAN WAKTU KELULUSAN MAHASISWA  
(Studi Kasus: Program Studi Matematika UIN  
Maulana Malik Ibrahim Malang)**

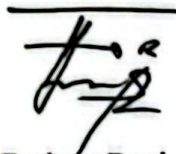
**SKRIPSI**

Oleh  
**Muhamad Bayu Ramdhani**  
NIM. 210601110027

Telah Disetujui untuk Diuji

Malang, 12 Desember 2025

Dosen Pembimbing I



Dr. Fachrur Rozi, M.Si.  
NIP. 19800527 200801 1 012

Dosen Pembimbing II



Erna Herawati, M.Pd.  
NIPPPK. 19760723 202321 2 006



Mengetahui,  
Ketua Program Studi Matematika

Dr. Fachrur Rozi, M.Si.  
NIP. 19800527 200801 1 012

**PEMODELAN REGRESI LOGISTIK MULTINOMIAL UNTUK  
KETEPATAN WAKTU KELULUSAN MAHASISWA  
(Studi Kasus: Program Studi Matematika UIN  
Maulana Malik Ibrahim Malang)**

**SKRIPSI**

**Oleh  
Muhamad Bayu Ramdhani  
NIM. 210601110027**

Telah Dipertahankan di Depan Dewan Penguji Skripsi  
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan  
untuk Memperoleh Gelar Sarjana Matematika (S.Mat)

Tanggal 22 Desember 2025

Ketua Penguji : Abdul Aziz, M.Si.

Anggota Penguji 1 : Ria Dhea Layla Nur Karisma, M.Si.

Anggota Penguji 2 : Dr. Fachrur Rozi, M.Si.

Anggota Penguji 3 : Erna Herawati, M.Pd.



Mengetahui,  
Ketua Program Studi Matematika



Dr. Fachrur Rozi, M.Si.  
NIP. 19800527 200801 1 012

## PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan dibawah ini:

Nama : Muhamad Bayu Ramdhani

NIM : 210601110027

Program Studi : Matematika

Judul Skripsi : Pemodelan Regresi Logistik Multinomial Untuk Ketepatan Waktu Kelulusan Mahasiswa (Studi Kasus: Program Studi Matematika UIN Maulana Malik Ibrahim Malang)

Menyatakan dengan sebenarnya bahwa skripsi yang saya tulis benar-benar merupakan hasil karya sendiri, bukan merupakan pengambilan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan dan pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka. Apabila di kemudian hari terbukti atau dapat dibuktikan skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 22 Desember 2025

Yang membuat pernyataan,



Muhamad Bayu Ramdhani

NIM. 210601110027

## **MOTO**

*“The truth is always simple and honest,  
while a lie is always complicated and convoluted”*

- Leo Tolstoy

## **PERSEMBAHAN**

Skripsi ini penulis persembahkan dengan penuh rasa terima kasih kepada: Keluarga tercinta, khususnya kepada Bapak Muhammad Ali dan Ibu Nia Sriyati, yang dengan doa, kasih sayang, serta nasihat yang bijaksana selalu menjadi sumber kekuatan saya dalam menempuh pendidikan ini. Kepada saudara-saudara saya yang selalu memberi dukungan dan motivasi dalam setiap proses yang saya jalani. Ucapan terima kasih saya sampaikan kepada sahabat-sahabat yang senantiasa membantu, dan memberi semangat ketika saya menghadapi kesulitan. Semoga segala doa dan dukungan tersebut menjadi pendorong bagi saya dalam meraih masa depan yang saya impikan.

## KATA PENGANTAR

*Assalamu 'alaikum Warahmatullahi Wabarakatuh.*

Segala puji bagi Allah SWT atas limpahan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Pemodelan Regresi Logistik Multinomial Untuk Ketepatan Waktu Kelulusan Mahasiswa (Studi Kasus: Program Studi Matematika UIN Maulana Malik Ibrahim Malang)”. Skripsi ini disusun sebagai salah satu syarat kelulusan dan memperoleh gelar Sarjana dalam bidang Matematika di Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.

Penulis menyadari bahwa selama penyusunan skripsi ini tidak terlepas dari bantuan, bimbingan, dan dukungan dari berbagai pihak. Oleh karena itu, penulis ingin memberikan ucapan terimakasih kepada:

1. Prof. Dr. Hj. Ilfi Nurdiana, M.Si., CAHRM., CRMP., selaku Rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Dr. Agus Mulyono, M.Kes., selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Dr. Fachrur Rozi, M.Si., selaku Ketua Program Studi Matematika sekaligus Dosen Pembimbing I, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Terimakasih telah meluangkan waktu, tenaga, dan pikiran dalam memberikan saran, arahan, nasihat, serta motivasi selama proses penyusunan skripsi.
4. Erna Herawati, M.Pd., selaku Dosen Pembimbing II yang telah meluangkan waktu, tenaga, dan pikiran untuk memberikan saran, arahan, dan nasihat selama proses penyusunan skripsi.
5. Abdul Aziz, M.Si., selaku Ketua Penguji dalam Ujian Skripsi, yang telah memberikan saran, arahan, dan nasihat yang bermanfaat bagi penulis.
6. Ria Dhea Layla Nur Karisma, M.Si., selaku Anggota Penguji I dalam Ujian Skripsi, yang telah memberikan saran, arahan, dan nasihat yang bermanfaat bagi penulis.
7. Ari Kusumastuti, M.Pd., M.Si., selaku Dosen Wali yang senantiasa memberikan bimbingan, arahan, dan ilmu yang bermanfaat.



8. Seluruh Dosen Program Studi Matematika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim.
9. Kedua orang tua tercinta, Bapak Muhamad Ali dan Ibu Nia Sriyati, atas doa, dukungan moral dan dukungan materi, serta kasih sayang yang tidak pernah terputus kepada penulis.
10. Muhamad Resno Setiadi, dan Muhamad Novianto Isra Rahmani selaku kakak-kakak penulis yang selalu memberikan dukungan dan doa kepada penulis.
11. Seluruh teman-teman seperjuangan di Program Studi Matematika angkatan 2021 khususnya teman-teman konsorsium statistika yang telah memberikan bantuan serta semangat.
12. Semua pihak yang telah memberikan bantuan moril maupun materil kepada penulis dalam menyelesaikan skripsi ini yang tidak dapat penulis sebutkan satu per satu.

Semoga Allah SWT melimpahkan taufik dan hidayah-Nya kepada kita semua. Penulis menyadari masih terdapat kekurangan dalam penyusunan draf skripsi ini. Penulis berharap skripsi ini dapat memberikan manfaat bagi penulis dan pembaca serta dapat menjadi sumbangsih bagi perkembangan ilmu pengetahuan, khususnya di bidang Matematika.

*Wassalamu'alaikum Warahmatullahi Wabarakatuh.*

Malang, 22 Desember 2025

Penulis

## DAFTAR ISI

HALAMAN JUDUL .....	i
HALAMAN PENGANTAR .....	ii
HALAMAN PERSETUJUAN .....	iii
HALAMAN PENGESAHAN.....	iv
PERNYATAAN KEASLIAN TULISAN .....	iv
MOTO .....	vi
PERSEMBAHAN.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR TABEL .....	xii
DAFTAR GAMBAR.....	xiii
DAFTAR SIMBOL .....	xiv
DAFTAR LAMPIRAN .....	xvi
ABSTRAK .....	xvii
ABSTRACT .....	xviii
مستخلصا البحث.....	xix
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	8
1.3 Tujuan Penelitian .....	8
1.4 Manfaat Penelitian .....	9
1.5 Batasan Masalah .....	10
<b>BAB II KAJIAN TEORI .....</b>	<b>12</b>
2.1 Teori Pendukung.....	12
2.1.1 Standardisasi .....	12
2.1.2 Uji Multikolinearitas .....	13
2.1.3 Model Regresi Logistik Multinomial.....	14
2.1.4 Evaluasi Kinerja Model dan <i>Confusion Matrix</i> .....	23
2.1.5 Ketepatan Waktu Kelulusan Mahasiswa.....	26
2.2 Kajian Integrasi Ketepatan Waktu Kelulusan dengan Al-Quran / Hadits.....	30
2.3 Kajian Ketepatan Waktu Kelulusan dengan Teori Pendukung .....	32
<b>BAB III METODE PENELITIAN .....</b>	<b>37</b>
3.1 Jenis Penelitian .....	37
3.2 Data dan Sumber Data .....	37
3.3 Teknik Analisis Data .....	38
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>	<b>46</b>
4.1 Persiapan Data .....	46
4.2 Statistika Deskriptif .....	47
4.2.1 Deskripsi Variabel Prediktor Numerik.....	48
4.2.2 Deskripsi Variabel Prediktor Kategorik.....	50
4.3 Uji Multikolinearitas.....	56
4.4 Standardisasi .....	57
4.5 Pembagian Data <i>Training</i> dan <i>Testing</i> .....	58

4.6	Membangun Model Regresi Logistik Multinomial .....	59
4.7	Estimasi Parameter Model Regresi Logistik Multinomial .....	61
4.8	Uji Signifikansi Parameter .....	62
4.8.1	Uji Signifikansi Parameter Secara Simultan .....	62
4.8.2	Uji Signifikansi Parameter Secara Parsial.....	63
4.9	Uji Kesesuaian Model.....	69
4.10	Pengukuran Kebaikan Model .....	70
4.11	Evaluasi Kinerja Model .....	70
4.11.1	Metrik Kinerja Keseluruhan.....	71
4.11.2	Metrik Kinerja Per-kategori .....	73
4.12	Interpretasi Model.....	77
4.12.1	Interpretasi Model Logit 1 (Terlambat vs Tepat Waktu) .....	77
4.12.2	Interpretasi Model Logit 2 (Kritis vs Tepat Waktu) .....	79
4.13	Simulasi Prediksi pada Studi Kasus .....	81
4.14	Kajian Penelitian dalam Perspektif Islam.....	84
<b>BAB V</b>	<b>PENUTUP .....</b>	<b>89</b>
5.1	Kesimpulan .....	89
5.2	Saran .....	91
<b>DAFTAR PUSTAKA .....</b>		<b>92</b>
<b>LAMPIRAN.....</b>		<b>95</b>
<b>RIWAYAT HIDUP .....</b>		<b>107</b>

## DAFTAR TABEL

Tabel 3.1	Variabel Penelitian.....	38
Tabel 4.1	Variabel Prediktor Numerik.....	48
Tabel 4.2	Tabulasi Silang dari Variabel Prediktor Kategorik terhadap Ketepatan Waktu Kelulusan .....	53
Tabel 4.3	Nilai VIF .....	56
Tabel 4.4	Pembagian Data <i>Training</i> dan <i>Testing</i> .....	58
Tabel 4.5	Estimasi Parameter.....	61
Tabel 4.6	Uji Signifikansi Simultan.....	63
Tabel 4.7	Uji Signifikansi Parsial .....	65
Tabel 4.8	Uji Signifikansi Simultan Model Revisi .....	66
Tabel 4.9	Uji Signifikansi Parsial Model Revisi.....	67
Tabel 4.10	Uji Kesesuaian Model.....	69
Tabel 4.11	Pseudo R-Square .....	70
Tabel 4.12	<i>Confusion Matrix</i> .....	71
Tabel 4.13	<i>Multiclass Confusion Matrix</i> untuk Kategori Tepat Waktu.....	73
Tabel 4.14	<i>Multiclass Confusion Matrix</i> untuk Kategori Terlambat .....	74
Tabel 4.15	<i>Multiclass Confusion Matrix</i> untuk Kategori Kritis .....	75
Tabel 4.16	Nilai <i>Odds Ratio</i> Logit 1 .....	78
Tabel 4.17	Nilai <i>Odds Ratio</i> Logit 2.....	80

## DAFTAR GAMBAR

Gambar 3.1 <i>Flowchart</i> Penelitian .....	45
Gambar 4.1 Ketepatan Waktu Kelulusan.....	47
Gambar 4.2 Variabel Prediktor Kategorik .....	51

## DAFTAR SIMBOL

$Y$	: Variabel respon Ketepatan Waktu Kelulusan
$x_1$	: Variabel IPK
$x_2$	: Variabel Jenis Kelamin
$x_3$	: Variabel Jenis Asal Sekolah
$x_4$	: Variabel Jenis Pembiayaan
$x_5$	: Variabel Jalur Masuk
$x_6$	: Variabel Asal Daerah
$x_7$	: Variabel Rasio SKS Lulus
$X_i$	: Nilai awal data ke- $i$
$\bar{X}$	: Nilai rata-rata (Mean)
$S$	: Simpangan baku (Standar Deviasi) sampel
$Z_i$	: Nilai data yang telah distandardisasi untuk observasi ke- $i$
$(VIF)_k$	: <i>Variance Inflation Factor</i> untuk variabel prediktor ke- $k$
$R_k^2$	: Koefisien determinasi ganda dari variabel prediktor $X_k$ dengan seluruh prediktor lainnya
$p$	: Banyaknya variabel prediktor dalam model
$n$	: Jumlah total observasi
$Y_{ji}$	: Variabel biner yang dikodekan (0 atau 1) untuk menunjukkan keanggotaan kelompok kategori $j$ pada observasi $i$
$x'$	: Vektor variabel prediktor (kovariat)
$x_0$	: Suku konstanta ( <i>intercept</i> ), di mana $x_0 = 1$
$x_{ki}$	: Nilai variabel prediktor ke- $k$ untuk observasi ke- $i$
$\pi_j(x)$	: Probabilitas bersyarat untuk kategori respons ke- $j$ berdasarkan vektor kovariat $x$
$g(x)$	: Fungsi Logit
$g_j(x)$	: Fungsi Logit yang membandingkan kategori $Y = j$ dengan kategori referensi $Y = 0$
$\beta$	: Vektor semua estimasi parameter regresi (koefisien)
$\beta_i$	: Estimasi koefisien regresi logistik untuk variabel prediktor ke- $i$
$\ln$	: Logaritma natural
$L(\beta)$	: Fungsi <i>log-likelihood</i>
$l(b)$	: Fungsi <i>likelihood</i>
$\hat{\beta}$	: <i>Maximum Likelihood Estimator</i> (MLE)
$\widehat{SE}(\hat{\beta}_i)$	: Taksiran <i>Standard Error</i> dari estimasi parameter
$G^2$	: Statistik uji <i>Likelihood Ratio Test</i> (Uji G)
$W$	: Statistik uji Wald
$\chi^2$	: Distribusi <i>Chi-square</i>
$\chi^2_{(v,\alpha)}$	: Nilai kritis Distribusi <i>Chi-square</i>
$v$	: Derajat bebas ( <i>degrees of freedom</i> )
$\hat{C}$	: Statistik uji Hosmer and Lemeshow ( <i>Goodness of Fit</i> )
$O_k$	: Observasi pada kategori ke- $k$ untuk uji kesesuaian model

$\bar{\pi}_k$	: Rata-rata taksiran peluang ke- $k$ untuk uji kesesuaian model
$R_N^2$	: Nilai Pseudo R-Square Nagelkerke
$R_{CS}^2$	: Nilai Pseudo R-Square Cox & Snell
$LL(M_{awal})$	: Nilai log-likelihood dari model awal (hanya <i>intercept</i> )
$OR$	: <i>Odds Ratio</i> (Rasio Peluang)
$OR_j(a, b)$	: <i>Odds Ratio</i> yang membandingkan kategori $Y = j$ terhadap $Y = 0$ pada $x = a$ vs $x = b$
$TP$	: True Positive (Prediksi positif yang benar)
$FP_k$	: False Positive untuk kelas ke- $k$
$FN_k$	: False Negative untuk kelas ke- $k$
<i>Accuracy</i>	: Akurasi model keseluruhan
$Precision_k$	: Presisi untuk kelas ke- $k$
$Recall_k$	: Recall (True Positive Rate) untuk kelas ke- $k$
$F1-Score_k$	: F1-Score untuk kelas ke- $k$

## DAFTAR LAMPIRAN

Lampiran 1. Data dan <i>Coding</i> .....	95
Lampiran 2. <i>Output</i> untuk Uji Multikolinearitas .....	99
Lampiran 3. <i>Output</i> untuk Estimasi Parameter, Uji Parsial .....	100
Lampiran 4. <i>Output</i> untuk Estimasi Parameter Model Revisi, Uji Parsial, dan <i>Odds Ratio</i> .....	101
Lampiran 5. <i>Output</i> untuk Uji Signifikansi Parameter Secara Simultan .....	102
Lampiran 6. <i>Output</i> untuk Uji Signifikansi Parameter Secara Simultan Model Revisi .....	103
Lampiran 7. <i>Output</i> untuk Uji Kesesuaian Model .....	104
Lampiran 8. <i>Output</i> untuk Pengukuran Kebaikan Model .....	105
Lampiran 9. <i>Output</i> untuk Evaluasi Kinerja Model .....	106



## ABSTRAK

Ramdhani, Muhamad Bayu. 2025. **Pemodelan Regresi Logistik Multinomial Untuk Ketepatan Waktu Kelulusan Mahasiswa (Studi Kasus: Program Studi Matematika UIN Maulana Malik Ibrahim Malang)**. Skripsi. Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Dr. Fachrur Rozi, M.Si. (II) Erna Herawati, M.Pd.

**Kata Kunci:** Regresi Logistik Multinomial; Ketepatan Waktu Kelulusan; Klasifikasi; Prediksi; *Confusion Matrix*.

Ketepatan waktu kelulusan mahasiswa merupakan indikator penting dalam menilai efektivitas proses akademik dan pencapaian mutu program studi. Oleh karena itu, penelitian ini membangun model prediksi serta menganalisis faktor-faktor yang memengaruhi ketepatan waktu kelulusan mahasiswa Program Studi Matematika UIN Maulana Malik Ibrahim Malang menggunakan regresi logistik multinomial. Variabel dependen diklasifikasikan ke dalam kategori Tepat Waktu, Terlambat, dan Kritis, sedangkan tujuh variabel independen meliputi IPK, jenis kelamin, jenis asal sekolah, jenis pembiayaan, jalur masuk, asal daerah, dan rasio SKS lulus. Sebanyak 331 data mahasiswa diproses melalui pembersihan data, pembentukan dummy, pengujian asumsi, standarisasi variabel numerik, pembagian data *training* dan *testing* dengan *stratified sampling*, serta estimasi parameter menggunakan *Maximum Likelihood Estimator*. Hasil analisis menunjukkan bahwa pada Logit 1 (Terlambat vs Tepat Waktu) hanya IPK yang signifikan, sedangkan rasio SKS lulus tidak signifikan. Pada Logit 2 (Kritis vs Tepat Waktu), IPK dan rasio SKS lulus signifikan dan keduanya menurunkan peluang mahasiswa masuk kategori Kritis. Evaluasi model menunjukkan akurasi 64,18%, dengan F1-Score per kategori sebesar 75,55% (Tepat Waktu), 42,10% (Terlambat), dan 28,57% (Kritis). Dengan demikian, model yang dihasilkan dapat digunakan sebagai alat prediksi awal untuk mengidentifikasi potensi keterlambatan studi serta mendukung perumusan kebijakan akademik.

## ABSTRACT

Ramdhani, Muhamad Bayu. 2025. **Multinomial Logistic Regression Modeling for Student Graduation Timeliness (Case Study: Mathematics Study Program, UIN Maulana Malik Ibrahim Malang)**. Undergraduate Thesis. Department of Mathematics, Faculty of Science and Technology, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Advisors: (I) Dr. Fachrur Rozi, M.Si. (II) Erna Herawati, M.Pd.

**Keywords:** Multinomial Logistic Regression; Graduation Timeliness; Classification; Prediction; Confusion Matrix.

Timely graduation is an important indicator for assessing the effectiveness of academic processes and the overall quality of academic programs. Therefore, this study aims to develop a predictive model and analyze the factors influencing the graduation timeliness of students in the Mathematics Study Program at UIN Maulana Malik Ibrahim Malang using multinomial logistic regression. The dependent variable is categorized into On Time, Late, and Critical, while the seven independent variables include GPA, gender, type of prior school, type of funding, admission pathway, region of origin, and ratio of completed credits. A total of 331 student data were processed through data cleaning, dummy variable construction, assumption testing, numerical variable standardization, stratified sampling for training–testing split, and parameter estimation using the Maximum Likelihood Estimator. The results show that in Logit 1 (Late vs On Time), only GPA is significant, while the ratio of completed credits is not significant. In Logit 2 (Critical vs On Time), both GPA and the ratio of completed credits are significant and negatively associated with the likelihood of students entering the Critical category. Model evaluation produced an accuracy of 64.18%, with F1-Scores of 75.55% (On Time), 42.10% (Late), and 28.57% (Critical). Thus, the developed model can serve as an early prediction tool to identify potential delays in student graduation and support academic policy formulation.

## مستخلصا البحث

رمضاني، محمد بايو. ٢٠٢٥. نمذجة الانحدار اللوجستي المتعدد للفئات لتحديد دقة التخرج في الوقت المناسب للطلبة (دراسة حالة: قسم الرياضيات في جامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانج). البحث الجامعي. قسم الرياضيات، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانج. المشرف: (١) الدكتور. فخر الرّازي، الماجستير في العلوم. (٢) إيرنا هراواتي، الماجستير في التربية.

**الكلمات الأساسية:** الانحدار اللوجستي المتعدد للفئات؛ دقة التخرج؛ التصنيف؛ التنبؤ؛ مصفوفة الالتباس.

تُعدّ دقة التخرج في الوقت المحدد مؤشرًا مهمًا في تقييم فاعلية العملية الأكاديمية وجودة البرامج الدراسية. وبناءً على ذلك، تهدف هذه الدراسة إلى بناء نموذج تنبؤي وتحليل العوامل المؤثرة في دقة التخرج في الوقت المناسب لطلبة برنامج دراسة الرياضيات في جامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانج باستخدام الانحدار اللوجستي المتعدد للفئات. قُيِّم المتغير التابع إلى ثلاث فئات: التخرج في الوقت المناسب، التأخر، والحالة الحرجة، بينما اشتملت المتغيرات المستقلة السبعة على المعدل التراكمي، الجنس، نوع المدرسة السابقة، نوع التمويل، مسار القبول، منطقة السكن، ونسبة الساعات الدراسية المجتازة. شملت البيانات ٣٣١ طالبًا، وتمت معالجتها من خلال تنظيف البيانات، وبناء المتغيرات الوهمية، واختبار الافتراضات، وتوحيد المتغيرات العددية، وتقسيم البيانات إلى تدريب واختبار باستخدام أخذ العينات الطبقيّة، إضافةً إلى تقدير المعلمات باستخدام طريقة الاحتمالية العظمى. أظهرت نتائج التحليل أن المعدل التراكمي كان المتغير الوحيد المؤثر في الفئة المتأخرة مقارنةً بالفئة في الوقت المناسب، بينما لم تكن نسبة الساعات المجتازة ذات دلالة إحصائية. أمّا في الفئة الحرجة، فقد كان كلٌّ من المعدل التراكمي ونسبة الساعات المجتازة ذا دلالة إحصائية وبمعامل سالب، مما يدل على أن ارتفاع هذين المتغيرين يقلل من احتمال دخول الطلبة في الفئة الحرجة. كما أظهر تقييم النموذج دقة مقدارها ٦٤,١٨٪، مع قيم  $F1$  لفئات النموذج بلغت ٧٥,٥٥٪ للفئة في الوقت المناسب، و ٤٢,١٠٪ للفئة المتأخرة، و ٢٨,٥٧٪ للفئة الحرجة. وبذلك يمكن اعتماد النموذج كأداة أولية للتنبؤ باحتمالية تأخر التخرج ودعم تطوير السياسات الأكاديمية.

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Model logistik atau logit adalah nama lain dari regresi logistik dalam statistik. Dengan menggunakan fungsi logit dari kurva logistik, model ini menghitung kemungkinan suatu peristiwa akan terjadi. Cara lain untuk memikirkan regresi logistik adalah sebagai teknik untuk membuat model prediksi. Para peneliti menggunakan regresi logistik untuk memprediksi variabel dependen dikotomis. Data nominal dengan dua kategori, seperti ya atau tidak, baik atau buruk, dan tinggi atau rendah, disebut sebagai variabel dikotomi. Selain itu, karena regresi logistik menggunakan distribusi logistik, maka tidak diperlukan asumsi bahwa varians kesalahan (residual) memiliki distribusi normal (Varamita, 2017).

Guna memodelkan koneksi fungsional antar variabel, digunakan sebuah metode statistik yang dikenal sebagai analisis regresi, di mana metode ini memberikan penekanan utama pada identifikasi adanya hubungan kausalitas. Metode ini mampu mengakomodasi berbagai jenis data, baik yang bersifat kategorik maupun numerik. Dalam konteks penelitian ini, dipilih varian regresi logistik multinomial sebagai alat analisis utama mengingat karakteristik data yang dimiliki.

Regresi logistik multinomial adalah metode analisis statistik yang digunakan untuk menjelaskan hubungan antara variabel independen ( $x$ ) dan variabel dependen ( $y$ ) yang memiliki lebih dari dua kategori (Agresti, 2002). Analisis ini menggunakan model logit untuk memperkirakan probabilitas terjadinya suatu

kategori berdasarkan nilai variabel independen. Menurut Hosmer & Lemeshow (2000), variabel dependen ( $y$ ) dalam analisis ini memiliki lebih dari dua kelompok, yang pada umumnya diberi kode numerik seperti 0, 1, atau 2. Dibandingkan dengan jenis regresi lainnya, regresi logistik multinomial memberikan berbagai keunggulan, seperti kemampuan untuk memperkirakan probabilitas suatu kategori, mengelola variabel dependen berskala multinomial, mengidentifikasi hubungan antara variabel independen dan dependen, serta menggunakan variabel dependen berskala nominal dengan pendekatan yang lebih akurat.

Ketika dihadapkan pada variabel respon yang sifatnya nominal dan memiliki lebih dari dua kategori, serta variabel prediktor yang bisa berbentuk kontinu maupun kategorik, maka metode analisis yang paling sesuai untuk digunakan adalah regresi logistik multinomial (Hosmer & Lemeshow, 2000). Metode ini tepat digunakan untuk kasus seperti pemodelan ketepatan waktu kelulusan mahasiswa, di mana variabel dependennya berskala nominal dengan klasifikasi tidak berjenjang. Dalam penelitian ini, ketepatan waktu kelulusan dikategorikan menjadi tiga kelompok diskrit: Tepat waktu untuk mahasiswa yang menyelesaikan studi dalam waktu 7-8 semester, Terlambat untuk mahasiswa yang menyelesaikan studi dalam waktu 9-10 semester, dan Kritis untuk mahasiswa yang menyelesaikan studi dalam waktu 11 semester atau lebih. Kategori ini bersifat diskrit dan tidak menunjukkan hubungan tingkatan (non-ordinal), namun memenuhi kriteria saling lepas (*mutually exclusive*) yang menjadi syarat utama dalam analisis multinomial.

Penetapan kategori ketepatan waktu kelulusan menjadi 'Tepat Waktu' (7-8 semester), 'Terlambat' (9-10 semester), dan 'Kritis' (11 semester atau lebih) merupakan sebuah definisi operasional yang dirumuskan untuk keperluan analisis

dalam penelitian ini. Kategori 'Tepat Waktu' didasarkan pada standar nasional yang berlaku. Sementara itu, pembagian kategori 'Terlambat' dan 'Kritis' didasarkan pada sistem klasifikasi dan evaluasi masa studi yang ditetapkan oleh Universitas (Siakad UIN Malang). Sistem tersebut secara internal mengklasifikasikan mahasiswa yang melewati batas studi normal ke dalam beberapa tingkatan monitoring, seperti 'Waspada' untuk masa studi 9-10 semester dan 'Kritis' untuk 11-14 semester.

Perguruan tinggi memegang peranan krusial dalam mencetak sumber daya manusia yang unggul di tengah persaingan global yang semakin kompetitif. Dengan demikian, menjadikan peningkatan kualitas dan pencapaian akreditasi sebagai sebuah keharusan bagi institusi perguruan tinggi. Berdasarkan penilaian BAN-PT (2019), mahasiswa merupakan salah satu komponen penentu dalam evaluasi mutu sekaligus status akreditasi dari sebuah institusi. Indikator-indikator terkait mahasiswa meliputi durasi studi pada setiap program studi serta persentase kelulusan tepat waktu di seluruh program studi yang ada. Konsekuensinya, sebagai entitas penyelenggara pendidikan, setiap perguruan tinggi harus menjadikan tingkat kelulusan tepat waktu mahasiswanya sebagai prioritas utama, guna mendongkrak penilaian kualitas dan akreditasi institusi.

Indikator kelulusan tepat waktu bagi mahasiswa ditentukan berdasarkan jenjang program studi yang ditempuh di perguruan tinggi. Sebagai ilustrasi, Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang sebagaimana tercantum dalam "Pedoman Pendidikan 2019 UIN Malang" institusi tersebut menyelenggarakan program akademik pada jenjang Sarjana (S1), Magister (S2), dan Doktor (S3). Merujuk pada standar yang ditetapkan oleh Kementerian Riset, Teknologi, dan Pendidikan Tinggi, Direktorat Jenderal Pembelajaran dan

Kemahasiswaan (2013), batas waktu penyelesaian studi untuk kategori lulus tepat waktu pada jenjang S1, S2, dan S3 masing-masing adalah maksimal empat tahun, dua tahun, dan tiga tahun. Sebaliknya, mahasiswa yang menyelesaikan studi melebihi batas waktu tersebut dikategorikan sebagai tidak lulus tepat waktu. Selain durasi studi, persyaratan akademik berupa beban studi minimum juga menjadi penentu kelulusan, yaitu pencapaian minimal 144 SKS (Satuan Kredit Semester) untuk S1, 36 SKS untuk S2, dan 42 SKS untuk S3 (Tim UIN Maulana Malik Ibrahim Malang, 2019).

UIN Malang menyelenggarakan berbagai program studi jenjang sarjana, termasuk di dalamnya Program Studi Matematika. Berdasarkan data statistik dari Badan Administrasi Akademik UIN Malang, tercatat bahwa tingkat kelulusan tepat waktu pada program studi ini masih berada pada angka yang relatif rendah. Dari total 336 mahasiswa angkatan 2018-2021, hanya 192 individu (57,14%) yang berhasil menyelesaikan studi dalam waktu yang telah ditetapkan yakni 7-8 semester. Rendahnya persentase kelulusan tepat waktu ini menjadi perhatian serius mengingat dampaknya terhadap penilaian kualitas akademik dan proses akreditasi institusi pendidikan tinggi. Kondisi demikian mengindikasikan adanya kebutuhan untuk melakukan evaluasi mendalam terhadap faktor-faktor yang memengaruhi durasi studi mahasiswa.

Ketepatan waktu kelulusan mahasiswa merupakan aspek penting dalam penyelesaian studi, mengingat institusi pendidikan telah menetapkan batas waktu yang harus dipenuhi. Fenomena ini selaras dengan prinsip ketepatan waktu yang diajarkan dalam Islam, sebagaimana termaktub dalam QS. An-Nisa ayat 103:

فَإِذَا قَضَيْتُمُ الصَّلَاةَ فَادْكُرُوا اللَّهَ قِيَامًا وَقُعُودًا وَعَلَىٰ جُنُوبِكُمْ فَإِذَا اطْمَأْنَنْتُمْ فَأَقِيمُوا الصَّلَاةَ إِنَّ الصَّلَاةَ  
كَانَتْ عَلَى الْمُؤْمِنِينَ كِتَابًا مَّوْقُوتًا ﴿١٠٣﴾

“Apabila kamu telah menyelesaikan salat, berzikirlah kepada Allah (mengingat dan menyebut-Nya), baik ketika kamu berdiri, duduk, maupun berbaring. Apabila kamu telah merasa aman, laksanakanlah salat itu (dengan sempurna). Sesungguhnya salat itu merupakan kewajiban yang waktunya telah ditentukan atas orang-orang mukmin.” (QS. An-Nisa: 103)

Dalam tafsir Al-Misbah, Quraish Shihab menjelaskan bahwa QS. An-Nisa ayat 103 menegaskan kewajiban salat sebagai ibadah yang terikat dengan waktu (*mauqūt*). Penekanan pada frasa *kitābam mauqūtā* (kewajiban yang ditentukan waktunya) menunjukkan bahwa salat hanya sah apabila dilaksanakan dalam batas waktu syar’i yang telah ditetapkan. Shihab menjelaskan bahwa ketepatan waktu dalam ibadah mencerminkan kedisiplinan dan kepatuhan kepada Allah, sekaligus menjadi simbol keteraturan dalam Islam. Lebih jauh, ia menghubungkan prinsip ini dengan sunnatullah yang mengatur segala sesuatu secara teratur, termasuk fase kehidupan manusia (Shihab, 2002). Dalam konteks akademik, prinsip ini dapat dianalogikan dengan sistem kelulusan mahasiswa. Mahasiswa yang menyelesaikan studi tepat waktu mencerminkan ketaatan terhadap aturan institusional, sebagaimana Muslim taat pada waktu salat. Keterlambatan kelulusan, menurut perspektif ini, dapat dipandang sebagai pelanggaran terhadap prinsip disiplin waktu yang Islami, karena menghambat kontribusi sosial dan profesional sesuai timeline yang telah ditetapkan.

Pada penelitian terdahulu mengenai ketepatan waktu kelulusan, Muhendra (2021) terkait prediksi tingkat kelulusan mahasiswa teknik industri Universitas Bhayangkara Jakarta Raya dengan menggunakan metode *Decision Tree* (DT). Penelitian tersebut memanfaatkan tujuh variabel, yaitu jenis kelamin, jumlah SKS



lulus, predikat, kriteria usia masuk kuliah, bidang penelitian, dan kepemilikan sertifikat. Hasil penelitian tersebut menunjukkan tingkat akurasi model lebih dari 80%. Selanjutnya, Mashfia (2022) terkait prediksi waktu kelulusan mahasiswa menggunakan metode *Naive Bayes Classifier* di Program Studi Matematika UIN Malang dengan lima variabel prediktor, yaitu IPK, jenis kelamin, jenis sekolah, jenis pembiayaan, dan jalur masuk. Hasil penelitian tersebut menunjukkan tingkat akurasi sebesar 90% dengan tingkat kesalahan (APER) sebesar 10%. Selanjutnya, Syahranita (2022) terkait prediksi kategori kelulusan mahasiswa menggunakan metode regresi logistik multinomial di Program Studi Teknik Informatika UIN Malang dengan tiga variabel prediktor (IP, jenis kelamin, dan jalur masuk). Temuan penelitian ini menghasilkan nilai akurasi 85,5% dengan *precision* 78,5%, *recall* 93,9%, dan *micro f1-score* 89,8%. Selanjutnya, Fadilah (2023) terkait optimasi *k-nearest neighbors* menggunakan *fuzzy c-means* pada ketepatan waktu kelulusan berdasarkan tiga variabel prediktor (IPS, jenis kelamin, dan jenis pembiayaan). Hasil penelitian tersebut menunjukkan bahwa model dasar tanpa optimasi memiliki akurasi sebesar 73,65%, sedangkan model prediksi terbaik dengan penambahan optimasi menggunakan *Fuzzy c-Means*, menghasilkan akurasi 74,8%.

Berdasarkan tinjauan terhadap penelitian terdahulu oleh Muhendra (2021), Mashfia (2022), Syahranita (2022), dan Fadilah (2023), terdapat celah penelitian yang signifikan. Secara spesifik, belum ada penelitian yang menerapkan metode Regresi Logistik Multinomial (RLM) pada objek Program Studi Matematika UIN Malang. Penelitian sebelumnya pada objek yang sama menggunakan metode lain seperti *Naive Bayes Classifier* (Mashfia, 2022), sementara penelitian yang menggunakan RLM dilakukan pada program studi yang berbeda (Syahranita,

2022). Selain itu, model-model yang ada cenderung menggunakan set variabel prediktor yang lebih terbatas. Oleh karena itu, kebaruan utama dari penelitian ini terletak pada penerapan metode Regresi Logistik Multinomial dengan menggunakan set variabel yang lebih komprehensif, mencakup tujuh prediktor, termasuk faktor akademik baru seperti Rasio SKS Lulus dan faktor demografis Asal Daerah, pada konteks spesifik Program Studi Matematika UIN Malang. Pendekatan ini tidak hanya mengisi celah metodologis dan kontekstual tersebut, tetapi juga berpotensi menghasilkan model prediktif yang lebih akurat dan relevan secara praktis bagi institusi.

Berdasarkan uraian di atas, penelitian ini bertujuan untuk menganalisis faktor-faktor yang memengaruhi ketepatan waktu kelulusan mahasiswa Program Studi Matematika UIN Malang melalui pendekatan regresi logistik multinomial. Variabel dependen dalam penelitian ini adalah ketepatan waktu kelulusan yang diklasifikasikan menjadi tiga kategori: tepat waktu (7-8 semester), terlambat (9-10 semester), dan kritis (11 semester atau lebih). Sedangkan variabel independen meliputi faktor akademik (IPK, rasio SKS lulus), demografis (jenis kelamin, asal daerah), institusional (jenis asal sekolah, jalur masuk), serta finansial (jenis pembiayaan).

Dengan pemodelan regresi logistik multinomial, penelitian ini mengkaji faktor-faktor yang berdampak pada ketepatan waktu kelulusan mahasiswa Program Studi Matematika UIN Malang, sehingga hasilnya memiliki arti penting baik dari segi akademis maupun praktis. Dengan variabel dependen berupa kategori ketepatan waktu kelulusan (tepat waktu, terlambat, dan kritis) serta variabel independen meliputi IPK, jenis kelamin, jenis asal sekolah, jenis pembiayaan, jalur

masuk, asal daerah, dan rasio SKS lulus, hasil penelitian ini menghasilkan sebuah model prediktif yang mampu menghitung besarnya pengaruh setiap faktor terhadap probabilitas kelulusan mahasiswa. Temuan ini dapat menjadi landasan saintifik untuk merancang kebijakan intervensi dini yang lebih efektif dan berbasis bukti. Penerapan kebijakan ini pada gilirannya akan meningkatkan angka kelulusan tepat waktu, yang merupakan komponen esensial dalam asesmen mutu dan penguatan status akreditasi Program Studi Matematika UIN Maulana Malik Ibrahim Malang.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan sebelumnya, rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana model regresi logistik multinomial untuk ketepatan waktu kelulusan mahasiswa Program Studi Matematika UIN Maulana Malik Ibrahim Malang?
2. Berapa kinerja model regresi logistik multinomial tersebut dalam mengklasifikasikan ketepatan waktu kelulusan mahasiswa, yang diukur menggunakan *confusion matrix* beserta metrik *accuracy*, *precision*, *recall*, dan *F1-score*?
3. Bagaimana hasil prediksi probabilitas ketepatan waktu kelulusan mahasiswa menggunakan model yang telah terbentuk untuk sebuah studi kasus?

## 1.3 Tujuan Penelitian

Berikut ini adalah tujuan dari penelitian ini, yang didasarkan pada rumusan masalah yang telah diberikan sebelumnya:

1. Menentukan model regresi logistik multinomial untuk ketepatan waktu kelulusan mahasiswa Program Studi Matematika UIN Maulana Malik Ibrahim Malang.
2. Menghitung dan menganalisis kinerja model regresi logistik multinomial dalam mengklasifikasikan ketepatan waktu kelulusan mahasiswa dengan menggunakan *confusion matrix* untuk menghitung nilai *accuracy*, *precision*, *recall*, dan *F1-score*.
3. Menunjukkan contoh penerapan model untuk memprediksi probabilitas ketepatan waktu kelulusan mahasiswa pada sebuah studi kasus.

#### **1.4 Manfaat Penelitian**

Manfaat yang dapat diperoleh dari penelitian ini adalah sebagai berikut:

##### **1. Bagi Penulis**

Hasil penelitian ini memberikan pemahaman mendalam tentang penerapan metode regresi logistik multinomial untuk menganalisis faktor-faktor yang memengaruhi ketepatan waktu kelulusan mahasiswa. Penulis juga mendapatkan pengalaman dalam mengolah data statistik dan menyusun model prediktif yang relevan.

##### **2. Bagi Pembaca**

Hasil penelitian ini diharapkan dapat menjadi referensi ilmiah bagi mahasiswa, dosen, atau peneliti lain yang tertarik pada pemodelan statistik, khususnya dalam konteks pendidikan tinggi. Penelitian ini juga dapat menjadi dasar untuk studi lanjutan terkait optimasi masa studi mahasiswa atau pengembangan model prediksi lainnya dengan metode serupa.

3. Bagi Program Studi Matematika UIN Maulana Malik Ibrahim Malang  
Informasi yang dihasilkan dari penelitian ini dapat digunakan sebagai sebuah model prediktif fungsional bagi Program Studi Matematika UIN Maulana Malik Ibrahim Malang. Model ini berfungsi untuk mengidentifikasi mahasiswa yang berpotensi mengalami keterlambatan studi sejak dini, sehingga memungkinkan penerapan tindakan preventif yang efektif. Hasil analisis ini dapat menjadi landasan untuk evaluasi dan perumusan kebijakan sistem akademik yang lebih adaptif, yang bertujuan untuk meningkatkan tingkat kelulusan mahasiswa sesuai durasi studi normal. Implementasi kebijakan berbasis prediksi ini secara langsung akan berkontribusi pada peningkatan hasil asesmen kualitas dan status akreditasi program studi tersebut.

### **1.5 Batasan Masalah**

Batasan-batasan masalah yang diterapkan dalam penelitian ini meliputi:

1. Studi hanya mencakup mahasiswa Program Studi Matematika UIN Malang pada angkatan 2018-2021 yang telah menyelesaikan studi.
2. Data status kelulusan mahasiswa yang digunakan dalam penelitian ini adalah data yang tercatat hingga akhir semester genap tahun ajaran 2024-2025.
3. Proses estimasi dalam penelitian ini dilakukan dengan mengadopsi pendekatan *Maximum Likelihood Estimation* (MLE).
4. Kinerja model dievaluasi menggunakan *confusion matrix* untuk menghitung metrik performa secara keseluruhan (*accuracy, micro F1-score*) dan per-kategori kelulusan (*precision, recall, F1-score*).

5. Uji Multikolinearitas hanya dilakukan pada variabel-variabel numerik.

## BAB II

### KAJIAN TEORI

#### 2.1 Teori Pendukung

##### 2.1.1 Standardisasi

Standardisasi merupakan proses transformasi data untuk menyetarakan skala antarvariabel sehingga tidak ada variabel tertentu yang mendominasi analisis akibat perbedaan satuan atau rentang nilai. Dalam konteks analisis statistik multivariat seperti regresi logistik, perbedaan skala dapat menyebabkan ketidakstabilan numerik dan bias dalam estimasi parameter, terutama ketika variabel prediktor memiliki varians yang sangat berbeda (Field, 2013). Standardisasi dilakukan untuk mengonversi nilai asli data ke dalam satuan baku sehingga interpretasi dan kontribusi variabel dalam model menjadi lebih seimbang. Menurut Hair dkk. (2019), standardisasi juga dapat memperbaiki efisiensi proses estimasi menggunakan *Maximum Likelihood Estimation* (MLE) dengan memastikan setiap variabel memiliki skala yang konsisten dalam iterasi perhitungan model.

Metode standardisasi yang paling umum digunakan adalah Z-score, yaitu pendekatan yang mengubah suatu nilai berdasarkan posisi relatifnya terhadap rata-rata dan simpangan baku dari variabel tersebut. Nilai hasil transformasi menggambarkan seberapa jauh observasi berada dari rata-rata dalam satuan simpangan baku. Rumus Z-score dapat dilihat pada persamaan (2.1) sebagai berikut:

$$Z_i = \frac{X_i - \bar{X}}{S} \quad (2.1)$$

Keterangan:

$Z_i$  : Nilai data yang telah distandarisasi untuk observasi ke- $i$

$X_i$  : Nilai awal ke- $i$

$\bar{X}$  : Nilai rata-rata

$S$  : Simpangan baku sampel

Pendekatan ini penting digunakan ketika variabel prediktor digunakan secara simultan dalam suatu model statistik karena menghasilkan distribusi dengan rata-rata 0 dan simpangan baku 1, sehingga memudahkan perbandingan efek relatif setiap variabel dalam model (Hosmer dkk., 2013). Dengan demikian, standardisasi menjadi langkah fundamental dalam analisis data untuk memastikan kesetaraan skala serta meningkatkan keandalan dan akurasi estimasi model statistik.

### 2.1.2 Uji Multikolinearitas

Dalam konteks model regresi, multikolinieritas didefinisikan sebagai adanya korelasi linier antar variabel prediktor. Pengujian terhadap kondisi ini menjadi penting, terutama pada model regresi logistik yang melibatkan banyak prediktor, mengingat model tersebut sangat peka terhadap keberadaan multikolinieritas. Menurut Gujarati (2003), dampak utama dari multikolinieritas adalah meningkatnya nilai *standard error* pada estimasi koefisien regresi.

Asumsi fundamental dalam regresi logistik adalah tidak adanya multikolinearitas. Pelanggaran asumsi ini akan menyebabkan *standard error* koefisien regresi membengkak, yang pada gilirannya meningkatkan probabilitas



uji Wald untuk keliru menyimpulkan bahwa sebuah variabel prediktor tidak memiliki pengaruh signifikan. Deteksi multikolinearitas dapat dilakukan melalui perhitungan nilai VIF pada tiap prediktor. Data dianggap memiliki masalah multikolinearitas jika nilai VIF-nya melampaui angka 10. Berdasarkan panduan dari Kutner dkk. (2005), sebuah model dianggap bebas dari multikolinearitas jika nilai VIF hasil pengujiannya berada di bawah 10. Hipotesis untuk pengujian ini dirumuskan sebagai berikut:

Hipotesis:

$H_0$ : Tidak terdapat multikolinearitas antar variabel prediktor

$H_1$ : Terdapat multikolinearitas antar variabel prediktor

Statistik Uji:

$$(VIF)_k = \frac{1}{1 - R_k^2} \quad (2.2)$$

di mana  $k = 1, 2, \dots, p$

Keterangan:

$p$  : Banyaknya variabel prediktor

$R_k^2$  : Koefisien determinasi ganda dari variabel prediktor  $X_k$  dengan seluruh variabel prediktor lainnya.

### 2.1.3 Model Regresi Logistik Multinomial

Metode regresi logistik pada dasarnya digunakan untuk menganalisis variabel respon yang bersifat kategori, dengan prediktor yang bisa berupa kontinu maupun kategori. Secara spesifik, ketika variabel dependen tersebut berskala nominal dan terdiri atas lebih dari dua kategori (dikenal juga sebagai polikotomus), maka pendekatan yang paling sesuai adalah regresi logistik

multinomial. Metode ini pada dasarnya adalah sebuah ekstensi dari model regresi logistik biner, yang diperluas untuk menangani variabel dependen dengan kategori jamak (Hosmer & Lemeshow, 2000).

Dalam penelitian ini, dengan asumsi bahwa setiap kategori dari variabel  $Y$  diberi label kode numerik, misalnya 0, 1, dan 2. Dalam penerapannya, pemilihan kategori referensi menjadi penting, di mana biasanya kategori dasar (misalnya  $Y = 0$ ) digunakan sebagai pembanding untuk kategori lainnya ( $Y = 1$  dan  $Y = 2$ ). Model ini memerlukan pembentukan beberapa fungsi logit, dengan setiap logit membandingkan satu kategori terhadap kategori referensi. Perbedaan antara dua logit tersebut juga dapat digunakan untuk memodelkan perbandingan antar kategori non-referensi (misalnya  $Y = 2$  vs  $Y = 1$ ). Implementasi model ini membutuhkan perhatian terhadap pengkodean variabel, karena beberapa perangkat lunak statistik memiliki persyaratan tertentu terkait penomoran kategori (Hosmer & Lemeshow, 2000). Bentuk umum model regresi logistik untuk kasus ini dapat dilihat pada Persamaan (2.3).

$$\pi(x) = \frac{e^{g(x)}}{1 + e^{g(x)}} \quad (2.3)$$

dengan  $g(x) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$

Untuk mengembangkan model, diasumsikan terdapat  $p$  variabel kovariat dan sebuah suku konstanta yang direpresentasikan sebagai vektor  $\mathbf{x}$  dengan panjang  $p + 1$ , di mana  $x_0 = 1$ . Kedua fungsi logit tersebut dinyatakan sebagai berikut (Hosmer & Lemeshow, 2000).

$$g_1(x) = \ln \left[ \frac{P(Y = 1|\mathbf{x})}{P(Y = 0|\mathbf{x})} \right] = \beta_{10} + \beta_{11}x_1 + \beta_{12}x_2 + \dots + \beta_{1p}x_p \quad (2.4)$$

$$\begin{aligned}
&= \mathbf{x}'\boldsymbol{\beta}_1 \\
g_2(\mathbf{x}) &= \ln \left[ \frac{P(Y = 2|\mathbf{x})}{P(Y = 0|\mathbf{x})} \right] = \beta_{20} + \beta_{21}x_1 + \beta_{22}x_2 + \cdots + \beta_{2p}x_p \\
&= \mathbf{x}'\boldsymbol{\beta}_2
\end{aligned} \tag{2.5}$$

Dengan demikian, probabilitas bersyarat untuk setiap kategori respons berdasarkan vektor kovariat dapat dinyatakan sebagai berikut.

$$P(Y = 0|\mathbf{x}) = \frac{1}{1 + \exp g_1(\mathbf{x}) + \exp g_2(\mathbf{x})} \tag{2.6}$$

$$P(Y = 1|\mathbf{x}) = \frac{\exp g_1(\mathbf{x})}{1 + \exp g_1(\mathbf{x}) + \exp g_2(\mathbf{x})} \tag{2.7}$$

$$P(Y = 2|\mathbf{x}) = \frac{\exp g_2(\mathbf{x})}{1 + \exp g_1(\mathbf{x}) + \exp g_2(\mathbf{x})} \tag{2.8}$$

Mengikuti dengan konvensi model biner, didefinisikan  $\pi_j(\mathbf{x}) = P(Y = j|\mathbf{x})$  untuk  $j = 0, 1, 2$ . Setiap probabilitas merupakan fungsi dari vektor parameter  $\boldsymbol{\beta}' = (\boldsymbol{\beta}'_1, \boldsymbol{\beta}'_2)$  dengan dimensi  $2(p + 1)$  (Hosmer & Lemeshow, 2000).

Menurut Agresti (2002), parameter dalam model regresi logistik diestimasi menggunakan *Maximum Likelihood Estimation* (MLE). Prinsip kerja metode ini adalah memaksimalkan fungsi likelihood untuk memperoleh nilai taksiran  $\boldsymbol{\beta}$ .

Untuk menyusun fungsi likelihood, tiga variabel biner ( $Y_0, Y_1, Y_2$ ) yang dikodekan sebagai 0 atau 1 dibuat untuk menunjukkan keanggotaan kelompok suatu observasi (Hosmer & Lemeshow, 2000). Perlu ditekankan bahwa variabel-variabel ini diperkenalkan hanya untuk memperjelas formulasi fungsi *likelihood*

dan tidak digunakan dalam analisis aktual regresi logistik multinomial. Pengkodean variabel tersebut adalah sebagai berikut: jika  $Y = 0$  maka  $Y_0 = 1$ ,  $Y_1 = 0$ , dan  $Y_2 = 0$ ; jika  $Y = 1$  maka  $Y_0 = 0$ ,  $Y_1 = 1$ , dan  $Y_2 = 0$ ; serta jika  $Y = 2$  maka  $Y_0 = 0$ ,  $Y_1 = 0$ , dan  $Y_2 = 1$  (Hosmer & Lemeshow, 2000). Suatu sifat penting dari variabel-variabel ini adalah bahwa jumlah nilainya akan selalu sama dengan ( $\sum_{j=0}^2 Y_j = 1$ ) untuk semua nilai  $Y$  yang mungkin. Berdasarkan notasi ini, jika terdapat sebuah sampel dengan  $n$  amatan yang bersifat independen, maka fungsi likelihood bersyaratnya bisa dinyatakan dalam formulasi berikut (Hosmer & Lemeshow, 2000).

$$l(\boldsymbol{\beta}) = \prod_{i=1}^n [\pi_0(\mathbf{x}_i)^{y_{0i}} \pi_1(\mathbf{x}_i)^{y_{1i}} \pi_2(\mathbf{x}_i)^{y_{2i}}] \quad (2.9)$$

Dengan menerapkan log dan menggunakan fakta bahwa  $\sum y_{ji} = 1$  untuk setiap  $i$ , menghasilkan fungsi log-likelihood sebagai berikut (Hosmer & Lemeshow, 2000).

$$L(\boldsymbol{\beta}) = \sum_{i=1}^n y_{1i} g_1(\mathbf{x}_i) + y_{2i} g_2(\mathbf{x}_i) - \ln(1 + e^{g_1(\mathbf{x}_i)} + 1 + e^{g_2(\mathbf{x}_i)}) \quad (2.10)$$

Persamaan likelihood diperoleh melalui proses diferensiasi parsial pertama fungsi  $L(\boldsymbol{\beta})$  terhadap masing-masing dari  $2(p + 1)$  parameter yang belum diketahui. Untuk mempermudah notasi, didefinisikan  $\pi_{ji} = \pi_j(\mathbf{x}_i)$  sebagai penyederhanaan. Secara umum, bentuk persamaan-persamaan tersebut dapat dinyatakan sebagai:

$$\frac{\partial L(\boldsymbol{\beta})}{\partial \beta_{jk}} = \sum_{i=1}^n x_{ki} (y_{ji} - \pi_{ji}) \quad (2.11)$$

dengan  $j = 1, 2$  dan  $k = 0, 1, 2, \dots, p$  dengan  $x_{0i} = 1$  untuk setiap observasi (Hosmer & Lemeshow, 2000).

Penduga kemungkinan maksimum,  $\hat{\boldsymbol{\beta}}$ , diperoleh dengan menyamakan persamaan-persamaan tersebut dengan nol dan menyelesaikannya untuk  $\boldsymbol{\beta}$ . Solusi ini memerlukan komputasi iteratif yang serupa dengan yang digunakan pada kasus variabel respon biner (Hosmer & Lemeshow, 2000).

Matriks turunan parsial kedua diperlukan untuk memperoleh matriks informasi serta penduga matriks kovarians dari penduga kemungkinan maksimum. Secara umum, elemen-elemen dalam matriks turunan parsial kedua ini memiliki bentuk sebagai berikut:

$$\frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_{jk} \partial \beta_{j'k'}} = - \sum_{i=1}^n x_{k'i} x_{ki} \pi_{ji} (1 - \pi_{ji}) \quad (2.12)$$

dan

$$\frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_{jk} \partial \beta_{j'k'}} = \sum_{i=1}^n x_{k'i} x_{ki} \pi_{ji} \pi_{j'i} \quad (2.13)$$

dengan  $j$  dan  $j' = 1, 2$  dan  $k$  dan  $k' = 0, 1, 2, \dots, p$  (Hosmer & Lemeshow, 2000).

Mengingat persamaan (2.11) bersifat non-linear dalam  $\boldsymbol{\beta}$ , solusi analitik tidak dapat diperoleh secara langsung. Oleh karena itu, estimasi parameter dilakukan menggunakan metode numerik iteratif, yaitu metode Newton-Raphson. (Agresti, 2002).

Metode Newton-Raphson memperbarui nilai taksiran parameter dari iterasi ke- $t$ , yaitu  $\boldsymbol{\beta}^{(t)}$ , menuju nilai baru  $\boldsymbol{\beta}^{(t+1)}$  menggunakan rumus berikut:

$$\boldsymbol{\beta}^{(t+1)} = \boldsymbol{\beta}^{(t)} - (\mathbf{H}^{(t)})^{-1} \mathbf{u}^{(t)} \quad (2.14)$$

Keterangan:

$\mathbf{u}^{(t)}$  : vektor skor yang berisi elemen-elemen turunan parsial pertama

$\mathbf{H}^{(t)}$  : matriks hessian yang elemen-elemennya merupakan turunan parsial kedua

Proses iterasi ini dimulai dengan nilai awal (biasanya 0) dan terus berlanjut hingga selisih antara  $\beta^{(t+1)}$  dan  $\beta^{(t)}$  konvergen (mendekati nol). Nilai akhir dari iterasi inilah yang menjadi MLE untuk koefisien  $\beta$ .

Hosmer & Lemeshow (2000) menyatakan bahwa pengujian statistik harus dilakukan untuk mengetahui signifikansi pengaruh variabel prediktor terhadap variabel respon. Terdapat dua macam pengujian dalam model regresi logistik multinomial, yaitu uji parameter secara simultan menggunakan uji  $G$  dan uji parameter secara parsial menggunakan uji Wald.

#### 1. Pengujian Signifikansi Parameter Secara Simultan

Uji  $G$ , atau disebut juga *likelihood ratio test*, merupakan metode yang digunakan untuk melakukan pengujian signifikansi parameter guna mengevaluasi pengaruh gabungan dari semua variabel prediktor pada variabel respon.

Hipotesis :

$H_0$ :  $\beta_1 = \beta_2 = \dots = \beta_p = 0$  (tidak ada variabel prediktor yang memengaruhi model).

$H_1$ : Setidaknya ada satu  $\beta_i \neq 0$ , dengan  $i = 1, 2, \dots, p$ .

Statistik Uji :

$$G^2 = -2 \ln \left[ \frac{\left(\frac{n_0}{n}\right)^{n_0} \left(\frac{n_1}{n}\right)^{n_1} \left(\frac{n_2}{n}\right)^{n_2}}{\prod_{i=1}^n [\pi_0(x_i)^{y_{0i}} \pi_1(x_i)^{y_{1i}} \pi_2(x_i)^{y_{2i}}]} \right] \quad (2.15)$$

di mana,

$$n_0 = \sum_{i=1}^n y_{0i}; n_1 = \sum_{i=1}^n y_{1i}; n_2 = \sum_{i=1}^n y_{2i}; n = n_0 + n_1 + n_2$$

Keterangan :

$n_0$  : Jumlah observasi dengan kategori  $y = 0$

$n_1$  : Jumlah observasi dengan kategori  $y = 1$

$n_2$  : Jumlah observasi dengan kategori  $y = 2$

$n$  : Jumlah banyaknya observasi

Daerah Kritis:

Jika nilai  $G^2 > \chi_{(v,\alpha)}^2$  atau  $p\text{-value} < \alpha$ , maka tolak  $H_0$ , di mana  $v$  adalah derajat bebas.

## 2. Pengujian Signifikansi Parameter Secara Parsial

Untuk mengetahui signifikansi pengaruh masing-masing variabel prediktor terhadap variabel respon, dilakukan pengujian signifikansi parameter menggunakan uji Wald. Tujuan pengujian ini adalah untuk sebagai dasar pengambilan keputusan mengenai perlu atau tidaknya suatu variabel prediktor dimasukkan ke dalam model (Agresti, 2002).

Hipotesis :

$H_0: \beta_i = 0$  (tidak ada pengaruh yang signifikan).

$H_1: \beta_i \neq 0$ , di mana  $i = 1, 2, \dots, p$  (ada pengaruh yang signifikan).

Statistik Uji:

$$W = \left( \frac{\hat{\beta}_i}{\widehat{SE}(\hat{\beta}_i)} \right)^2 \quad (2.16)$$

Keterangan:

$(\hat{\beta}_i)$  : Nilai koefisien regresi untuk variabel prediktor ke- $i$ .

$\widehat{SE}(\hat{\beta}_i)$  : Taksiran *standard error* dari estimasi parameter  $\hat{\beta}_i$ .

Daerah Kritis:

Hipotesis nol ( $H_0$ ) akan ditolak apabila nilai statistik  $W > \chi^2_{(v,\alpha)}$  atau jika  $p\text{-value} < \alpha$ , di mana  $v$  adalah derajat bebas.

Untuk mengevaluasi kecocokan model regresi logistik yang telah diestimasi, diperlukan pengukuran sejauh mana model tersebut mampu menjelaskan variabel respon. Gagasan untuk menggunakan Uji *Goodness of Fit* dalam menilai kecocokan model dikemukakan oleh Hosmer & Lemeshow (2000), yang melibatkan hipotesis sebagai berikut.

Hipotesis:

$H_0$  : Model sesuai

$H_1$  : Model tidak sesuai

Statistik Uji:

$$\hat{C} = \sum_{k=1}^g \frac{(O_k - n'_k \bar{\pi}_k)^2}{n'_k \bar{\pi}_k (1 - \bar{\pi}_k)} \quad (2.17)$$

Keterangan:

$O_k$  : Observasi pada kategori ke- $k$

$n'_k$  : Jumlah observasi pada kategori ke- $k$

$\bar{\pi}_k$  : Rata-rata taksiran peluang ke- $k$

$g$  : Jumlah kategori

Daerah Kritis:

Tolak  $H_0$  jika  $\hat{C} > \chi^2_{(v,\alpha)}$  atau  $p\text{-value} < \alpha$ , dengan  $v = p - (k + 1)$  sebagai derajat bebas, di mana  $p$  menunjukkan jumlah parameter dan  $k$  merupakan banyaknya variabel prediktor dalam model.



Untuk mengevaluasi kebaikan model regresi logistik, diperlukan pengukuran sejauh mana variabel-variabel independen mampu menjelaskan variasi pada variabel dependen. Karena koefisien determinasi standar tidak dapat diterapkan pada model regresi logistik, maka digunakan ukuran Pseudo R-Square. Gagasan untuk menggunakan nilai ini adalah untuk meniru interpretasi  $R^2$  pada regresi linear (Agresti, 2002). Salah satu ukuran yang paling umum digunakan dan dilaporkan adalah Nagelkerke  $R^2$ , karena nilai ini telah dimodifikasi agar memiliki rentang antara 0 dan 1 (Widarjono, 2010). Kriteria pengambilan keputusannya adalah sebagai berikut:

1. Jika  $R^2$  mendekati 0, maka terdapat pengaruh lemah yang diberikan variabel independen terhadap variabel dependen dalam model regresi.
2. Jika  $R^2$  mendekati 1, maka terdapat pengaruh kuat yang diberikan variabel independen terhadap variabel dependen dalam model regresi.

Statistik Uji:

$$R_N^2 = \frac{R_{CS}^2}{1 - (LL(M_{awal}))^{(\frac{2}{n})}} \quad (2.18)$$

Keterangan:

$R_N^2$  : Nilai Pseudo R-Square Nagelkerke

$R_{CS}^2$  : Nilai Pseudo R-Square Cox & Snell

$LL(M_{awal})$  : Nilai *log-likelihood* dari model awal (hanya dengan intersep)

$n$  : Jumlah total observasi

Generalisasi notasi yang biasa digunakan untuk hasil biner diperlukan untuk menyederhanakan penjelasan mengenai estimasi dan interpretasi *odds ratio* saat dihadapkan pada konteks hasil multinomial, sebagaimana dikemukakan oleh

Hosmer & Lemeshow (2000). Generalisasi ini mencakup perbandingan hasil serta nilai kovariat. Dalam hal ini, hasil yang diberi label  $Y = 0$  ditetapkan sebagai kategori referensi. Subskrip pada *odds ratio* menunjukkan hasil mana yang dibandingkan terhadap hasil referensi. *Odds ratio* untuk hasil  $Y = j$  dibandingkan  $Y = 0$  pada nilai kovariat  $x = a$  versus  $x = b$  didefinisikan sebagai berikut.

$$OR_j(a, b) = \Psi_{ab} = \frac{P(Y = j|x = a)/P(Y = 0|x = a)}{P(Y = j|x = b)/P(Y = 0|x = b)} \quad (2.19)$$

Dalam kasus khusus ketika kovariat bersifat biner (dikodekan 0 atau 1), notasi dapat disederhanakan menjadi  $OR_j = OR_j(1,0)$  (Hosmer & Lemeshow, 2000).

#### 2.1.4 Evaluasi Kinerja Model dan *Confusion Matrix*

Hasil prediksi oleh model kemudian dibandingkan dengan data *testing*. Hal ini dilakukan untuk dapat menghitung nilai numerik metrik kinerja yang dibutuhkan dari sistem yang telah dibangun. Evaluasi kinerja model menggunakan *multiclass confusion matrix* karena penelitian ini termasuk kedalam *multiclass classification*.

Menurut Grandini dkk. (2020), *confusion matrix* adalah alat evaluasi fundamental yang merangkum semua informasi relevan mengenai kinerja algoritma dengan menyajikan rincian prediksi yang benar dan yang salah. Dari *confusion matrix* ini, nilai *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) untuk setiap kelas akan dihitung dan diagregasi (dijumlahkan) untuk mendapatkan ukuran performa model secara keseluruhan.

Dari *confusion matrix* 3x3, akan dihitung dua jenis metrik kinerja: metrik keseluruhan yang memberikan gambaran umum, dan metrik per-kategori yang memberikan analisis mendalam.

## 1. Metrik Kinerja Keseluruhan

### a. Akurasi (*Accuracy*)

Akurasi adalah salah satu metrik paling populer dalam klasifikasi multi-kelas, yang mengukur proporsi prediksi benar dari seluruh data uji. Seperti yang dijelaskan oleh Grandini dkk. (2020), akurasi pada dasarnya adalah probabilitas bahwa prediksi model itu benar jika kita memilih satu unit data secara acak.

$$Accuracy = \frac{TP}{\text{Total Data Testing}} \quad (2.20)$$

### b. *Micro-Average Metrics*

Metrik ini dihitung dengan menjumlahkan semua *True Positive* (TP), *False Positive* (FP), dan *False Negative* (FN) dari semua kelas, kemudian menghitung *Precision*, *Recall*, dan *F1-Score* dari total tersebut. Pendekatan ini cenderung memberikan bobot lebih pada kelas yang lebih besar (kelas mayoritas). Nilai *Micro-F1 Score* identik dengan nilai Akurasi (Grandini dkk., 2020).

## 2. Metrik Kinerja Per-Kategori

### a. Presisi per Kelas (*Precision per Class*)

Presisi mendefinisikan proporsi data yang diprediksi sebagai positif, yang ternyata memang benar-benar positif. Dalam konteks ini, presisi menjawab seberapa besar kita bisa mempercayai model ketika ia

memprediksi seorang mahasiswa masuk ke dalam kategori kelulusan tertentu.

$$Precision_k = \frac{TP_k}{TP_k + FP_k} \quad (2.21)$$

b. *Recall per Kelas (Recall per Class)*

*Recall* mengukur kemampuan model untuk menemukan kembali semua data yang relevan dalam sebuah kelas. Metrik ini, yang oleh Grandini dkk. (2020), juga disebut sebagai *True Positive Rate* (TPR) atau sensitivitas, menghitung proporsi data positif aktual yang berhasil diidentifikasi dengan benar oleh model.

$$Recall_k = \frac{TP_k}{TP_k + FN_k} \quad (2.22)$$

c. *F1-Score per Kelas (F1-Score per Class)*

*F1-Score* dihitung sebagai rata-rata harmonik antara *precision* dan *recall*, yang berguna untuk menemukan keseimbangan terbaik antara kedua metrik tersebut. Ketika *precision* dan *recall* dihitung dari total TP, FP, dan FN (metode *micro-averaging*), maka *F1-Score* yang dihasilkan disebut sebagai *Micro F1-Score*. Grandini dkk. (2020), menunjukkan bahwa dalam skenario ini, nilai *Micro F1-Score* akan identik dengan nilai akurasi.

$$F1 - Score_k = 2 \times \frac{Recall_k \times Precision_k}{Recall_k + Precision_k} \quad (2.23)$$

Dengan menyajikan metrik kinerja keseluruhan (Akurasi) dan metrik per-kategori (*Precision*, *Recall*, *F1-Score* untuk setiap kelas), analisis kinerja model akan menjadi sangat komprehensif dan mampu menunjukkan kekuatan serta kelemahan model secara detail.

### **2.1.5 Ketepatan Waktu Kelulusan Mahasiswa**

Keberhasilan sebuah institusi pendidikan tinggi salah satunya tecermin dari durasi studi dan tingkat kelulusan tepat waktu mahasiswanya. Hal ini menjadi krusial karena kualitas lulusan yang dihasilkan akan membangun reputasi dan kepercayaan di dunia kerja, sehingga perusahaan tidak ragu untuk merekrut talenta dari institusi tersebut. Dengan demikian, memastikan mahasiswa lulus tepat waktu merupakan bagian integral dari strategi institusi pendidikan tinggi untuk menghasilkan lulusan yang mampu bersaing di kancah internasional.

Untuk dapat dinyatakan lulus dari program Sarjana (S1), mahasiswa diwajibkan untuk menuntaskan setidaknya 144 SKS dalam rentang waktu 4 hingga 7 tahun (8-14 semester). Ketentuan ini didasarkan pada Pasal 27 dalam Peraturan Menteri Ristekdikti mengenai Standar Nasional Pendidikan Tinggi (Pedoman Pendidikan UIN Malang 2019). Kriteria kelulusan mahasiswa dapat dibagi menjadi dua, yaitu berbasis kompetensi dan berbasis waktu. Dari segi kompetensi, mahasiswa harus memiliki IPK minimal 2.00 dan menyelesaikan berbagai kewajiban seperti KKM, PKL, tugas akhir, serta publikasi ilmiah (UU RI No.12, 2012). Dari segi waktu, khususnya untuk program S1, kelulusan dianggap tepat waktu jika dicapai dalam empat tahun atau kurang (Kemenristekdikti, 2013).

Penetapan kategori ketepatan waktu kelulusan menjadi 'Tepat Waktu' (7-8 semester), 'Terlambat' (9-10 semester), dan 'Kritis' (11 semester atau lebih) merupakan sebuah definisi operasional yang dirumuskan untuk keperluan analisis dalam penelitian ini. Kategori 'Tepat Waktu' didasarkan pada standar nasional yang berlaku. Sementara itu, pembagian kategori 'Terlambat' dan 'Kritis'

didasarkan pada sistem klasifikasi dan evaluasi masa studi yang ditetapkan oleh universitas. Sistem tersebut secara internal mengklasifikasikan mahasiswa yang melewati batas studi normal ke dalam beberapa tingkatan monitoring, seperti 'Waspada' untuk masa studi 9-10 semester dan 'Kritis' untuk 11-14 semester. Oleh karena itu, pengkategorian dalam penelitian ini selaras dengan praktik administratif dan sistem peringatan dini yang sudah diterapkan oleh institusi. Pemilihan atribut-atribut dalam penelitian ini dilandasi oleh sejumlah justifikasi yang logis. Pertama, untuk data akademik, diasumsikan bahwa data mahasiswa dari semester 1 hingga 6 relatif homogen. Asumsi ini merujuk pada Pedoman Pendidikan UIN Malang yang menempatkan mata kuliah wajib pada periode tersebut, sehingga diharapkan dapat mengurangi jumlah *outliers* atau data anomali.

Kedua, pemilihan atribut jenis kelamin dilandasi oleh teori psikologi mengenai perbedaan kemampuan kognitif. Meskipun Santrock (2014) tidak menemukan perbedaan intelektual secara umum, beberapa penelitian (misalnya Suprpto dkk., 2018) menunjukkan adanya sedikit keunggulan laki-laki di bidang sains, yang berpotensi memengaruhi ketepatan waktu kelulusan. Pilihan ini juga didukung oleh penelitian sebelumnya (Syahrani, 2022).

Ketiga, pemilihan atribut jenis asal sekolah didasarkan pada pengaruh latar belakang pendidikan sebelumnya terhadap kesiapan akademik mahasiswa. Menurut penelitian Djudin (2018), mahasiswa yang berasal dari sekolah dengan kurikulum berbasis sains (seperti SMA IPA) cenderung lebih adaptif dalam menghadapi mata kuliah matematika dan sains di perguruan tinggi dibandingkan lulusan SMA IPS atau SMK. Hal ini mendukung pemilihan atribut jenis asal

sekolah sebagai salah satu faktor yang berpotensi memengaruhi ketepatan waktu kelulusan.

Keempat, pemilihan atribut jenis pembiayaan dipilih karena adanya persyaratan khusus pada beasiswa Bidikmisi yang secara inheren mendorong mahasiswa untuk lulus tepat waktu. Menurut situs resmi Ristekdikti (2019), penerima beasiswa Bidikmisi memiliki sejumlah kewajiban yang dipantau oleh Belmawa. Kewajiban tersebut antara lain adalah menjaga stabilitas indeks prestasi setiap semester, mengikuti pembinaan agar dapat menyelesaikan studi sesuai target, mempertahankan status sebagai mahasiswa aktif, dan mematuhi seluruh peraturan akademik yang berlaku.

Kelima, pemilihan atribut jalur masuk merujuk pada perbedaan selektivitas dan karakteristik mahasiswa berdasarkan mekanisme penerimaan. Mahasiswa yang diterima melalui jalur undangan (SNMPTN) rata-rata menyelesaikan studi lebih singkat daripada mereka yang diterima melalui tes tulis (SBMPTN) atau jalur mandiri. Menurut Djudin (2018), hal ini terjadi karena seleksi SNMPTN lebih mengutamakan rekam jejak akademik yang konsisten. Temuan serupa diungkapkan oleh Imaslihkah dkk. (2013) yang menyatakan bahwa predikat kelulusan mahasiswa di ITS Surabaya secara signifikan terkait dengan jalur masuk, dengan mahasiswa SNMPTN menunjukkan performa lebih stabil. Atribut ini juga digunakan dalam penelitian Mashfia (2022) sebagai prediktor ketepatan waktu kelulusan.

Keenam, pemilihan atribut asal daerah didasarkan pada pengaruh faktor geografis dan lingkungan terhadap adaptasi akademik mahasiswa. Penelitian Mashfia (2022) merekomendasikan penggunaan variabel asal daerah dalam

analisis ketepatan waktu kelulusan, mengingat perbedaan akses pendidikan dan fasilitas pendukung belajar antar daerah dapat berdampak pada kesiapan mahasiswa menghadapi perkuliahan. Selain itu, faktor budaya dan jarak tempuh dari daerah asal ke kampus juga memengaruhi motivasi dan interaksi mahasiswa dengan lingkungan akademik.

Ketujuh, pemilihan atribut rasio SKS lulus didasarkan pada perannya sebagai indikator fundamental yang mengukur kemajuan dan efisiensi akademik mahasiswa. Variabel ini melengkapi IPK dengan memberikan gambaran kuantitatif mengenai progresi studi, sementara IPK mengukur kualitas pemahaman. Rasio SKS Lulus mengukur kecepatan dan kuantitas pemenuhan syarat kelulusan. Hal ini sejalan dengan penelitian oleh Malay dkk. (2025) yang menemukan bahwa jumlah SKS yang diambil per semester berpengaruh signifikan terhadap peluang kelulusan tepat waktu. Selain itu, definisi variabel yang memperhitungkan SKS "tanpa mengulang" secara langsung menangkap aspek efisiensi, di mana pengulangan mata kuliah merupakan salah satu faktor risiko keterlambatan (Malay dkk., 2025). Penggunaan jumlah SKS sebagai prediktor juga telah terbukti relevan dalam berbagai studi prediksi akademik, seperti yang dilakukan oleh Muhendra (2021) dan Wahyuni dkk. (2018). Oleh karena itu, rasio SKS lulus dimasukkan sebagai prediktor kunci karena kemampuannya dalam merepresentasikan kelancaran mahasiswa dalam memenuhi beban studi yang disyaratkan.



## 2.2 Kajian Integrasi Ketepatan Waktu Kelulusan dengan Al-Quran / Hadits

Permasalahan ketepatan waktu kelulusan mahasiswa tidak hanya bersifat administratif, tetapi juga memiliki dimensi spiritual dalam Islam. Konsep disiplin waktu dan tanggung jawab dalam menyelesaikan kewajiban tercermin dalam firman Allah Surat Al-Mu'minin ayat 1-3:

قَدْ أَفْلَحَ الْمُؤْمِنُونَ ﴿١﴾ الَّذِينَ هُمْ فِي صَلَاتِهِمْ خَاشِعُونَ ﴿٢﴾ وَالَّذِينَ هُمْ عَنِ اللَّغْوِ مُعْرِضُونَ ﴿٣﴾

“*Sungguh, beruntunglah orang-orang mukmin. (Yaitu) orang-orang yang khusyuk dalam shalatnya, orang-orang yang meninggalkan (perbuatan dan perkataan) yang tidak berguna.*” (QS. Al-Mu'minin: 1-3)

Dalam tafsir Al-Misbah, Quraish Shihab menjelaskan bahwa QS. Al-Mu'minin ayat 1-3 menekankan konsep kesuksesan sejati yang dicapai melalui kedisiplinan waktu dan konsentrasi penuh dalam beribadah (*khusyu'*), di mana kualitas keimanan seseorang tercermin dari kemampuannya mengelola waktu secara produktif. Shihab menegaskan bahwa penghindaran dari perbuatan sia-sia (*al-laghwu*) dalam ayat ini merupakan manifestasi dari kehidupan yang teratur dan berorientasi tujuan, sebagaimana diajarkan dalam Islam (Shihab, 2002). Konsep ini memiliki relevansi kuat dengan sistem pendidikan tinggi, di mana ketepatan waktu penyelesaian studi menjadi indikator kedisiplinan dan keseriusan akademik, sebagaimana ketepatan waktu dalam melaksanakan ibadah salat mencerminkan ketaatan seorang muslim. Hal ini sejalan dengan sabda Rasulullah yang diriwayatkan Abu Daud no. 426:

عَنْ أُمِّ فَرْوَةَ قَالَتْ سَأَلْتُ رَسُولَ اللَّهِ -صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ- أَيُّ الْأَعْمَالِ أَفْضَلُ قَالَ (الصَّلَاةُ فِي أَوَّلِ وَفَيْهَا)

*Dari Ummu Farwah, ia berkata, “Rasulullah shallallahu ‘alaihi wa sallam pernah ditanya, amalan apakah yang paling afdhol. Beliau pun menjawab, “Shalat di awal waktunya.” (HR. Abu Daud no. 426)*

Hadits ini menunjukkan keutamaan menyelesaikan kewajiban pada awal waktu. Dalam konteks akademik, mahasiswa yang mampu menyelesaikan studi tepat waktu atau bahkan lebih cepat mencerminkan kedisiplinan dan efektivitas dalam mengelola waktu, sebagaimana keutamaan salat di awal waktu. Kedua konsep ini sama-sama menekankan pentingnya manajemen waktu yang baik sebagai indikator kesuksesan, baik dalam dimensi spiritual maupun akademik.

Permasalahan ketepatan waktu kelulusan tidak hanya terkait dengan manajemen waktu, tetapi juga dipengaruhi oleh faktor motivasi. Dalam Islam, motivasi untuk bersungguh-sungguh dalam menuntut ilmu dan menyelesaikan kewajiban tercermin dalam firman Allah Surat At-Taubah ayat 105:

وَقُلِ اعْمَلُوا فَسَيَرَى اللَّهُ عَمَلَكُمْ وَرَسُولُهُ وَالْمُؤْمِنُونَ وَسُرَدُّونَ إِلَىٰ عِلْمِ الْغَيْبِ وَالشَّهَادَةِ فَيُنبِّئُكُمْ بِمَا كُنْتُمْ تَعْمَلُونَ ﴿١٠٥﴾

*Katakanlah (Nabi Muhammad), “Bekerjalah! Maka, Allah, rasul-Nya, dan orang-orang mukmin akan melihat pekerjaanmu. Kamu akan dikembalikan kepada (Zat) yang mengetahui yang gaib dan yang nyata. Lalu, Dia akan memberitakan kepada kamu apa yang selama ini kamu kerjakan.” (QS. At-Taubah: 105)*

Ayat ini mengajarkan bahwa setiap usaha manusia berada dalam pengawasan Allah SWT, Rasul-Nya, dan orang-orang beriman. Kesadaran ini membentuk sistem pertanggungjawaban yang mencakup dimensi spiritual dan sosial. Bagi mahasiswa, pemahaman tersebut dapat menjadi dorongan untuk menyelesaikan studi dengan baik dan tepat waktu. Di satu sisi, kesadaran akan pentingnya prestasi akademik dapat menjadikan mahasiswa sebagai teladan bagi sesama muslim. Di

sisi lain, keyakinan bahwa menyelesaikan pendidikan sesuai jadwal merupakan bentuk tanggung jawab terhadap amanah yang diberikan oleh Allah SWT.

Dalam tafsir Al-Misbah, Quraish Shihab (2002) menjelaskan bahwa perintah “bekerjalah” dalam ayat ini menekankan pentingnya kesungguhan dalam menjalani aktivitas kehidupan, dengan keyakinan bahwa setiap amal akan memperoleh balasan yang adil dari Allah, baik di dunia maupun di akhirat. Dalam konteks pendidikan, khususnya terkait ketepatan waktu kelulusan, penafsiran ini menunjukkan bahwa kesungguhan mahasiswa tidak hanya diukur dari capaian akademik, tetapi juga dari proses yang mencerminkan tanggung jawab, etos kerja, dan komitmen yang konsisten. Ketepatan waktu dalam menyelesaikan studi bukan semata-mata target administratif, melainkan bagian dari perwujudan iman dan integritas pribadi dalam memaknai serta mempertanggungjawabkan amanah waktu dan ilmu di hadapan Allah SWT.

### **2.3 Kajian Ketepatan Waktu Kelulusan dengan Teori Pendukung**

Penelitian mengenai analisis regresi logistik multinomial dilakukan melalui beberapa tahapan yang sistematis. Penelitian ini diawali dengan tahap pengumpulan data. Data yang digunakan bersifat sekunder, yaitu berupa rekam jejak akademik dari mahasiswa Program Studi Matematika UIN Malang angkatan 2018-2021. Data tersebut terdiri dari satu variabel dependen (status kelulusan: tepat waktu, terlambat, dan kritis) dan beberapa variabel independen, mencakup IPK, jenis kelamin, jenis asal sekolah, jenis pembiayaan, jalur masuk, asal daerah, serta rasio SKS lulus.

Setelah data terkumpul, dilakukan persiapan data yang meliputi penanganan *missing* data melalui penghapusan, dan pengkodean variabel kategorik ke dalam bentuk *dummy variable* khususnya untuk jenis kelamin, jenis asal sekolah, jenis pembiayaan, jalur masuk, dan asal daerah. Proses ini bertujuan untuk memenuhi asumsi dasar analisis dan meningkatkan akurasi hasil pemodelan. Tahap selanjutnya, dilakukan statistika deskriptif untuk menggambarkan karakteristik data secara umum, seperti distribusi frekuensi, nilai mean, standar deviasi, minimum dan maksimum. Tujuan dari tahap ini adalah untuk mendapatkan pemahaman pendahuluan terkait pola serta tendensi yang terdapat pada data penelitian.

Tahap selanjutnya, sebelum membangun model regresi logistik multinomial, dilakukan terlebih dahulu uji multikolinearitas. Pengujian multikolinearitas dilaksanakan pada variabel-variabel prediktor numerik dengan tujuan untuk memverifikasi ketiadaan korelasi linier yang kuat di antara variabel-variabel tersebut berdasarkan persamaan (2.2). Tahap selanjutnya dilakukan standarisasi variabel numerik yaitu IPK ( $x_1$ ) dan Rasio SKS Lulus ( $x_7$ ) agar skala data seragam. Proses standarisasi ini dilakukan menggunakan metode *Z-score* sesuai dengan Persamaan (2.1), di mana setiap nilai data diubah berdasarkan posisi relatifnya terhadap rata-rata dan simpangan baku.

Tahap selanjutnya adalah pembagian data, data dibagi menjadi dua bagian, yaitu data latih (*training*) dan data uji (*testing*), dengan proporsi 80% untuk keperluan pembangunan model dan 20% sisanya dialokasikan untuk mengevaluasi performa model tersebut. Selanjutnya, membangun model regresi logistik multinomial dengan menggunakan *dummy variable*. Tahap ini melibatkan penetapan kategori referensi ( $Y = 0$ ) sebagai dasar pembandingan untuk kategori

lainnya ( $Y = 1$  dan  $Y = 2$ ), serta perumusan fungsi logit untuk setiap kategori non-referensi sesuai persamaan (2.4) dan (2.5). Model dibangun dengan memasukkan variabel prediktor yang relevan ke dalam vektor kovariat  $\mathbf{x}$  dengan panjang  $p + 1$ , di mana  $x_0 = 1$  merepresentasikan konstanta. Probabilitas bersyarat untuk setiap kategori respon dihitung berdasarkan persamaan (2.6) hingga (2.8), dengan memastikan total probabilitas seluruh kategori sama dengan satu.

Tahap selanjutnya, dilakukan estimasi parameter dengan menggunakan metode MLE sebagaimana tercantum dalam persamaan likelihood (2.9) dan fungsi log-likelihood (2.10). Untuk mendapatkan taksiran koefisien regresi, dilakukan penyelesaian pada persamaan (2.11). Mengingat persamaan ini bersifat non-linear, solusinya diperoleh secara numerik melalui proses iteratif menggunakan metode Newton-Raphson berdasarkan persamaan (2.14). Dalam metode tersebut, nilai  $\hat{\beta}$  diperbarui pada setiap iterasi dengan memanfaatkan turunan parsial pertama dalam persamaan (2.11) dan turunan parsial kedua dalam persamaan (2.12 dan 2.13) hingga mencapai konvergensi. Hasil estimasi ini bertujuan untuk menghasilkan taksiran koefisien regresi yang paling mampu dalam memodelkan keterkaitan variabel independen dan dependen.

Setelah parameter diestimasi, berikutnya dilaksanakan pengujian signifikansi parameter, yang mencakup uji serentak (simultan) dengan uji G berdasarkan persamaan (2.15) untuk menilai apakah model secara keseluruhan signifikan, serta uji parsial menggunakan uji *wald* berdasarkan persamaan (2.16) untuk menguji kontribusi masing-masing variabel independen. Jika model tidak signifikan, maka perlu dilakukan perbaikan model dengan meninjau kembali pemilihan variabel atau variabel tidak dimasukkan dalam model.

Tahap selanjutnya adalah uji kesesuaian model berdasarkan persamaan (2.17) untuk memastikan bahwa model dapat menggambarkan data dengan baik, dilakukan dengan menggunakan uji *Goodness Of Fit*. Tahap selanjutnya adalah pengukuran kebaikan model berdasarkan persamaan (2.18) untuk mengukur kekuatan penjelasan model. Pengukuran kebaikan model ini akan dilakukan dengan menghitung nilai Pseudo R-Square, seperti Nagelkerke  $R^2$ . Nilai ini akan menunjukkan seberapa besar persentase variasi variabel dependen yang mampu dijelaskan oleh variabel-variabel independen yang ada di dalam model. Selanjutnya, dilakukan evaluasi kinerja model secara komprehensif dengan menghitung metrik keseluruhan (Akurasi) berdasarkan persamaan (2.20), dilengkapi dengan *Micro-Average Metrics* untuk mendapatkan skor tunggal *precision*, *recall*, dan *F1-score* dari agregasi semua kelas, dan metrik per-kategori (*Precision*, *Recall*, dan *F1-Score*) berdasarkan persamaan (2.21) hingga (2.23).

Tahap selanjutnya, dilakukan interpretasi model dengan menghitung odds ratio berdasarkan persamaan (2.19). Dalam analisis ini, koefisien regresi diubah menjadi odds ratio untuk mengukur pengaruh variabel prediktor terhadap probabilitas kategori ketepatan waktu kelulusan. Signifikansi dari hasil penelitian ini terletak pada kontribusinya sebagai data pendukung untuk evaluasi kebijakan akademik, yang berorientasi pada peningkatan proporsi mahasiswa lulus tepat waktu. Keberhasilan dalam upaya ini secara langsung akan menyempurnakan penilaian kualitas dan status akreditasi Program Studi Matematika di UIN Maulana Malik Ibrahim Malang.

Tahap terakhir, setelah interpretasi model, adalah melakukan simulasi prediksi pada studi kasus. Tahap ini bertujuan untuk mengilustrasikan penerapan

praktis dari model yang telah terbentuk dengan membuat sebuah studi kasus dari data *testing* yang merepresentasikan karakteristik mahasiswa tertentu. Nilai-nilai variabel independen dari kasus tersebut kemudian disubstitusikan ke dalam persamaan model regresi logistik multinomial yang telah diestimasi. Hasil dari simulasi ini adalah nilai probabilitas bagi mahasiswa tersebut untuk masuk ke dalam setiap kategori kelulusan (Tepat Waktu, Terlambat, dan Kritis), di mana kategori dengan probabilitas tertinggi akan menjadi hasil akhir prediksi model untuk studi kasus tersebut. Langkah ini bertujuan untuk menunjukkan bagaimana model yang telah dibangun dapat digunakan sebagai alat bantu prediksi dalam konteks nyata.

## **BAB III**

### **METODE PENELITIAN**

#### **3.1 Jenis Penelitian**

Penelitian ini berfokus pada analisis data numerik untuk mendefinisikan, mengelompokkan, dan menguji hubungan antar variabel penelitian sebelum penarikan kesimpulan dilakukan. Untuk membangun landasan teoretis yang kuat, penelitian ini didukung oleh tinjauan pustaka dari berbagai sumber referensi seperti buku, jurnal, dan artikel ilmiah mengenai regresi logistik multinomial serta faktor-faktor yang memengaruhi ketepatan waktu kelulusan.

#### **3.2 Data dan Sumber Data**

Penelitian ini menggunakan data sekunder yang bersumber dari Program Studi Matematika UIN Maulana Malik Ibrahim Malang. Populasi dalam penelitian ini adalah seluruh mahasiswa angkatan 2018-2021 yang berjumlah 336 orang. Dari populasi tersebut, dilakukan proses pembersihan data (*data cleaning*) untuk menangani data yang tidak lengkap atau tidak memenuhi kriteria batasan masalah. Setelah proses tersebut, diperoleh populasi akhir yang valid dan siap untuk dianalisis sebanyak 331 mahasiswa. Rincian lengkap data yang digunakan dalam penelitian ini disajikan pada Lampiran 1 pada nomor 1.

Penelitian ini melibatkan satu variabel dependen dan sejumlah variabel independen yang dijadikan dasar dalam analisis regresi. Variabel dependen adalah status ketepatan waktu kelulusan mahasiswa. Sementara itu, variabel independennya meliputi IPK, jenis kelamin, jenis asal sekolah, jenis pembiayaan,



jalur masuk, asal daerah, dan rasio SKS lulus. Struktur dan pengelompokan seluruh variabel tersebut disajikan pada Tabel 3.1.

**Tabel 3.1** Variabel Penelitian

Variabel	Kategori	Keterangan	Skala
Ketepatan Waktu Kelulusan (Y)	0 = Tepat Waktu	Lulus 7-8 semester	Nominal
	1 = Terlambat	Lulus 9-10 semester	
	2 = Kritis	Lulus 11-14 semester	
IPK (Semester 1-6) ( $X_1$ )	Numerik	Rata-rata IPK Kumulatif hingga semester 6	Interval
Jenis Kelamin ( $X_2$ )	1 = Laki-laki		Nominal
	2 = Perempuan		
Jenis Asal Sekolah ( $X_3$ )	1 = MA		Nominal
	2 = SMA		
	3 = SMK		
Jenis Pembiayaan ( $X_4$ )	1 = Bidikmisi		Nominal
	2 = Non Bidikmisi		
Jalur Masuk ( $X_5$ )	1 = SNMPTN		Nominal
	2 = SBMPTN		
	3 = Mandiri		
Asal Daerah ( $X_6$ )	1 = Kota Malang		Nominal
	2 = Luar Kota Malang		
Rasio SKS Lulus ( $X_7$ )	Numerik	Rasio total SKS lulus hingga semester 6 tanpa SKS mengulang terhadap total SKS lulus yang disyaratkan oleh Program Studi	Rasio

### 3.3 Teknik Analisis Data

Berikut adalah beberapa tahapan yang dilalui dalam proses analisis data pada penelitian ini:

#### 1. Persiapan Data

Tahap awal dalam analisis regresi logistik multinomial adalah persiapan data yang melibatkan serangkaian langkah krusial. Dimulai dengan penanganan data yang hilang (*missing data*) dengan menggunakan metode penghapusan.

Setelah data bersih, dilakukan pengkodean variabel kategorik meliputi jenis kelamin, jenis asal sekolah, jenis pembiayaan, jalur masuk, dan asal daerah menjadi *dummy variable* agar dapat diproses secara matematis dalam model.

## 2. Statistika Deskriptif

Tahap selanjutnya dilakukan statistika deskriptif pada data akademik mahasiswa Program Studi Matematika angkatan 2018-2021. Tahap ini bertujuan untuk memahami karakteristik data, seperti distribusi frekuensi, mean, standar deviasi, nilai maksimum, dan minimum dari variabel-variabel yang relevan. Selain itu, visualisasi data juga akan digunakan untuk memperjelas temuan-temuan awal terkait distribusi dan karakteristik variabel.

## 3. Uji Multikolinearitas

Tahap selanjutnya dilakukan uji multikolinearitas antar variabel prediktor numerik bertujuan untuk memverifikasi tidak adanya korelasi tinggi antara variabel-variabel prediktor berdasarkan persamaan (2.2). Nilai VIF digunakan sebagai indikator dalam uji ini, di mana nilai VIF yang melebihi batas tertentu ( $VIF > 10$ ) menunjukkan adanya multikolinearitas. Untuk mengatasi terjadinya multikolinearitas, dilakukan langkah-langkah sebagai berikut: (1) penghapusan variabel prediktor yang memiliki nilai VIF tertinggi secara bertahap, (2) analisis ulang VIF setelah setiap penghapusan variabel, dan (3) iterasi proses ini hingga semua variabel yang tersisa memiliki nilai  $VIF \leq 10$ . Proses eliminasi ini dilakukan dengan pertimbangan teoritis untuk mempertahankan variabel-variabel yang paling relevan secara substansial dalam model.

#### 4. Standardisasi

Tahap selanjutnya adalah standardisasi data, yang difokuskan khusus pada variabel numerik, yaitu IPK ( $x_1$ ) dan Rasio SKS Lulus ( $x_7$ ). Standardisasi ini bertujuan untuk menyeragamkan skala pengukuran data sehingga memiliki nilai rata-rata nol dan simpangan baku satu (*Z-score normalization*), guna meminimalisir bias yang mungkin timbul akibat perbedaan rentang nilai antar variabel dan membantu konvergensi model. Proses standardisasi ini dilakukan dengan mengacu pada metode *Z-score* seperti yang ditunjukkan pada persamaan (2.1), di mana setiap nilai observasi ( $x_i$ ) dikurangi rata-rata ( $\bar{x}$ ) dan dibagi dengan simpangan baku ( $S$ ).

#### 5. Pembagian Data

Data kemudian dibagi menjadi dua bagian menggunakan proporsi 80:20, yaitu 80% data training untuk membangun model dan 20% data testing untuk evaluasi model. Pembagian ini dilakukan secara acak dengan *stratified sampling* untuk memastikan validitas hasil analisis.

#### 6. Membangun Model Regresi Logistik Multinomial

Sebelum melakukan estimasi parameter, tahap yang harus diselesaikan adalah membangun model regresi logistik multinomial dengan menggunakan *dummy variable*. Tahap ini melibatkan penetapan kategori referensi ( $Y = 0$ ) sebagai dasar pembandingan untuk kategori lainnya ( $Y = 1$  dan  $Y = 2$ ), serta perumusan fungsi logit untuk setiap kategori non-referensi sesuai persamaan (2.4) dan (2.5). Model dibangun dengan memasukkan variabel prediktor yang relevan ke dalam vektor kovariat  $\mathbf{x}$  dengan panjang  $p + 1$ , di mana  $x_0 = 1$  merepresentasikan konstanta. Probabilitas bersyarat untuk setiap kategori

respon dihitung berdasarkan persamaan (2.6) hingga (2.8), dengan memastikan total probabilitas seluruh kategori sama dengan satu.

#### 7. Estimasi Parameter

Tahap selanjutnya adalah melakukan estimasi parameter menggunakan metode Maximum Likelihood Estimation (MLE). Berdasarkan persamaan likelihood (2.9) dan fungsi log-likelihood (2.10), proses ini bertujuan untuk mendapatkan taksiran koefisien regresi ( $\hat{\beta}$ ) yang memaksimalkan fungsi log-likelihood. Nilai  $\hat{\beta}$  tersebut diperoleh dengan menyelesaikan persamaan (2.11) secara numerik menggunakan metode iteratif Newton-Raphson hingga mencapai konvergensi berdasarkan persamaan (2.14). Hasil estimasi ini akan memberikan nilai koefisien regresi yang paling mungkin menggambarkan keterkaitan antara serangkaian variabel independen dan satu variabel dependen kategorik dalam model regresi logistik multinomial.

#### 8. Uji Signifikansi Parameter

Tahap selanjutnya dilakukan uji signifikansi parameter secara simultan dan parsial. Uji simultan, yang menggunakan metode likelihood ratio test atau uji G berdasarkan persamaan (2.15), bertujuan untuk mengevaluasi signifikansi model secara keseluruhan. Jika model tidak signifikan, maka variabel perlu ditinjau ulang atau dikeluarkan dari pemodelan. Sementara itu, uji parsial yang menerapkan Uji Wald berdasarkan persamaan (2.16) digunakan untuk menganalisis kontribusi individual dari setiap variabel independen, yang didasarkan pada nilai p-value masing-masing koefisien.

#### 9. Uji Kesesuaian Model

Tahap selanjutnya untuk memvalidasi kecocokan antara model regresi logistik multinomial yang dikembangkan dengan data penelitian, dilakukan uji kesesuaian model menggunakan uji Hosmer-Lemeshow berdasarkan persamaan (2.17). Uji ini membandingkan frekuensi yang diharapkan dengan frekuensi yang diamati dalam kelompok-kelompok risiko yang berbeda. Jika nilai  $p\text{-value} > \alpha$ , maka model dianggap sesuai dan dapat digunakan untuk analisis lebih lanjut.

#### 10. Pengukuran Kebaikan Model

Setelah model terbukti sesuai dengan data, tahap selanjutnya adalah pengukuran kebaikan model untuk mengukur kekuatan penjelasan model. Pengukuran kebaikan model ini akan dilakukan dengan menghitung nilai Pseudo R-Square, seperti Nagelkerke  $R^2$  berdasarkan persamaan (2.18). Nilai ini akan menunjukkan seberapa besar persentase variasi variabel dependen yang mampu dijelaskan oleh variabel-variabel independen yang ada di dalam model.

#### 11. Evaluasi Kinerja Model dan *Confusion Matrix*

Tahap selanjutnya adalah evaluasi kinerja model, akan diukur secara komprehensif menggunakan metrik keseluruhan dan per-kategori. Secara keseluruhan, Akurasi (*Accuracy*) akan dihitung untuk mengukur proporsi total prediksi yang benar berdasarkan persamaan (2.20), dilengkapi dengan *Micro-Average Metrics* untuk mendapatkan skor tunggal *precision*, *recall*, dan *F1-score* dari agregasi semua kelas. Untuk analisis yang lebih mendalam, akan dihitung juga metrik untuk setiap kategori secara spesifik, yang meliputi Presisi per Kelas berdasarkan persamaan (2.21), *Recall* per Kelas berdasarkan

persamaan (2.22), serta *F1-Score* per Kelas berdasarkan persamaan (2.23). Kombinasi metrik ini akan menghasilkan analisis yang utuh mengenai kekuatan dan kelemahan model secara detail.

## 12. Interpretasi Model

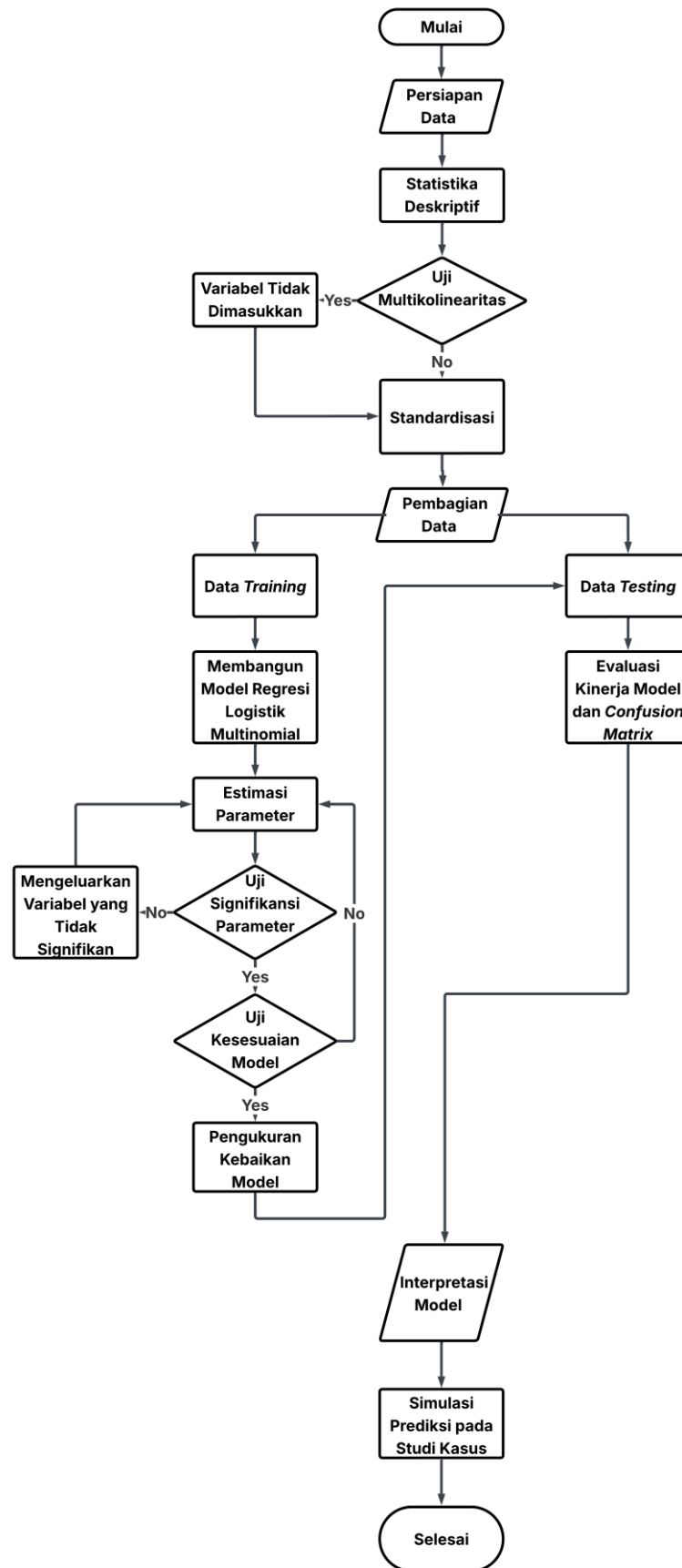
Tahap terakhir adalah menginterpretasikan hasil dari model regresi logistik multinomial dengan menghitung odds ratio berdasarkan persamaan (2.19). Koefisien regresi diubah menjadi odds ratio untuk mengukur pengaruh variabel prediktor terhadap probabilitas kategori ketepatan waktu kelulusan. Hasil ini penting untuk memberikan informasi yang mendukung evaluasi kebijakan sistem akademik dalam upaya menambah jumlah mahasiswa yang lulus tepat waktu sehingga memperbaiki penilaian kualitas dan akreditasi Program Studi Matematika UIN Maulana Malik Ibrahim Malang.

## 13. Simulasi Prediksi pada Studi Kasus

Tahap terakhir adalah mengilustrasikan penerapan praktis dari model yang telah terbentuk. Pada tahap ini, akan dibuat sebuah studi kasus dari data *testing* yang merepresentasikan seorang mahasiswa dengan karakteristik tertentu (misalnya: IPK 3.70, jenis kelamin perempuan, asal sekolah MA, jalur masuk SNMPTN, pembiayaan non-bidikmisi, asal daerah dari luar kota Malang, rasio SKS lulus 0.85, dan ketepatan waktu kelulusan tepat waktu). Nilai-nilai dari variabel independen ini akan disubstitusikan ke dalam persamaan model regresi logistik multinomial yang telah diestimasi. Hasil dari simulasi ini adalah nilai probabilitas bagi mahasiswa tersebut untuk masuk ke dalam setiap kategori kelulusan:  $P(Y=\text{Tepat Waktu})$ ,  $P(Y=\text{Terlambat})$ , dan  $P(Y=\text{Kritis})$ . Kategori dengan probabilitas tertinggi

akan menjadi hasil akhir prediksi model untuk studi kasus tersebut. Langkah ini bertujuan untuk menunjukkan bagaimana model dapat digunakan sebagai alat bantu prediksi di dunia nyata.

Tahapan-tahapan penelitian tersebut divisualisasikan dalam *flowchart* pada Gambar 3.1.



**Gambar 3.1** *Flowchart* Penelitian



## BAB IV

### HASIL DAN PEMBAHASAN

#### 4.1 Persiapan Data

Persiapan data diawali dengan pengumpulan data sekunder yang bersumber dari Program Studi Matematika UIN Maulana Malik Ibrahim Malang. Populasi penelitian ini adalah seluruh mahasiswa angkatan 2018-2021 yang berjumlah 336 orang. Dari populasi tersebut, dilakukan proses pembersihan data untuk menangani data yang tidak lengkap, sehingga diperoleh populasi akhir yang valid dan siap dianalisis sebanyak 331 orang.

Selanjutnya adalah pengkodean variabel prediktor kategorik. Berdasarkan pada Tabel 3.1, dilakukan proses pengkodean dengan mengkodekan variabel-variabel prediktor kategorik meliputi jenis kelamin, jenis asal sekolah, jenis pembiayaan, jalur masuk, dan asal daerah menjadi *dummy variable*. Kategori referensi pada setiap variabel kategorik ditetapkan berdasarkan kategori dengan frekuensi terbesar (kategori mayoritas) pada masing-masing variabel. Dengan demikian, kategori yang digunakan sebagai referensi adalah:

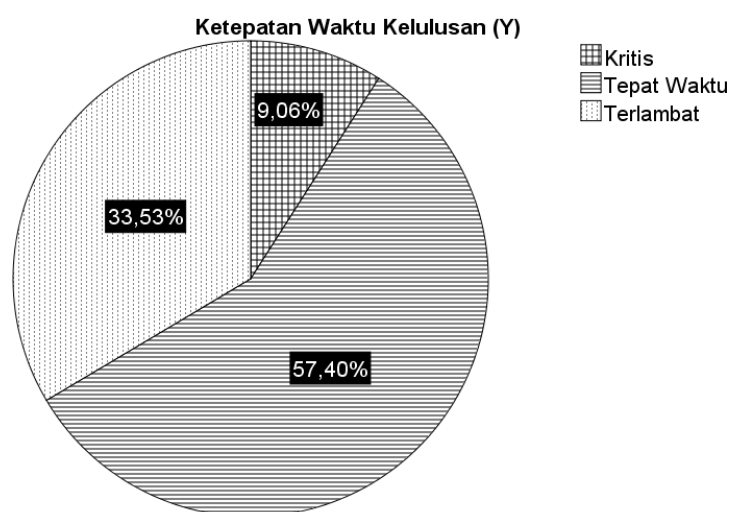
1. Kategori 2 (Perempuan) untuk variabel Jenis Kelamin,
2. Kategori 2 (SMA) untuk variabel Jenis Asal Sekolah,
3. Kategori 2 (Non Bidikmisi) untuk variabel Jenis Pembiayaan,
4. Kategori 2 (SBMPTN) untuk variabel Jalur Masuk, dan
5. Kategori 2 (Luar Kota Malang) untuk variabel Asal Daerah.

Dalam proses pembentukan *dummy*, setiap kategori referensi dikodekan sebagai 0, sedangkan kategori lainnya dikodekan sebagai 1, sehingga model dapat

membandingkan pengaruh masing-masing kategori terhadap kategori referensi. Hasil lengkap pengkodean *dummy variable* untuk seluruh variabel prediktor kategorik dapat dilihat pada Lampiran 1 pada nomor 2.

## 4.2 Statistika Deskriptif

Penelitian ini menggunakan data sekunder mahasiswa Program Studi Matematika UIN Malang angkatan 2018-2021 yang diperoleh dari Program Studi Matematika UIN Malang. Dataset terdiri dari 331 mahasiswa setelah proses pembersihan data, dengan informasi yang dikumpulkan meliputi 7 variabel prediktor IPK Semester 1-6 ( $X_1$ ), Jenis Kelamin ( $X_2$ ), Jenis Asal Sekolah ( $X_3$ ), Jenis Pembiayaan ( $X_4$ ), Jalur Masuk ( $X_5$ ), Asal Daerah ( $X_6$ ), dan Rasio SKS Lulus ( $X_7$ ). Sementara itu, variabel respon dalam data ini adalah Ketepatan Waktu Kelulusan ( $Y$ ) yang memiliki kategori Tepat Waktu (7-8 semester), Terlambat (9-10 semester), dan Kritis ( $\geq 11$  semester). Deskripsi mengenai ketepatan waktu kelulusan di Program Studi Matematika UIN Malang disajikan dalam Gambar 4.1.



**Gambar 4.1** Ketepatan Waktu Kelulusan

Berdasarkan Gambar 4.1, persentase mahasiswa yang lulus Tepat Waktu lebih banyak yaitu sebesar 57,4% atau sebanyak 190 mahasiswa dalam klasifikasi Tepat Waktu. Sedangkan persentase mahasiswa yang Terlambat sebesar 33,53% atau sebanyak 111 mahasiswa, dan yang Kritis sebesar 9,06% atau sebanyak 30 mahasiswa. Hal tersebut menunjukkan bahwa sekitar 4 dari 10 mahasiswa (42,6%) dalam sampel ini tidak lulus tepat waktu (Terlambat atau Kritis). Meskipun mayoritas mahasiswa lulus tepat waktu, persentase mahasiswa yang tidak lulus tepat waktu yang mencapai lebih dari 40% menandakan adanya masalah terkait durasi studi yang perlu mendapat perhatian serius. Hal tersebut yang dapat menjadi dasar untuk melakukan analisis lebih lanjut untuk mengidentifikasi faktor-faktor yang berpengaruh terhadap ketepatan waktu kelulusan mahasiswa, seperti IPK, jenis kelamin, jenis asal sekolah, jenis pembiayaan, jalur masuk, asal daerah, dan rasio SKS lulus.

#### 4.2.1 Deskripsi Variabel Prediktor Numerik

IPK ( $x_1$ ) dan Rasio SKS Lulus ( $x_7$ ) merupakan dua faktor akademik yang dianalisis pengaruhnya terhadap ketepatan waktu kelulusan mahasiswa. Distribusi IPK kumulatif hingga semester enam serta distribusi Rasio SKS Lulus pada sampel penelitian disajikan pada Tabel 4.1 sebagai gambaran umum karakteristik variabel prediktor numerik yang digunakan dalam penelitian ini.

**Tabel 4.1** Variabel Prediktor Numerik

Variabel	Rata-Rata	Standar Deviasi	Minimum	Maksimum
$x_1$	3,5022	0,18968	2,47	3,92
$x_7$	0,8943	0,03147	0,69	0,93

Berdasarkan Tabel 4.1, dapat dilihat bahwa nilai rata-rata IPK mahasiswa adalah 3,5022. Nilai rata-rata ini menunjukkan bahwa secara umum, mahasiswa dalam sampel ini memiliki prestasi akademik yang tergolong sangat baik, mendekati predikat *cum laude*. Untuk melihat sebaran data IPK di sekitar nilai rata-ratanya, disajikan standar deviasi sebesar 0,18968. Nilai standar deviasi yang relatif kecil ini mengindikasikan bahwa sebagian besar nilai IPK mahasiswa cenderung terkonsentrasi di sekitar nilai rata-rata, menunjukkan tingkat variasi yang tidak terlalu besar dalam prestasi akademik di antara sampel. Nilai IPK minimum yang tercatat adalah 2,47, sementara nilai IPK maksimum mencapai 3,92. Rentang ini menggambarkan adanya variasi dalam pencapaian akademik, mulai dari mahasiswa dengan IPK yang cukup hingga mahasiswa dengan IPK yang sangat tinggi mendekati sempurna.

Sementara itu, nilai rata-rata Rasio SKS Lulus mahasiswa adalah 0,8943. Nilai rata-rata yang tinggi ini menunjukkan bahwa secara umum, mahasiswa dalam sampel ini telah menyelesaikan sebagian besar (hampir 90%) dari total SKS yang disyaratkan hingga semester 6, mengindikasikan progresi studi yang baik. Untuk melihat sebaran data Rasio SKS Lulus di sekitar nilai rata-ratanya, disajikan standar deviasi sebesar 0,03147. Nilai standar deviasi yang sangat kecil ini mengindikasikan bahwa sebagian besar nilai Rasio SKS Lulus mahasiswa sangat terkonsentrasi di sekitar nilai rata-rata, menunjukkan tingkat variasi yang sangat rendah dalam hal kemajuan penyelesaian SKS di antara sampel. Nilai Rasio SKS Lulus minimum yang tercatat adalah 0,69, sementara nilai Rasio SKS Lulus maksimum mencapai 0,93. Rentang ini menggambarkan bahwa meskipun ada variasi, semua mahasiswa dalam sampel telah menyelesaikan setidaknya 69%

SKS yang disyaratkan, dengan beberapa mahasiswa hampir menyelesaikan seluruh SKS non-skripsi hingga semester 6.

#### **4.2.2 Deskripsi Variabel Prediktor Kategorik**

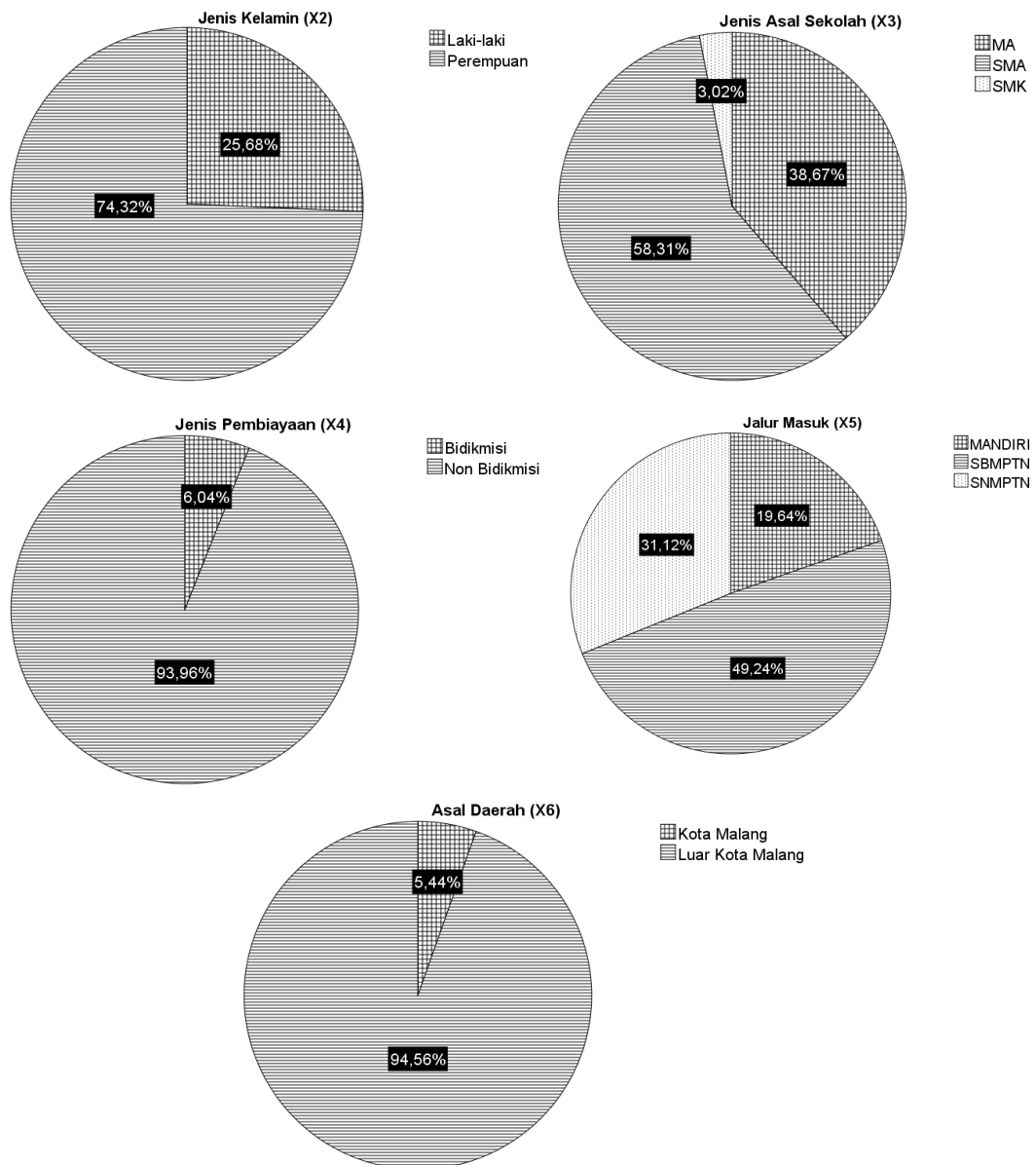
Variabel prediktor kategorik yang digunakan dalam penelitian ini terdiri dari Jenis Kelamin, Jenis Asal Sekolah, Jenis Pembiayaan, Jalur Masuk, dan Asal Daerah. Deskripsi distribusi frekuensi dan persentase untuk masing-masing variabel tersebut disajikan pada Gambar 4.2.

Berdasarkan Gambar 4.2, terlihat bahwa mahasiswa yang berjenis kelamin Perempuan lebih banyak dibandingkan dengan mahasiswa Laki-laki. Dari total 331 mahasiswa, sebesar 74,3% atau sebanyak 246 mahasiswa berjenis kelamin Perempuan. Proporsi ini menunjukkan bahwa komposisi mahasiswa pada Program Studi Matematika didominasi oleh perempuan, sedangkan mahasiswa Laki-laki hanya sebesar 25,7% atau sebanyak 85 orang.

Distribusi jenis asal sekolah mahasiswa menunjukkan bahwa mayoritas responden berasal dari SMA dibandingkan dengan MA maupun SMK. Dari 331 mahasiswa, sebesar 58,3% atau 193 mahasiswa berasal dari SMA, sedangkan 38,7% atau 128 mahasiswa berasal dari MA. Sisanya sebesar 3,0% atau sebanyak 10 mahasiswa berasal dari SMK, sehingga dapat disimpulkan bahwa SMA merupakan jalur asal sekolah yang paling dominan pada populasi penelitian.

Distribusi jenis pembiayaan menunjukkan bahwa mahasiswa dengan pembiayaan Non Bidikmisi jauh lebih banyak dibandingkan mahasiswa penerima Bidikmisi. Dari total 331 mahasiswa, sebesar 94,0% atau sebanyak 311 mahasiswa merupakan pengguna pembiayaan Non Bidikmisi. Sementara itu,

hanya 6,0% atau sebanyak 20 mahasiswa yang merupakan penerima Bidikmisi, sehingga kategori Non Bidikmisi menjadi kelompok yang sangat dominan dalam populasi ini.



**Gambar 4.2** Variabel Prediktor Kategorik

Distribusi jalur masuk memperlihatkan bahwa mahasiswa yang diterima melalui jalur SBMPTN jauh lebih banyak dibandingkan SNMPTN dan Mandiri.

Sebesar 49,2% atau 163 mahasiswa masuk melalui jalur SBMPTN, sedangkan 31,1% atau 103 mahasiswa berasal dari jalur SNMPTN. Adapun sisanya, sebesar 19,6% atau 65 mahasiswa, masuk melalui jalur Mandiri, yang menjadikannya kategori dengan proporsi paling kecil.

Distribusi asal daerah menunjukkan bahwa mahasiswa yang berasal dari Luar Kota Malang jauh lebih banyak dibandingkan mahasiswa dari Kota Malang. Dari total 331 mahasiswa, sebesar 94,6% atau sebanyak 313 mahasiswa berasal dari luar wilayah Kota Malang. Sementara itu, hanya 5,4% atau sebanyak 18 mahasiswa yang berasal dari Kota Malang, sehingga dapat dikatakan bahwa sebagian besar mahasiswa program studi ini merupakan pendatang dari luar daerah.

Deskripsi lebih lanjut mengenai variabel prediktor kategorik, meliputi jenis kelamin, jenis asal sekolah, jenis pembiayaan, jalur masuk, dan asal daerah terhadap ketepatan waktu kelulusan disajikan dalam tabulasi silang pada Tabel 4.2.

Tabel 4.2 menunjukkan deskripsi jenis kelamin berdasarkan ketepatan waktu kelulusan mahasiswa dalam sampel penelitian. Untuk jenis kelamin Laki-laki (total 85 mahasiswa), sebanyak 50,6% (43 mahasiswa) lulus Tepat Waktu, 36,5% (31 mahasiswa) lulus Terlambat, dan 12,9% (11 mahasiswa) lulus Kritis. Sementara itu, untuk jenis kelamin Perempuan (total 246 mahasiswa), sebanyak 59,8% (147 mahasiswa) lulus Tepat Waktu, 32,5% (80 mahasiswa) lulus Terlambat, dan 7,7% (19 mahasiswa) lulus Kritis. Oleh karena itu, secara proporsional, persentase mahasiswa yang lulus Tepat Waktu lebih tinggi pada kelompok Perempuan (59,8%) dibandingkan Laki-laki (50,6%). Sebaliknya,

persentase mahasiswa yang lulus tidak tepat waktu (gabungan Terlambat dan Kritis) lebih tinggi pada kelompok Laki-laki (49,4%) dibandingkan Perempuan (40,2%).

**Tabel 4.2** Tabulasi Silang dari Variabel Prediktor Kategorik terhadap Ketepatan Waktu Kelulusan

		Ketepatan Waktu Kelulusan			Total
		Tepat Waktu	Terlambat	Kritis	
Jenis Kelamin	Laki-laki	43 [50,6%]	31 [36,5%]	11 [12,9%]	85 [100%]
	Perempuan	147 [59,8%]	80 [32,5%]	19 [7,7%]	246 [100%]
Jenis Asal Sekolah	MA	72 [56,3%]	47 [36,7%]	9 [7%]	128 [100%]
	SMA	111 [57,5%]	62 [32,1%]	20 [10,4%]	193 [100%]
	SMK	7 [70%]	2 [20%]	1 [10%]	10 [100%]
Jenis Pembiayaan	Bidikmisi	14 [70%]	5 [25%]	1 [5%]	20 [100%]
	Non Bidikmisi	176 [56,6%]	106 [34,1%]	29 [9,3%]	311 [100%]
Jalur Masuk	SNMPTN	31 [47,7%]	26 [40%]	8 [12,3%]	65 [100%]
	SBMPTN	101 [62%]	46 [28,2%]	16 [9,8%]	163 [100%]
	MANDIRI	58 [56,3%]	39 [37,9%]	6 [5,8%]	103 [100%]
Asal Daerah	Kota Malang	12 [66,7%]	6 [33,3%]	0 [0%]	18 [100%]
	Luar Kota Malang	178 [56,9%]	105 [33,5%]	30 [9,6%]	313 [100%]

Selanjutnya, Tabel 4.2 juga menunjukkan deskripsi jenis asal sekolah berdasarkan ketepatan waktu kelulusan mahasiswa dalam sampel penelitian. Terlihat dari tabel tersebut, untuk mahasiswa yang berasal dari MA (total 128 mahasiswa), sebanyak 56,3% (72 mahasiswa) lulus Tepat Waktu, 36,7% (47 mahasiswa) lulus Terlambat, dan 7,0% (9 mahasiswa) lulus Kritis. Sementara itu,



untuk mahasiswa yang berasal dari SMA (total 193 mahasiswa), sebanyak 57,5% (111 mahasiswa) lulus Tepat Waktu, 32,1% (62 mahasiswa) lulus Terlambat, dan 10,4% (20 mahasiswa) lulus Kritis. Adapun untuk mahasiswa yang berasal dari SMK (total 10 mahasiswa), sebanyak 70,0% (7 mahasiswa) lulus Tepat Waktu, 20,0% (2 mahasiswa) lulus Terlambat, dan 10,0% (1 mahasiswa) lulus Kritis. Oleh karena itu, secara proporsional, persentase mahasiswa yang lulus Tepat Waktu paling tinggi pada kelompok yang berasal dari SMK (70,0%). Sebaliknya, persentase mahasiswa yang lulus tidak tepat waktu (gabungan Terlambat dan Kritis) paling tinggi pada kelompok yang berasal dari MA (43,7%).

Selanjutnya, Tabel 4.2 juga menunjukkan deskripsi jenis pembiayaan berdasarkan ketepatan waktu kelulusan mahasiswa dalam sampel penelitian. Terlihat dari tabel tersebut, untuk mahasiswa penerima Bidikmisi (total 20 mahasiswa), sebanyak 70,0% (14 mahasiswa) lulus Tepat Waktu, 25,0% (5 mahasiswa) lulus Terlambat, dan 5,0% (1 mahasiswa) lulus Kritis. Sementara itu, untuk mahasiswa Non Bidikmisi (total 311 mahasiswa), sebanyak 56,6% (176 mahasiswa) lulus Tepat Waktu, 34,1% (106 mahasiswa) lulus Terlambat, dan 9,3% (29 mahasiswa) lulus Kritis. Oleh karena itu, secara proporsional, persentase mahasiswa yang lulus Tepat Waktu lebih tinggi pada kelompok Bidikmisi (70,0%) dibandingkan Non Bidikmisi (56,6%). Sebaliknya, persentase mahasiswa yang lulus tidak tepat waktu (gabungan Terlambat dan Kritis) lebih tinggi pada kelompok Non Bidikmisi (43,4%) dibandingkan Bidikmisi (30,0%).

Selanjutnya, Tabel 4.2 juga menunjukkan deskripsi jalur masuk berdasarkan ketepatan waktu kelulusan mahasiswa dalam sampel penelitian. Terlihat dari tabel tersebut, untuk mahasiswa yang masuk melalui jalur SNMPTN

(total 65 mahasiswa), sebanyak 47,7% (31 mahasiswa) lulus Tepat Waktu, 40,0% (26 mahasiswa) lulus Terlambat, dan 12,3% (8 mahasiswa) lulus Kritis. Sementara itu, untuk mahasiswa yang masuk melalui jalur SBMPTN (total 163 mahasiswa), sebanyak 62,0% (101 mahasiswa) lulus Tepat Waktu, 28,2% (46 mahasiswa) lulus Terlambat, dan 9,8% (16 mahasiswa) lulus Kritis. Adapun untuk mahasiswa yang masuk melalui jalur MANDIRI (total 103 mahasiswa), sebanyak 56,3% (58 mahasiswa) lulus Tepat Waktu, 37,9% (39 mahasiswa) lulus Terlambat, dan 5,8% (6 mahasiswa) lulus Kritis. Oleh karena itu, secara proporsional, persentase mahasiswa yang lulus Tepat Waktu paling tinggi pada kelompok yang masuk melalui jalur SBMPTN (62,0%). Sebaliknya, persentase mahasiswa yang lulus tidak tepat waktu (gabungan Terlambat dan Kritis) paling tinggi pada kelompok yang masuk melalui jalur SNMPTN (52,3%).

Terakhir, Tabel 4.2 juga menunjukkan deskripsi asal daerah berdasarkan ketepatan waktu kelulusan mahasiswa dalam sampel penelitian. Terlihat dari tabel tersebut, untuk mahasiswa yang berasal dari Kota Malang (total 18 mahasiswa), sebanyak 66,7% (12 mahasiswa) lulus Tepat Waktu, 33,3% (6 mahasiswa) lulus Terlambat, dan 0% (0 mahasiswa) lulus Kritis. Sementara itu, untuk mahasiswa yang berasal dari Luar Kota Malang (total 313 mahasiswa), sebanyak 56,9% (178 mahasiswa) lulus Tepat Waktu, 33,5% (105 mahasiswa) lulus Terlambat, dan 9,6% (30 mahasiswa) lulus Kritis. Oleh karena itu, secara proporsional, persentase mahasiswa yang lulus Tepat Waktu lebih tinggi pada kelompok yang berasal dari Kota Malang (66,7%) dibandingkan Luar Kota Malang (56,9%). Sebaliknya, persentase mahasiswa yang lulus tidak tepat waktu (gabungan Terlambat dan

Kritis) lebih tinggi pada kelompok yang berasal dari Luar Kota Malang (43,1%) dibandingkan Kota Malang (33,3%).

### 4.3 Uji Multikolinearitas

Uji multikolinearitas dilakukan untuk mengetahui ada atau tidaknya hubungan linear yang kuat antar variabel prediktor pada model regresi. Pengujian ini dilakukan menggunakan persamaan (2.2) untuk mengetahui nilai VIF.

1. Hipotesis:

$H_0$ : Tidak terdapat multikolinearitas antar variabel prediktor

$H_1$ : Terdapat multikolinearitas antar variabel prediktor

2. Taraf signifikansi:

Pengujian menggunakan taraf signifikansi  $\alpha = 5\%$  (0,05)

3. Daerah kritis:

Tolak  $H_0$  jika  $VIF > 10$  yang artinya terjadi pelanggaran multikolinearitas.

Nilai VIF untuk setiap variabel prediktor disajikan pada Tabel 4.3.

**Tabel 4.3** Nilai VIF

<b>Prediktor</b>	<b>VIF</b>
$x_{1IPK}$	1,970
$x_{2Ik}$	1,072
$x_{3MA}$	1,096
$x_{3SMK}$	1,048
$x_{4Bidikmisi}$	1,052
$x_{5SNMPTN}$	1,227
$x_{5Mandiri}$	1,267
$x_{6Kota\ Malang}$	1,029
$x_{7Rasio\ SKS\ Lulus}$	1,917

Berdasarkan Tabel 4.3 dapat diketahui bahwa nilai VIF untuk masing-masing variabel prediktor kurang dari 10 maka gagal tolak  $H_0$ , sehingga dapat disimpulkan bahwa pada data yang digunakan tidak terdapat gejala multikolinearitas antar variabel prediktor.

#### 4.4 Standardisasi

Standardisasi dilakukan bertujuan untuk menyeragamkan skala pengukuran data sehingga memiliki nilai rata-rata nol dan simpangan baku satu (*Z-score normalization*), guna meminimalisir bias yang mungkin timbul akibat perbedaan rentang nilai antar variabel dan membantu konvergensi model. Proses standardisasi ini dilakukan dengan mengacu pada metode *Z-score* seperti yang ditunjukkan pada persamaan (2.1).

Sebagai contoh, berikut adalah perhitungan manual standardisasi untuk variabel IPK ( $x_1$ ). Proses ini menggunakan nilai rata-rata IPK (3,5022) dan standar deviasi IPK (0,18968) yang diperoleh dari hasil statistika deskriptif pada Tabel 4.1, untuk seorang mahasiswa dengan profil IPK: 3,70.

$$\begin{aligned} Z_i &= \frac{X_i - \bar{X}}{S} \\ &= \frac{3,70 - 3,5022}{0,18968} \\ &= \frac{0,1978}{0,18968} \\ &= 1,04 \end{aligned}$$

Dengan cara yang sama, proses standardisasi ini diterapkan untuk semua variabel numerik (IPK dan rasio SKS lulus) pada seluruh data. Hasil lengkap

standardisasi yang digunakan dalam penelitian dapat dilihat pada Lampiran 1 pada nomor 3.

#### 4.5 Pembagian Data *Training* dan *Testing*

Pembagian data dilakukan dengan menggunakan teknik *stratified random sampling* untuk menjaga keseimbangan proporsi kategori variabel dependen pada setiap subset data. Data dibagi menjadi data *training* sebesar 80% dan data *testing* sebesar 20% dari total observasi. Pemilihan rasio 80:20 didasarkan pada pertimbangan metodologis bahwa proporsi tersebut mampu memberikan kompromi optimal antara kompleksitas model dan kemampuan generalisasi, sehingga menghasilkan akurasi dan kestabilan model yang lebih baik dibandingkan dengan rasio pembagian lainnya. Proses pembagian data dilakukan secara acak terstratifikasi menggunakan *cloud service* Google Colab, dengan hasil pembagian data disajikan pada Tabel 4.4.

**Tabel 4.4** Pembagian Data *Training* dan *Testing*

		Pembagian Data		Total
		Training (80%)	Testing (20%)	
Ketepatan Waktu Kelulusan	Tepat Waktu	152	38	190
	Terlambat	88	23	111
	Kritis	24	6	30
Total		264	67	331

Tabel 4.4 menunjukkan pembagian data untuk proses *training* dan *testing* terhadap data ketepatan waktu kelulusan. Data keseluruhan berjumlah 331, yang kemudian dibagi menjadi dua bagian yaitu 80% untuk data *training* (sebanyak 264 data) dan 20% untuk data *testing* (sebanyak 67 data). Berdasarkan total 190 data

pada kategori Tepat Waktu, sebanyak 152 digunakan untuk *training* dan 38 digunakan untuk *testing*. Selanjutnya, dari 111 data pada kategori Terlambat, 88 digunakan untuk *training* dan 23 untuk *testing*. Sementara itu, dari 30 data pada kategori Kritis, 24 digunakan untuk *training* dan 6 untuk *testing*. Hasil lengkap pembagian data yang digunakan dalam penelitian dapat dilihat pada Lampiran 1 pada nomor 4.

#### **4.6 Membangun Model Regresi Logistik Multinomial**

Pemodelan regresi logistik multinomial pada penelitian ini menggunakan model *main effect* dengan metode *enter*. Model *main effect* digunakan karena penelitian ini hanya mempertimbangkan pengaruh langsung dari masing-masing variabel independen terhadap variabel dependen tanpa adanya interaksi antarvariabel. Dengan demikian, setiap variabel bebas diasumsikan berkontribusi secara independen terhadap peluang perubahan kategori variabel dependen. Pendekatan ini sejalan dengan panduan analisis yang dikemukakan oleh Hosmer, Lemeshow, dan Sturdivant (2013), yang menyatakan bahwa model *main effect* lebih sesuai digunakan pada penelitian inferensial yang berfokus pada hubungan sebab-akibat antara faktor-faktor utama terhadap respon.

Metode *enter* digunakan pada tahap awal proses estimasi dengan memasukkan seluruh variabel independen ke dalam model secara bersamaan tanpa seleksi awal berbasis nilai signifikansi. Tahap ini menghasilkan model awal yang bersifat eksploratif untuk menggambarkan kontribusi masing-masing variabel secara simultan. Selanjutnya, untuk memperoleh model yang lebih efisien dan *parsimonious*, dilakukan proses eliminasi variabel secara sistematis dengan

mengeluarkan variabel-variabel yang tidak menunjukkan signifikansi statistik berdasarkan nilai *p-value* pada uji Wald. Proses ini diulang hingga diperoleh model yang hanya menyertakan variabel dengan kontribusi signifikan.

Model awal regresi logistik multinomial tersebut kemudian disederhanakan melalui proses eliminasi variabel, sehingga menghasilkan model akhir yang lebih stabil dan layak diinterpretasikan. Model akhir ini memastikan bahwa hanya variabel signifikan yang dipertahankan dalam analisis, sehingga interpretasi yang dihasilkan lebih akurat dalam menggambarkan hubungan antara variabel independen dan kategori ketepatan waktu kelulusan.

Berbeda dengan regresi biner yang hanya memiliki satu fungsi logit, model regresi logistik multinomial ini memerlukan  $(k-1)$  persamaan, di mana  $k$  adalah jumlah kategori variabel dependen. Pada penelitian ini, kategori 0 (Tepat Waktu) ditetapkan sebagai kategori referensi sehingga model membentuk dua fungsi logit. Berdasarkan Persamaan (2.4) dan (2.5), bentuk umum model logit untuk penelitian ini adalah sebagai berikut:

Logit ke-1 (Terlambat vs Tepat Waktu):

$$\begin{aligned} g_1(x) &= \ln \left[ \frac{P(Y = 1)}{P(Y = 0)} \right] \\ &= \beta_{10} + \beta_{11}x_1 + \beta_{12}x_2 + \beta_{13}x_3 + \beta_{14}x_4 + \beta_{15}x_5 + \beta_{16}x_6 + \beta_{17}x_7 \end{aligned}$$

Logit ke-2 (Kritis vs Tepat Waktu):

$$\begin{aligned} g_2(x) &= \ln \left[ \frac{P(Y = 2)}{P(Y = 0)} \right] \\ &= \beta_{20} + \beta_{21}x_1 + \beta_{22}x_2 + \beta_{23}x_3 + \beta_{24}x_4 + \beta_{25}x_5 + \beta_{26}x_6 + \beta_{27}x_7 \end{aligned}$$

#### 4.7 Estimasi Parameter Model Regresi Logistik Multinomial

Estimasi parameter dalam regresi logistik multinomial bertujuan untuk mengetahui arah serta besar pengaruh variabel independen terhadap peluang kategori variabel dependen. Proses estimasi dilakukan menggunakan metode *Maximum Likelihood Estimation (MLE)*, yang bertujuan memaksimalkan peluang (*likelihood*) agar model yang diperoleh memiliki tingkat kesesuaian terbaik dengan data observasi. Penentuan nilai koefisien ( $\beta$ ) dilakukan berdasarkan persamaan (2.14). Nilai koefisien ( $\beta$ ) menunjukkan arah pengaruh, di mana tanda positif mengindikasikan peningkatan peluang terhadap kategori tertentu, sedangkan tanda negatif menunjukkan penurunan peluang dibandingkan kategori referensi. Hasil estimasi parameter disajikan pada Tabel 4.5 berikut.

**Tabel 4.5** Estimasi Parameter

Logit	Variabel	Parameter	Estimasi	Std. Error
1	<i>Intercept</i>	$\beta_{10}$	-0,678	0,267
	$x_1$ (IPK std)	$\beta_{11}$	-1,143	0,239
	$x_2$ Laki-laki	$\beta_{12}$	0,144	0,349
	$x_3$ MA	$\beta_{13}$	-0,111	0,317
	$x_3$ SMK	$\beta_{14}$	-1,087	0,935
	$x_4$ Bidikmisi	$\beta_{15}$	-1,134	0,729
	$x_5$ SNMPTN	$\beta_{16}$	0,423	0,343
	$x_5$ Mandiri	$\beta_{17}$	0,437	0,411
	$x_6$ KotaMlg	$\beta_{18}$	-0,232	0,593
	$x_7$ (Rasio SKS Lulus std)	$\beta_{19}$	-0,014	0,263
2	<i>Intercept</i>	$\beta_{20}$	-1,779	0,413
	$x_1$ (IPK std)	$\beta_{21}$	-0,967	0,392
	$x_2$ Laki-laki	$\beta_{22}$	0,343	0,571
	$x_3$ MA	$\beta_{23}$	-0,725	0,597
	$x_3$ SMK	$\beta_{24}$	-1,774	1,576
	$x_4$ Bidikmisi	$\beta_{25}$	-1,038	1,320
	$x_5$ SNMPTN	$\beta_{26}$	-0,354	0,657
	$x_5$ Mandiri	$\beta_{27}$	0,171	0,706
	$x_6$ KotaMlg	$\beta_{28}$	-20,05	0,000
	$x_7$ (Rasio SKS Lulus std)	$\beta_{29}$	-0,804	0,366



## 4.8 Uji Signifikansi Parameter

Pengujian signifikansi parameter terbagi menjadi dua, yaitu uji serentak atau simultan yang mengevaluasi signifikansi variabel prediktor keseluruhan terhadap variabel respon dan uji parsial yang mengevaluasi signifikansi tiap-tiap variabel prediktor secara terpisah.

### 4.8.1 Uji Signifikansi Parameter Secara Simultan

Uji simultan dilakukan untuk mengidentifikasi secara keseluruhan bagaimana variabel-variabel saling berhubungan dan berpengaruh dalam suatu model. Uji yang digunakan adalah *likelihood ratio test* atau uji  $G^2$ .

1. Hipotesis:

$H_0: \beta_1 = \beta_2 = \dots = \beta_7 = 0$  (tidak ada variabel prediktor yang memengaruhi model).

$H_1$ : Setidaknya ada satu  $\beta_i \neq 0$ , dengan  $i = 1, 2, \dots, 7$ .

2. Taraf signifikansi:

Pengujian menggunakan taraf signifikansi  $\alpha = 5\%$  (0,05)

3. Daerah kritis:

Jika nilai  $G^2 > \chi^2_{(v,\alpha)}$  atau  $p\text{-value} < \alpha$ , maka tolak  $H_0$ , artinya variabel prediktor secara bersama-sama berpengaruh signifikan terhadap variabel dependen.

4. Statistik uji:

Hasil dari uji signifikansi parameter secara simultan disajikan dalam Tabel 4.6. Perhitungan manual uji  $G$  berdasarkan persamaan (2.15) sebagai berikut:

$$\begin{aligned}
 G^2 &= -2 \ln \left[ \frac{\left(\frac{n_0}{n}\right)^{n_0} \left(\frac{n_1}{n}\right)^{n_1} \left(\frac{n_2}{n}\right)^{n_2}}{\prod_{i=1}^n [\pi_0(x_i)^{y_{0i}} \pi_1(x_i)^{y_{1i}} \pi_2(x_i)^{y_{2i}}]} \right] \\
 &= 473,511 - 387,369 \\
 &= 86,142
 \end{aligned}$$

**Tabel 4.6** Uji Signifikansi Simultan

	<i>-2 Log Likelihood</i>	<i>Df</i>	<i>Sig.</i>	<i>G<sup>2</sup> (χ<sup>2</sup>)</i>	<i>χ<sup>2</sup><sub>(v,α)</sub></i>
<b>Intercept Only</b>	473,511				
<b>Final</b>	387,369	18	0,000	86,142	28,869

Tabel 4.6 menunjukkan hasil uji signifikansi parameter secara simultan. Nilai statistik uji  $G^2$  ( $\chi^2$ ) diperoleh sebesar 86,142, yang merupakan selisih dari  $-2 \text{ Log Likelihood}$  model 'Intercept Only' (473,511) dan model 'Final' (387,369). Uji ini memiliki derajat kebebasan (Df) sebanyak 18 dan menghasilkan  $p\text{-value} = 0,000$ . Karena  $p\text{-value} = 0,000 < \alpha = 0,05$  dan nilai uji  $G^2 = 86,142 > \chi_{tabel}^2 = 28,869$ , maka keputusannya adalah tolak  $H_0$ . Hal ini berarti seluruh variabel prediktor yaitu secara bersama-sama (simultan) berpengaruh signifikan terhadap model ketepatan waktu kelulusan mahasiswa.

#### 4.8.2 Uji Signifikansi Parameter Secara Parsial

Uji parsial dilakukan untuk menguji signifikansi pengaruh masing-masing variabel prediktor terhadap variabel respon. Uji yang digunakan adalah uji Wald.

1. Hipotesis:

$H_0: \beta_i = 0$  (tidak ada pengaruh yang signifikan).

$H_1: \beta_i \neq 0$ , di mana  $i = 1, 2, \dots, 7$  (ada pengaruh yang signifikan).

2. Taraf signifikansi:

Pengujian menggunakan taraf signifikansi  $\alpha = 5\%$  (0,05)

3. Daerah kritis:

Jika nilai  $W > \chi^2_{(v,\alpha)}$  atau  $p\text{-value} < \alpha$ , maka tolak  $H_0$ , artinya variabel prediktor secara parsial berpengaruh signifikan terhadap variabel respon.

4. Statistik uji:

Hasil dari uji signifikansi parameter secara parsial disajikan dalam Tabel 4.7. Perhitungan nilai Wald untuk logit 1 berdasarkan persamaan (2.16) sebagai berikut:

$$\begin{aligned} W &= \left( \frac{\hat{\beta}_i}{\widehat{SE}(\hat{\beta}_i)} \right)^2 \\ W_0 &= \left( \frac{\hat{\beta}_0}{\widehat{SE}(\hat{\beta}_0)} \right)^2 \\ &= \left( \frac{-0,678}{0,267} \right)^2 \\ &= 6,454 \end{aligned}$$

Dengan cara yang sama, untuk semua variabel dari logit 1 dan logit 2, diperoleh hasil sebagaimana disajikan pada Tabel 4.7.

Berdasarkan Tabel 4.7, dapat diketahui bahwa pada logit 1 terdapat satu variabel prediktor yang berpengaruh signifikan terhadap variabel respon yaitu variabel  $x_1$  (IPK\_std), karena menghasilkan  $p\text{-value} = 0,000 < \alpha = 0,05$ . Dengan demikian keputusannya adalah tolak  $H_0$  yang artinya variabel prediktor IPK memiliki pengaruh signifikan terhadap ketepatan waktu kelulusan. Sedangkan pada logit 2, ada dua variabel prediktor yang berpengaruh signifikan terhadap variabel respon yaitu  $x_1$  (IPK\_std), karena menghasilkan  $p\text{-value} = 0,014 < \alpha = 0,05$  dan  $x_7$  (Rasio SKS Lulus\_std), karena menghasilkan  $p\text{-value}$

=  $0,028 < \alpha = 0,05$ . Dengan demikian keputusannya adalah tolak  $H_0$  yang artinya variabel prediktor IPK dan Rasio SKS Lulus memiliki pengaruh signifikan terhadap ketepatan waktu kelulusan.

**Tabel 4.7 Uji Signifikansi Parsial**

Logit	Parameter	Estimasi	Std. Error	Wald	Exp ( $\beta$ )	Sig.	Keputusan
1	$\beta_{10}$	-0,678	0,267	6,454	-	0,011	Tolak $H_0$
	$\beta_{11}$	-1,143	0,239	22,830	0,319	0,000	Tolak $H_0$
	$\beta_{12}$	0,144	0,349	0,170	1,155	0,680	Terima $H_0$
	$\beta_{13}$	-0,111	0,317	0,123	0,895	0,725	Terima $H_0$
	$\beta_{14}$	-1,087	0,935	1,351	0,337	0,245	Terima $H_0$
	$\beta_{15}$	-1,134	0,729	2,421	0,322	0,120	Terima $H_0$
	$\beta_{16}$	0,423	0,343	1,523	1,527	0,217	Terima $H_0$
	$\beta_{17}$	0,437	0,411	1,133	1,549	0,287	Terima $H_0$
	$\beta_{18}$	-0,232	0,593	0,153	0,793	0,695	Terima $H_0$
	$\beta_{19}$	-0,014	0,263	0,003	0,986	0,958	Terima $H_0$
2	$\beta_{20}$	-1,779	0,413	18,564	-	0,000	Tolak $H_0$
	$\beta_{21}$	-0,967	0,392	6,089	0,380	0,014	Tolak $H_0$
	$\beta_{22}$	0,343	0,571	0,361	1,409	0,548	Terima $H_0$
	$\beta_{23}$	-0,725	0,597	1,474	0,485	0,225	Terima $H_0$
	$\beta_{24}$	-1,774	1,576	1,267	0,170	0,260	Terima $H_0$
	$\beta_{25}$	-1,038	1,320	0,619	0,354	0,431	Terima $H_0$
	$\beta_{26}$	-0,354	0,657	0,290	0,702	0,590	Terima $H_0$
	$\beta_{27}$	0,171	0,706	0,059	1,187	0,808	Terima $H_0$
	$\beta_{28}$	-20,05	0,000	-	$1,96(10^{-9})$	-	-
	$\beta_{29}$	-0,804	0,366	4,815	0,448	0,028	Tolak $H_0$

Berdasarkan uji signifikansi parsial diketahui bahwa terdapat 2 variabel prediktor yang signifikan dan terdapat 5 variabel prediktor yang tidak signifikan. Hal tersebut menunjukkan bahwa perlu dilakukan uji signifikansi model revisi dengan hanya melibatkan 2 variabel prediktor yang signifikan. Pengujian kemudian diawali dengan uji signifikansi simultan (Uji  $G$ ) model revisi untuk mengetahui bagaimana kedua variabel prediktor tersebut secara bersama-sama memengaruhi variabel ketepatan waktu kelulusan. Hasil dari uji signifikansi simultan model revisi pada kedua variabel prediktor yang signifikan disajikan

pada Tabel 4.8. Perhitungan manual uji  $G$  berdasarkan persamaan (2.15) sebagai berikut:

$$\begin{aligned}
 G^2 &= -2 \ln \left[ \frac{\left(\frac{n_0}{n}\right)^{n_0} \left(\frac{n_1}{n}\right)^{n_1} \left(\frac{n_2}{n}\right)^{n_2}}{\prod_{i=1}^n [\pi_0(x_i)^{y_{0i}} \pi_1(x_i)^{y_{1i}} \pi_2(x_i)^{y_{2i}}]} \right] \\
 &= 473,511 - 400,489 \\
 &= 73,022
 \end{aligned}$$

**Tabel 4.8** Uji Signifikansi Simultan Model Revisi

	<i>-2 Log Likelihood</i>	<i>Df</i>	<i>Sig.</i>	$G^2 (\chi^2)$	$\chi^2_{(v,\alpha)}$
<b>Intercept Only</b>	473,511				
<b>Final</b>	400,489	4	0,000	73,022	9,488

Tabel 4.8 menunjukkan hasil uji signifikansi parameter secara simultan model revisi. Nilai statistik uji  $G^2 (\chi^2)$  diperoleh sebesar 73,022, yang merupakan selisih dari *-2 Log Likelihood* model 'Intercept Only' (473,511) dan model 'Final' (400,489). Uji ini memiliki derajat kebebasan (*Df*) sebanyak 4 dan menghasilkan  $p\text{-value} = 0,000$ . Karena  $p\text{-value} = 0,000 < \alpha = 0,05$  dan nilai uji  $G^2 = 73,022 > \chi^2_{tabel} = 9,488$ , maka keputusannya adalah tolak  $H_0$ . Hal ini berarti variabel prediktor yaitu IPK, dan Rasio SKS Lulus secara bersama-sama (simultan) berpengaruh signifikan terhadap model ketepatan waktu kelulusan mahasiswa.

Selanjutnya, dilakukan uji signifikansi parsial model revisi untuk melihat kontribusi individual dari masing-masing variabel prediktor. Hasil uji signifikansi parsial model revisi disajikan pada Tabel 4.9.

Tabel 4.9 Uji Signifikansi Parsial Model Revisi

Logit	Parameter	Estimasi	Std. Error	Wald	Exp ( $\beta$ )	Sig.	Keputusan
1	$\beta_{10}$	-0,575	0,152	14,355	-	0,000	Tolak $H_0$
	$\beta_{11}$	-1,186	0,232	26,122	0,306	0,000	Tolak $H_0$
	$\beta_{19}$	0,063	0,252	0,063	1,065	0,802	Terima $H_0$
2	$\beta_{20}$	-2,172	0,280	60,099	-	0,000	Tolak $H_0$
	$\beta_{21}$	-0,945	0,367	6,620	0,389	0,010	Tolak $H_0$
	$\beta_{29}$	-0,780	0,359	4,711	0,459	0,030	Tolak $H_0$

Berdasarkan Tabel 4.9, hasil uji signifikansi parsial model revisi menunjukkan adanya perbedaan kontribusi variabel antara Logit 1 dan Logit 2. Pada Logit 1 (Terlambat vs Tepat Waktu), hanya variabel  $x_1$  (IPK\_std) menghasilkan  $p\text{-value} = 0,000 < \alpha = 0,05$ , sehingga keputusannya adalah tolak  $H_0$ . Sementara itu, variabel  $x_7$  (Rasio SKS Lulus\_std) menghasilkan  $p\text{-value} = 0,802 > \alpha = 0,05$ , sehingga keputusannya adalah gagal tolak  $H_0$ . Sedangkan pada Logit 2 (Kritis vs Tepat Waktu), kedua variabel yaitu  $x_1$  (IPK\_std) dan  $x_7$  (Rasio SKS Lulus\_std) sama-sama signifikan, yaitu  $x_1$  (IPK\_std), karena menghasilkan  $p\text{-value} = 0,010 < \alpha = 0,05$  dan  $x_7$  (Rasio SKS Lulus\_std), karena menghasilkan  $p\text{-value} = 0,030 < \alpha = 0,05$ , sehingga keputusannya adalah tolak  $H_0$ . Hal ini menyimpulkan bahwa IPK adalah prediktor yang paling kuat dalam memengaruhi peluang kelulusan Terlambat, sementara IPK dan Rasio SKS Lulus merupakan prediktor yang paling kuat dan konsisten dalam memengaruhi peluang mahasiswa masuk kategori Kritis.

Berdasarkan Tabel 4.9 didapatkan model akhir regresi logistik multinomial untuk Logit 1 dan Logit 2 berdasarkan persamaan (2.4) dan (2.5) sebagai berikut:

Logit 1 (Terlambat vs Tepat Waktu):

$$g_1(x) = \ln \left[ \frac{P(Y = 1)}{P(Y = 0)} \right] \quad (4.1)$$

$$= -0,575 - 1,186 (x_{1_{IPK_{std}}})$$

Logit 2 (Kritis vs Tepat Waktu):

$$\begin{aligned} g_2(x) &= \ln \left[ \frac{P(Y = 2)}{P(Y = 0)} \right] \\ &= -2,172 - 0,945 (x_{1_{IPK_{std}}}) - 0,780 (x_{7_{Rasio\ SKSLulus_{std}}}) \end{aligned} \quad (4.2)$$

Berdasarkan persamaan (4.1) diperoleh hasil estimasi model logit pertama ( $g_1(x)$ ) untuk model revisi. Diketahui bahwa variabel IPK memiliki koefisien negatif sebesar -1,186 dan berpengaruh signifikan terhadap peluang mahasiswa mengalami keterlambatan kelulusan dibandingkan dengan kategori Tepat Waktu. Hal ini menunjukkan bahwa peningkatan nilai IPK menurunkan peluang mahasiswa berada pada kategori Terlambat dibandingkan dengan kategori Tepat Waktu. Sementara itu, variabel Rasio SKS Lulus tidak signifikan, sehingga secara statistik tidak dapat disimpulkan memiliki pengaruh terhadap ketepatan waktu kelulusan pada kategori Terlambat dibandingkan Tepat Waktu.

Selanjutnya, berdasarkan persamaan (4.2) didapatkan hasil estimasi model logit kedua ( $g_2(x)$ ) untuk model revisi. Diketahui bahwa variabel IPK dan Rasio SKS Lulus masing-masing memiliki koefisien negatif sebesar -0,945 dan -0,780 serta keduanya signifikan secara statistik. Hasil ini menunjukkan bahwa peningkatan IPK dan rasio SKS lulus menurunkan peluang mahasiswa berada pada kategori Kritis dibandingkan dengan kategori Tepat Waktu. Dengan demikian, mahasiswa dengan IPK yang lebih tinggi dan tingkat penyelesaian SKS yang lebih baik cenderung memiliki peluang lebih besar untuk lulus Tepat Waktu dibandingkan mengalami keterlambatan yang bersifat kritis.

#### 4.9 Uji Kesesuaian Model

Uji kesesuaian model dilakukan untuk menentukan apakah ditemukan ketidaksamaan antara data hasil observasi dengan prediksi yang dihasilkan model.

1. Hipotesis:

$H_0$ : Model sudah sesuai.

$H_1$ : Model tidak sesuai.

2. Taraf signifikansi:

Pengujian menggunakan taraf signifikansi  $\alpha = 5\%$  (0,05)

3. Statistik uji:

$$\hat{C} = \sum_{k=1}^g \frac{(O_k - n'_k \bar{\pi}_k)^2}{n'_k \bar{\pi}_k (1 - \bar{\pi}_k)}$$

4. Daerah kritis:

Jika nilai  $\hat{C} > \chi^2_{(v,\alpha)}$  atau  $p\text{-value} < \alpha$ , maka tolak  $H_0$ , artinya model belum sesuai atau terdapat ketidaksamaan antara hasil observasi dengan peluang perkiraan model.

Hasil dari uji kesesuaian model disajikan pada Tabel 4.10.

**Tabel 4.10** Uji Kesesuaian Model

	$\hat{C}$	<i>Df</i>	<i>P-Value</i>	$\chi^2_{(v,\alpha)}$
<i>Deviance</i>	397,716	498	1,000	551,023

Tabel 4.10 menunjukkan hasil uji kesesuaian model yaitu  $p\text{-value} = 1,000 > \alpha = 0,05$ . Dapat dilihat juga bahwa nilai *deviance* ( $\hat{C}$ ) sebesar  $397,716 < \chi^2_{tabel} = 551,023$ . Dengan demikian, keputusannya adalah gagal tolak  $H_0$  yang berarti model yang digunakan telah sesuai.



#### 4.10 Pengukuran Keباikan Model

Pengukuran kebaikan model dalam penelitian ini dilakukan untuk melihat seberapa besar kemampuan variabel-variabel independen dalam menjelaskan variasi pada variabel dependen (kategori ketepatan waktu kelulusan). Hasil dari pengukuran kebaikan model disajikan pada Tabel 4.11.

**Tabel 4.11** Pseudo R-Square

Cox & Snell	Nagelkerke
0,242	0,289

Tabel 4.11 menunjukkan beberapa nilai Pseudo R-Square yang dihasilkan. Nilai yang digunakan untuk interpretasi adalah Nagelkerke R Square, yang menunjukkan nilai sebesar 0,289 atau 28,9%. Hal ini mengindikasikan bahwa sebesar 28,9% variasi dari variabel dependen (ketepatan waktu kelulusan) dapat dijelaskan oleh IPK ( $x_1$ ), Jenis Kelamin ( $x_2$ ), Jenis Asal Sekolah ( $x_3$ ), Jenis Pembiayaan ( $x_4$ ), Jalur Masuk ( $x_5$ ), Asal Daerah ( $x_6$ ), dan Rasio SKS Lulus ( $x_7$ ). Sedangkan sisanya sebesar 71,1% dipengaruhi oleh variabel-variabel lain yang tidak dimasukkan ke dalam model penelitian ini.

#### 4.11 Evaluasi Kinerja Model

Setelah model regresi logistik multinomial dibangun dan divalidasi menggunakan data training, langkah selanjutnya adalah melakukan evaluasi untuk menghitung kinerja model dalam melakukan generalisasi dan memprediksi data baru yang belum pernah dilihat sebelumnya. Evaluasi ini dilakukan dengan menguji model pada data *testing* yang belum pernah dilihat sebelumnya (20% dari sampel). Evaluasi ini menggunakan *multiclass confusion matrix* untuk mendapatkan nilai

numerik metrik kinerja keseluruhan (Akurasi) dan metrik per-kategori (*Precision*, *Recall*, dan *F1-Score*). Tabel 4.12 menyajikan *confusion matrix* dari data testing.

**Tabel 4.12** *Confusion Matrix*

		Prediksi			Total Aktual
		0	1	2	
Aktual	0	34	4	0	38
	1	15	8	0	23
	2	3	2	1	6
Total Prediksi		52	14	1	67

#### 4.11.1 Metrik Kinerja Keseluruhan

Metrik kinerja keseluruhan yang utama adalah Akurasi (*Accuracy*), yang mengukur proporsi total prediksi yang benar dari seluruh data *testing*. Selain itu, metrik keseluruhan juga dapat dihitung menggunakan *Micro-Average Metrics*. Berikut adalah hasil perhitungan untuk metrik kinerja keseluruhan:

##### 1. Akurasi (*Accuracy*)

Perhitungan akurasi didasarkan pada jumlah total prediksi yang benar (diagonal utama *confusion matrix*) dibagi dengan total data *testing* menggunakan persamaan (2.19).

$$\begin{aligned}
 Accuracy &= \frac{TP}{\text{Total Data Testing}} \\
 &= \frac{34 + 8 + 1}{67} \\
 &= \frac{43}{67} \\
 &= 0,6418 \text{ atau } 64,18\%
 \end{aligned}$$

##### 2. *Micro-Average Metrics*

Metrik *Micro-Average* dapat dihitung dengan menjumlahkan semua *True Positive* (TP), *False Positive* (FP), dan *False Negative* (FN) dari semua kelas, kemudian menghitung *Precision*, *Recall*, dan *F1-Score* dari total tersebut.

a. Perhitungan Total TP, FP, dan FN (*Micro*):

$$1) \text{ Total TP} = \text{TP}_0 + \text{TP}_1 + \text{TP}_2 = 34 + 8 + 1 = 43$$

$$2) \text{ Total FP} = \text{FP}_0 + \text{FP}_1 + \text{FP}_2 = (15 + 3) + (4 + 2) + (0 + 0) = 24$$

$$3) \text{ Total FN} = \text{FN}_0 + \text{FN}_1 + \text{FN}_2 = (4 + 0) + (15 + 0) + (3 + 2) = 24$$

b. *Precision*:

$$\text{Precision} = \frac{\text{Total TP}}{\text{Total TP} + \text{Total FP}} = \frac{43}{43 + 24} = 0,6418 = 64,18\%$$

c. *Recall*:

$$\text{Recall} = \frac{\text{Total TP}}{\text{Total TP} + \text{Total FN}} = \frac{43}{43 + 24} = 0,6418 = 64,18\%$$

d. *Micro F1-Score*:

$$\text{Micro F1 - Score} = 2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} = 0,6418 = 64,18\%$$

Hasil perhitungan metrik kinerja keseluruhan model pada data *testing* menunjukkan nilai akurasi sebesar 64,18%. Nilai ini mengindikasikan bahwa model berhasil memprediksi kategori kelulusan dengan benar untuk sekitar 64 dari 100 mahasiswa yang diuji. Hasil akurasi ini dikonfirmasi oleh perhitungan *micro-average*, di mana nilai *Micro-F1 Score* identik dengan nilai Akurasi, yaitu 0,6418 atau 64,18%. Konsistensi nilai ini selaras dengan teori yang dijelaskan oleh Grandini dkk. (2020), yang menyatakan bahwa dalam klasifikasi multikelas, nilai *Micro-F1 Score* memang identik dengan nilai Akurasi.

#### 4.11.2 Metrik Kinerja Per-kategori

Analisis yang lebih mendalam dilakukan untuk mengevaluasi kinerja model pada setiap kategori kelulusan secara spesifik. Perhitungan ini juga didasarkan pada *confusion matrix* dari Tabel 4.12, Berikut perhitungan manual untuk setiap kategori kelulusan menggunakan persamaan (2.21) untuk *Precision*, persamaan (2.22) untuk *Recall*, dan persamaan (2.23) untuk *F1-Score*.

1. Kategori Tepat Waktu (0):

**Tabel 4.13** *Multiclass Confusion Matrix* untuk Kategori Tepat Waktu

		Prediksi		
		0	1	2
Aktual	0	TP	FN	FN
	1	FP	TN	TN
	2	FP	TN	TN

- a. Perhitungan TP, TN, FP, FN untuk kategori tepat waktu:

- 1)  $TP = 34$
- 2)  $TN = 8 + 0 + 2 + 1 = 11$
- 3)  $FP = 15 + 3 = 18$
- 4)  $FN = 4 + 0 = 4$

- b. *Precision*:

$$Precision_0 = \frac{TP_0}{TP_0 + FP_0} = \frac{34}{34 + 18} = \frac{34}{52} = 0,6538 = 65,38\%$$

- c. *Recall*:

$$Recall_0 = \frac{TP_0}{TP_0 + FN_0} = \frac{34}{34 + 4} = \frac{34}{38} = 0,8947 = 89,47\%$$

- d. *F1-Score*:

$$F1 - Score_0 = 2 \times \frac{Recall_0 \times Precision_0}{Recall_0 + Precision_0} = 0,7555 = 75,55\%$$

Hasil analisis kinerja per-kategori menunjukkan bahwa kategori Tepat Waktu mencatatkan performa terkuat. Nilai *Recall* kategori ini terhitung sangat tinggi, yaitu 89,47%, yang berarti model berhasil menemukan hampir 89% dari semua mahasiswa yang sebenarnya lulus Tepat Waktu. Selanjutnya, nilai *Precision* sebesar 65,38% menunjukkan bahwa prediksi Tepat Waktu dari model benar 65,38% dari waktu. *F1-Score* yang mencapai 75,55% adalah skor keseimbangan yang kuat, mengindikasikan performa yang baik pada kategori ini.

2. Kategori Terlambat (1):

**Tabel 4.14** *Multiclass Confusion Matrix* untuk Kategori Terlambat

		Prediksi		
		0	1	2
Aktual	0	TN	FP	TN
	1	FN	TP	FN
	2	TN	FP	TN

a. Perhitungan TP, TN, FP, FN untuk kategori terlambat:

- 1) TP = 8
- 2) TN = 34 + 0 + 3 + 1 = 38
- 3) FP = 4 + 2 = 6
- 4) FN = 15 + 0 = 15

b. *Precision*:

$$Precision_1 = \frac{TP_1}{TP_1 + FP_1} = \frac{8}{8 + 6} = \frac{8}{15} = 0,5333 = 53,33\%$$

c. *Recall*:

$$Recall_1 = \frac{TP_1}{TP_1 + FN_1} = \frac{8}{8 + 15} = \frac{8}{23} = 0,3478 = 34,78\%$$

d. *F1-Score*:

$$F1 - Score_1 = 2 \times \frac{Recall_1 \times Precision_1}{Recall_1 + Precision_1} = 0,4210 = 42,10\%$$

Hasil analisis kinerja model untuk kategori Terlambat menghasilkan nilai *Recall* 34,78%. Nilai *Recall* yang rendah ini berarti model gagal mengidentifikasi lebih dari 60% mahasiswa yang sebenarnya lulus Terlambat (15 mahasiswa salah diklasifikasikan, kebanyakan sebagai Tepat Waktu). Nilai *Precision* 53,33% menunjukkan bahwa ketika model memprediksi seorang mahasiswa Terlambat, prediksi tersebut benar 53,33% dari waktu. Dengan *F1-Score* 42,10%, kinerja gabungan untuk kategori ini secara keseluruhan adalah rendah.

3. Kategori Kritis (2):

**Tabel 4.15** *Multiclass Confusion Matrix* untuk Kategori Kritis

		Prediksi		
		0	1	2
Aktual	0	TN	TN	FP
	1	TN	TN	FP
	2	FN	FN	TP

a. Perhitungan TP, TN, FP, FN untuk kategori kritis:

- 1)  $TP = 1$
- 2)  $TN = 34 + 4 + 15 + 8 = 61$
- 3)  $FP = 0 + 0 = 0$
- 4)  $FN = 3 + 2 = 5$

b. *Precision*:

$$Precision_2 = \frac{TP_2}{TP_2 + FP_2} = \frac{1}{1 + 0} = 1 = 100\%$$

c. *Recall*:

$$Recall_2 = \frac{TP_2}{TP_2 + FN_2} = \frac{1}{1 + 5} = \frac{1}{6} = 0,167 = 16,7\%$$

d. *F1-Score*:

$$F1 - Score_2 = 2 \times \frac{Recall_2 \times Precision_2}{Recall_2 + Precision_2} = 0,286 = 28,6\%$$

Hasil analisis kinerja model untuk kategori Kritis menghasilkan *F1-Score* terendah sebesar 28,6%, menunjukkan performa paling lemah. Meskipun model mencatatkan *Precision* 100% , nilai sempurna ini bersifat semu karena model hanya memprediksi 1 mahasiswa sebagai Kritis (dan kebetulan benar). Nilai *Recall* yang sangat rendah sebesar 16,7% menunjukkan bahwa model ini gagal mengidentifikasi 5 dari 6 mahasiswa yang sebenarnya masuk kategori Kritis. Dengan demikian, kinerja model dikonfirmasi tidak dapat diandalkan untuk memprediksi kategori Kritis.

Analisis per-kategori ini menunjukkan kelemahan signifikan yang tersembunyi di balik nilai akurasi keseluruhan 64,18%. Model ini menunjukkan kinerja yang kuat dalam mengidentifikasi kategori mayoritas (Tepat Waktu), terutama dalam hal *recall*. Namun, ditemukan bahwa model sangat lemah dalam mengidentifikasi kategori minoritas, yaitu Terlambat dan Kritis. Kelemahan ini terlihat dari kecenderungan model untuk salah mengklasifikasikan mahasiswa yang Terlambat dan Kritis sebagai Tepat Waktu. Hasil analisis ini mengonfirmasi bahwa, sebagai alat prediksi, model memiliki bias kuat terhadap kategori Tepat Waktu dan hampir tidak berguna untuk mendeteksi secara dini mahasiswa yang berisiko lulus Kritis.

## 4.12 Interpretasi Model

Interpretasi model regresi logistik dilakukan dengan berdasarkan nilai *odds ratio*. Tujuan dari interpretasi model adalah untuk menerjemahkan koefisien matematis ( $\beta$ ) yang telah diperoleh dari hasil uji signifikansi parsial model revisi (disajikan pada tabel 4.9) menjadi implikasi praktis yang lebih mudah dipahami.

Analisis ini dilakukan dengan menghitung nilai *odds ratio* ( $\text{Exp}(\beta)$ ) dengan merujuk pada Persamaan (2.19). Karena model ini bersifat multinomial dengan 3 kategori, interpretasi akan dibagi berdasarkan  $k-1$  ( $3 - 1 = 2$ ) fungsi logit yang terbentuk, dengan kategori Tepat Waktu ( $Y=0$ ) sebagai kategori referensi. Hasil nilai *odds ratio* yang dihitung untuk masing-masing logit disajikan pada tabel-tabel berikut.

### 4.12.1 Interpretasi Model Logit 1 (Terlambat vs Tepat Waktu)

Perhitungan manual nilai *odds ratio* pada model Logit 1 berdasarkan persamaan (2.19) sebagai berikut:

Perhitungan nilai *odds ratio* untuk variabel  $x_1$ :

$$OR_j(a, b) = \Psi_{ab} = \frac{P(Y = j|x = a)/P(Y = 0|x = a)}{P(Y = j|x = b)/P(Y = 0|x = b)}$$

$$OR_j = e^{\beta_j}$$

$$OR_1 = e^{-1,186}$$

$$= 0,306$$

Perhitungan nilai *odds ratio* untuk variabel  $x_7$ :

$$OR_j(a, b) = \Psi_{ab} = \frac{P(Y = j|x = a)/P(Y = 0|x = a)}{P(Y = j|x = b)/P(Y = 0|x = b)}$$

$$OR_j = e^{\beta_j}$$

$$OR_1 = e^{0,063}$$



$$= 1,065$$

Diperoleh hasil sebagaimana disajikan pada Tabel 4.16.

**Tabel 4.16** Nilai *Odds Ratio* Logit 1

Variabel	Estimasi	<i>Odds Ratio</i>
<i>Intercept</i>	-0,575	0,563
$x_1$	-1,186	0,306
$x_7$	0,063	1,065

Tabel 4.16 menunjukkan hasil estimasi model dan hasil perhitungan nilai *odds ratio* untuk Logit 1 (Terlambat vs Tepat Waktu). Diperoleh bahwa variabel IPK memiliki koefisien logit sebesar  $-1,186$ . Nilai koefisien negatif ini menunjukkan bahwa peningkatan nilai IPK berhubungan dengan penurunan peluang mahasiswa mengalami keterlambatan kelulusan dibandingkan dengan mahasiswa yang lulus tepat waktu. Secara substantif, setiap satu satuan peningkatan IPK menurunkan nilai logit keterlambatan kelulusan, yang kemudian tercermin dalam nilai *odds ratio* sebesar 0,306. Nilai *odds ratio* yang kurang dari 1 ini menunjukkan bahwa kenaikan IPK akan menurunkan peluang mahasiswa untuk lulus Terlambat sebesar  $(1 - 0,306)$  atau 69,4%, sehingga semakin tinggi IPK mahasiswa maka semakin kecil kemungkinan mereka lulus Terlambat (dibandingkan lulus Tepat Waktu). Dengan demikian, interpretasi koefisien logit dan *odds ratio* menghasilkan kesimpulan yang konsisten bahwa IPK merupakan prediktor yang menurunkan risiko keterlambatan kelulusan.

Selain itu, Tabel 4.16 juga menunjukkan bahwa variabel Rasio SKS Lulus memiliki koefisien logit sebesar 0,063 dengan nilai *odds ratio* 1,065. Koefisien logit yang bernilai positif mengindikasikan bahwa peningkatan satu satuan pada rasio SKS Lulus berhubungan dengan peningkatan peluang mahasiswa berada

dalam kategori Terlambat (dibandingkan Tepat Waktu). Nilai *odds ratio* 1,065 menunjukkan bahwa setiap kenaikan satu satuan rasio SKS lulus meningkatkan peluang mahasiswa lulus Terlambat sebesar  $(1,065 - 1)$  atau 6,5% dari *odds* sebelumnya. Dengan demikian, interpretasi koefisien logit dan *odds ratio* untuk variabel Rasio SKS Lulus memberikan kesimpulan bahwa semakin tinggi rasio SKS yang diselesaikan mahasiswa, semakin tinggi kemungkinan mereka masuk dalam kategori Terlambat dibandingkan Tepat Waktu.

#### 4.12.2 Interpretasi Model Logit 2 (Kritis vs Tepat Waktu)

Perhitungan manual nilai *odds ratio* pada model logit 2 berdasarkan persamaan (2.19) sebagai berikut:

Perhitungan nilai *odds ratio* untuk variabel  $x_1$ :

$$OR_j(a, b) = \Psi_{ab} = \frac{P(Y = j|x = a)/P(Y = 0|x = a)}{P(Y = j|x = b)/P(Y = 0|x = b)}$$

$$OR_j = e^{\beta_j}$$

$$OR_2 = e^{-0,945}$$

$$= 0,389$$

Perhitungan nilai *odds ratio* untuk variabel  $x_7$ :

$$OR_j(a, b) = \Psi_{ab} = \frac{P(Y = j|x = a)/P(Y = 0|x = a)}{P(Y = j|x = b)/P(Y = 0|x = b)}$$

$$OR_j = e^{\beta_j}$$

$$OR_2 = e^{-0,780}$$

$$= 0,459$$

Diperoleh hasil sebagaimana disajikan pada Tabel 4.17.

**Tabel 4.17** Nilai *Odds Ratio* Logit 2

<b>Variabel</b>	<b>Estimasi</b>	<b><i>Odds Ratio</i></b>
<i>Intercept</i>	-2,172	0,114
$x_1$	-0,945	0,389
$x_7$	-0,780	0,459

Tabel 4.17 menunjukkan hasil estimasi model dan hasil perhitungan nilai *odds ratio* untuk Logit 2 (Kritis vs Tepat Waktu). Diperoleh bahwa variabel IPK memiliki koefisien logit sebesar  $-0,945$ . Nilai koefisien negatif ini menunjukkan bahwa peningkatan nilai IPK berhubungan dengan penurunan peluang mahasiswa masuk dalam kategori Kritis dibandingkan dengan mahasiswa yang lulus Tepat Waktu. Secara substantif, setiap satu satuan peningkatan IPK menurunkan nilai logit mahasiswa masuk kategori Kritis, yang kemudian tercermin dalam nilai *odds ratio* sebesar 0,389. Nilai *odds ratio* yang kurang dari 1 ini menunjukkan bahwa kenaikan satu satuan IPK menurunkan peluang mahasiswa untuk lulus Kritis menjadi hanya 0,389 kali dari odds sebelumnya, atau setara dengan penurunan peluang sebesar  $(1 - 0,389)$  atau 61,1%. Dengan demikian, interpretasi koefisien logit dan *odds ratio* dari variabel IPK secara konsisten menunjukkan bahwa semakin tinggi IPK mahasiswa, semakin kecil kemungkinan mereka masuk dalam kategori Kritis.

Selain itu, Tabel 4.17 juga menunjukkan bahwa variabel Rasio SKS Lulus memiliki koefisien logit sebesar  $-0,780$  dengan nilai *odds ratio* 0,459. Koefisien logit yang bernilai negatif mengindikasikan bahwa peningkatan satu satuan pada rasio SKS Lulus berhubungan dengan penurunan peluang mahasiswa berada dalam kategori Kritis. Nilai *odds ratio* 0,459 menunjukkan bahwa setiap kenaikan satu satuan rasio SKS Lulus menurunkan peluang mahasiswa lulus Kritis menjadi hanya 0,459 kali dari *odds* sebelumnya, atau berkurang sebesar  $(1 - 0,459)$  atau

54,1%. Dengan demikian, interpretasi koefisien logit dan *odds ratio* untuk variabel Rasio SKS Lulus memberikan kesimpulan bahwa semakin tinggi rasio SKS yang diselesaikan mahasiswa, semakin kecil kemungkinan mereka masuk dalam kategori Kritis.

#### 4.13 Simulasi Prediksi pada Studi Kasus

Tahap terakhir adalah mengilustrasikan penerapan praktis dari model. Interpretasi dari simulasi ini adalah untuk menunjukkan bahwa model dalam tindakan. Setelah memvalidasi signifikansi dan kesesuaian model, langkah ini menunjukkan bagaimana model tersebut dapat digunakan sebagai alat prediksi untuk satu individu, mengubah koefisien abstrak menjadi probabilitas yang dapat ditindaklanjuti untuk pengambilan keputusan.

Digunakan data *testing* untuk studi kasus mahasiswa dengan profil sebagai berikut:

Tahun Masuk: 2018

IPK: 3,70

Jenis Kelamin: Perempuan

Jenis Asal Sekolah: MA

Jenis Pembiayaan: Non Bidikmisi

Jalur Masuk: SNMPTN

Asal Daerah: Luar Kota Malang

SKS (semester 1-6 tanpa mengulang): 128

SKS Total: 150

Rasio SKS Lulus:  $\frac{128}{150} = 0,85$

Ketepatan Waktu Kelulusan: Tepat Waktu

Berdasarkan profil studi kasus, nilai vektor  $X$  yang memuat variabel *dummy* serta variabel prediktor numerik yang telah distandardisasi ditentukan sebagai berikut:

$$X_1(\text{IPK}_{\text{std}}) = 1,04$$

$$X_2(\text{Laki} - \text{laki}) = 0$$

$$X_{3\text{MA}} = 1$$

$$X_{3\text{SMK}} = 0$$

$$X_{4\text{Bidikmisi}} = 0$$

$$X_{5\text{SNMPTN}} = 1$$

$$X_{5\text{Mandiri}} = 0$$

$$X_{6\text{Kota Malang}} = 0$$

$$X_7(\text{Rasio SKS}_{\text{std}}) = -1,41$$

Selanjutnya, nilai-nilai variabel prediktor tersebut disubstitusikan ke dalam persamaan model akhir untuk menghitung manual nilai logit  $g_1(x)$  dan  $g_2(x)$  berdasarkan persamaan (2.4) dan (2.5).

$$\begin{aligned} g_1(x) &= \ln \left[ \frac{P(Y = 1)}{P(Y = 0)} \right] \\ &= -0,575 - 1,186 \left( x_{1\text{IPK}_{\text{std}}} \right) \\ &= -0,575 - 1,186(1,04) \\ &= -1,808 \end{aligned}$$

$$g_2(x) = \ln \left[ \frac{P(Y = 2)}{P(Y = 0)} \right]$$

$$\begin{aligned}
&= -2,172 - 0,945 \left( x_{1_{IPK_{std}}} \right) - 0,780 \left( x_{7_{Rasio SKSLulus_{std}}} \right) \\
&= -2,172 - 0,945(1,04) - 0,780(-1,41) \\
&= -2,055
\end{aligned}$$

Dengan nilai logit  $g_1(x) = -1,808$  dan  $g_2(x) = -2,055$ , probabilitas untuk setiap kategori dihitung berdasarkan persamaan (2.6) sampai (2.8). Pertama, akan dihitung komponen eksponensial:

$$\exp(g_1(x)) = e^{-1,808} \approx 0,164$$

$$\exp(g_2(x)) = e^{-2,730} \approx 0,128$$

Kedua, akan dihitung denominator (penyebut) umum:

$$\text{Denominator} = 1 + \exp(g_1(x)) + \exp(g_2(x))$$

$$\text{Denominator} = 1 + 0,164 + 0,128 = 1,292$$

Ketiga, hitung probabilitas untuk setiap kategori:

$P(Y = 0 | \text{Tepat Waktu})$ :

$$P(Y = 0|x) = \frac{1}{1 + \exp g_1(x) + \exp g_2(x)} = \frac{1}{1,292} \approx 0,7740$$

$P(Y = 1 | \text{Terlambat})$ :

$$P(Y = 1|x) = \frac{\exp g_1(x)}{1 + \exp g_1(x) + \exp g_2(x)} = \frac{0,164}{1,292} \approx 0,1269$$

$P(Y = 2 | \text{Kritis})$ :

$$P(Y = 2|x) = \frac{\exp g_2(x)}{1 + \exp g_1(x) + \exp g_2(x)} = \frac{0,128}{1,292} \approx 0,0991$$

Berdasarkan hasil perhitungan manual. Probabilitas untuk mahasiswa dengan profil yang tertera adalah:

Peluang Lulus Tepat Waktu ( $Y=0$ ) : 77,40%

Peluang Lulus Terlambat ( $Y=1$ ) : 12,69%

Peluang Lulus Kritis (Y=2) : 9,91%

Model memprediksi bahwa kategori kelulusan yang paling mungkin (probabilitas tertinggi) untuk mahasiswa dengan profil tersebut adalah Tepat Waktu (Y=0).

#### 4.14 Kajian Penelitian dalam Perspektif Islam

Berdasarkan hasil penelitian prediksi ketepatan waktu kelulusan, diketahui bahwa variabel akademik seperti Indeks Prestasi Kumulatif (IPK) dan Rasio SKS Lulus memberikan pengaruh yang signifikan terhadap peluang mahasiswa lulus tepat waktu. Mahasiswa dengan IPK tinggi dan rasio SKS yang baik menunjukkan kesungguhan dalam menuntut ilmu, yang berdampak pada penyelesaian studi yang lebih cepat. Hal ini menunjukkan bahwa prestasi akademik bukan sekadar angka, melainkan cerminan dari derajat keilmuan seseorang. Hal tersebut selaras dengan ajaran Islam yang memuliakan para penuntut ilmu, sebagaimana tercantum dalam QS. Al-Mujadilah ayat 11 sebagai berikut:

يَا أَيُّهَا الَّذِينَ آمَنُوا إِذَا قِيلَ لَكُمْ تَفَسَّحُوا فِي الْمَجَالِسِ فَافْسَحُوا يَفْسَحِ اللَّهُ لَكُمْ وَإِذَا قِيلَ انشُرُوا فَانشُرُوا فَيَرْفَعِ

اللَّهُ الَّذِينَ آمَنُوا مِنْكُمْ وَالَّذِينَ أُوتُوا الْعِلْمَ دَرَجَاتٍ وَاللَّهُ بِمَا تَعْمَلُونَ خَبِيرٌ ﴿١١﴾

*"Hai orang-orang beriman apabila dikatakan kepadamu: 'Berlapang-lapanglah dalam majelis', maka lapangkanlah niscaya Allah akan memberi kelapangan untukmu. Dan apabila dikatakan: 'Berdirilah kamu', maka berdirilah, niscaya Allah akan meninggikan orang-orang yang beriman di antaramu dan orang-orang yang diberi ilmu pengetahuan beberapa derajat. Dan Allah Maha Mengetahui apa yang kamu kerjakan." (QS. Al-Mujadilah: 11)*

Berdasarkan ayat tersebut, dapat dipahami bahwa kesuksesan akademik (kelulusan tepat waktu) merupakan salah satu bentuk kelapangan dan peninggian

derajat yang Allah SWT berikan kepada mahasiswa yang serius dalam menuntut ilmu. IPK yang tinggi bukan hanya sekadar angka, melainkan manifestasi dari ikhtiar dan kesungguhan dalam memahami ilmu yang dipelajari. Dengan demikian, pencapaian akademik yang baik mencerminkan keseimbangan antara upaya intelektual dan spiritual, di mana ilmu yang diperoleh diharapkan membawa manfaat tidak hanya bagi diri sendiri, tetapi juga bagi masyarakat luas sesuai dengan prinsip kemaslahatan dalam Islam.

Selain faktor prestasi akademik, penelitian ini juga menyoroti aspek Ketepatan Waktu sebagai variabel utama. Dalam Islam, waktu adalah modal utama manusia yang tidak boleh disia-siakan. Mahasiswa yang lulus tepat waktu pada hakikatnya adalah mereka yang mampu manajemen waktu studinya dengan baik dan tidak membiarkan dirinya berada dalam kerugian akibat penundaan. Pentingnya disiplin waktu ini ditegaskan oleh Allah SWT dalam QS. Al-Asr ayat 1-3:

وَالْعَصْرِ ﴿١﴾ إِنَّ الْإِنْسَانَ لَفِي خُسْرٍ ﴿٢﴾ إِلَّا الَّذِينَ آمَنُوا وَعَمِلُوا الصَّالِحَاتِ وَتَوَاصَوْا بِالْحَقِّ ۖ وَتَوَاصَوْا  
بِالصَّبْرِ ﴿٣﴾

*"Demi masa. Sesungguhnya manusia itu benar-benar dalam kerugian, kecuali orang-orang yang beriman dan mengerjakan amal saleh dan nasehat menasehati supaya mentaati kebenaran dan nasehat menasehati supaya menetapi kesabaran."* (QS. Al-Asr: 1-3)

Ayat ini menjadi landasan teologis bahwa keterlambatan kelulusan tanpa alasan yang syar'i (uzur) merupakan bentuk kerugian waktu. Mahasiswa yang lulus tepat waktu telah mengamalkan prinsip menghargai waktu (demi masa) dan melakukan amal saleh berupa penyelesaian amanah studi tepat pada waktunya. Ayat ini juga mengajarkan pentingnya disiplin, kesungguhan, dan tanggung jawab dalam



mengelola waktu yang telah Allah anugerahkan. Setiap detik yang berlalu hendaknya dimanfaatkan untuk menuntut ilmu, beramal, dan memberikan kontribusi nyata bagi kemajuan diri serta lingkungan. Dengan demikian, keberhasilan akademik tidak hanya dinilai dari capaian angka IPK, tetapi juga dari sejauh mana seseorang mampu mengimplementasikan nilai-nilai keimanan, amal saleh, dan kesabaran dalam proses pendidikannya.

Hasil penelitian menunjukkan bahwa konsistensi dalam menyelesaikan beban studi (Rasio SKS Lulus) berpengaruh positif terhadap kelulusan. Hal ini mencerminkan sikap profesionalisme dan ketuntasan dalam bekerja atau belajar (*itqan*). Islam sangat mencintai umatnya yang ketika mengerjakan sesuatu, ia menyelesaikannya dengan tuntas dan berkualitas, tidak menunda-nunda atau membiarkannya terbengkalai. Hal ini sesuai dengan sabda Rasulullah SAW dalam sebuah hadits:

عَنْ عَائِشَةَ رَضِيَ اللَّهُ عَنْهَا قَالَتْ: قَالَ رَسُولُ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ: إِنَّ اللَّهَ تَعَالَى يُحِبُّ إِذَا عَمِلَ أَحَدُكُمْ عَمَلًا أَنْ يُتَمِّتَهُ (رواه الطبري والبيهقي)

*Dari Aisyah r.a., sesungguhnya Rasulullah s.a.w. bersabda: “Sesungguhnya Allah mencintai seseorang yang apabila bekerja, mengerjakannya secara profesional”. (HR. Thabrani, No: 891, Baihaqi, No: 334).*

Hadits ini relevan dengan temuan bahwa mahasiswa yang memiliki rasio kelulusan SKS tinggi (tidak banyak mengulang dan tuntas tiap semester) cenderung lulus tepat waktu. Perilaku akademik yang teratur dan tuntas ini adalah bentuk *itqan* dalam menuntut ilmu. Dengan demikian, hasil penelitian ini tidak hanya membuktikan hubungan statistik antar variabel, tetapi juga mengonfirmasi bahwa nilai-nilai kedisiplinan, kesungguhan menuntut ilmu, dan profesionalisme yang

diajarkan Islam memiliki korelasi positif dengan keberhasilan akademik mahasiswa.

Selain itu, hadits ini menegaskan bahwa setiap pekerjaan, termasuk kegiatan akademik, hendaknya dilakukan dengan kesempurnaan niat, perencanaan, dan pelaksanaan. Mahasiswa yang belajar dengan *itqan* tidak hanya mengejar hasil berupa nilai atau kelulusan, tetapi juga berorientasi pada keberkahan ilmu yang diperoleh. Nilai spiritual dalam *itqan* menumbuhkan kesadaran bahwa kesuksesan akademik adalah wujud ibadah kepada Allah SWT melalui dedikasi dan kejujuran intelektual. Dengan demikian, penerapan konsep *itqan* dalam konteks pendidikan modern mencerminkan integrasi antara ilmu dan iman, di mana kesungguhan belajar menjadi jalan untuk mencapai ridha Allah SWT sekaligus keberhasilan akademik yang bermakna.

Integrasi antara ilmu pendidikan dan nilai-nilai Islam tentang ketepatan waktu kelulusan menciptakan pendekatan yang tidak hanya berfokus pada aspek akademik, tetapi juga menghubungkan antara tanggung jawab intelektual dan spiritual. Praktiknya dalam kehidupan mahasiswa tercermin dari kesadaran akan amanah menuntut ilmu yang harus diselesaikan dengan penuh kedisiplinan dan tanggung jawab terhadap waktu. Ketepatan waktu dalam kelulusan menjadi wujud penghargaan terhadap nikmat waktu yang diberikan Allah SWT. Mahasiswa yang mampu mengelola waktu belajar dengan baik akan lebih optimal dalam menunaikan kewajiban akademik, berkontribusi bagi masyarakat, serta meraih keberkahan dalam setiap proses pendidikannya. Islam memandang ketepatan dan ketuntasan dalam studi sebagai bagian integral dari keberhasilan hidup dunia dan

akhirat yang harus dijaga dengan kesungguhan, keikhlasan, dan kesadaran penuh akan nilai ibadah dalam setiap langkah menuntut ilmu.

## BAB V

### PENUTUP

#### 5.1 Kesimpulan

Berdasarkan hasil analisis yang telah dilakukan, dapat disimpulkan sebagai berikut:

1. Model akhir regresi logistik multinomial disusun berdasarkan variabel-variabel signifikan, dengan kategori Tepat Waktu ( $Y = 0$ ) sebagai referensi. Dua fungsi logit yang dihasilkan disajikan pada persamaan (4.1) dan (4.2) sebagai berikut:

Logit 1 (Terlambat vs Tepat Waktu):

$$\begin{aligned}g_1(x) &= \ln \left[ \frac{P(Y = 1)}{P(Y = 0)} \right] \\ &= -0,575 - 1,186 \left( x_{1IPK_{std}} \right)\end{aligned}$$

Pada Logit 1, variabel IPK berpengaruh signifikan dengan koefisien  $-1,186$ , yang berarti semakin tinggi IPK mahasiswa, semakin kecil peluang mereka berada pada kategori Terlambat dibandingkan kategori Tepat Waktu. Sedangkan variabel Rasio SKS Lulus tidak signifikan, sehingga tidak memengaruhi perbedaan peluang antara kedua kategori tersebut.

Logit 2 (Kritis vs Tepat Waktu):

$$\begin{aligned}g_2(x) &= \ln \left[ \frac{P(Y = 2)}{P(Y = 0)} \right] \\ &= -2,172 - 0,945 \left( x_{1IPK_{std}} \right) - 0,780 \left( x_{7Rasio\ SKSLulus_{std}} \right)\end{aligned}$$

Pada Logit 2, IPK dan Rasio SKS Lulus berpengaruh signifikan, dengan koefisien masing-masing  $-0,945$  dan  $-0,780$ . Kedua nilai negatif tersebut

menunjukkan bahwa peningkatan IPK maupun Rasio SKS Lulus menurunkan peluang mahasiswa masuk kategori Kritis dibandingkan kategori Tepat Waktu.

2. Kinerja model regresi logistik multinomial secara keseluruhan pada data *testing* menghasilkan nilai akurasi sebesar 64,18%. Nilai ini identik dengan nilai *Micro F1-Score* (64,18%). Namun, analisis kinerja per-kategori menunjukkan hasil yang tidak seimbang sebagai berikut:

- a. Kategori Tepat Waktu ( $Y=0$ ) mencatatkan *F1-Score* tertinggi sebesar 75,55%.
- b. Kategori Terlambat ( $Y=1$ ) mencatatkan *F1-Score* yang rendah sebesar 42,10%.
- c. Kategori Kritis ( $Y=2$ ) mencatatkan *F1-Score* terendah sebesar 28,6%.

Kinerja yang timpang ini menunjukkan bahwa model memiliki bias kuat terhadap kategori mayoritas (Tepat Waktu) dan kesulitan mengidentifikasi kategori minoritas (Terlambat dan Kritis).

3. Hasil prediksi probabilitas studi kasus untuk studi kasus seorang mahasiswa dengan profil akademik yang baik (IPK: 3,70; Perempuan; MA; SNMPTN; Non Bidikmisi; Luar Kota Malang; Rasio SKS: 0,85), model memprediksi probabilitas sebagai berikut:

- a. Peluang Lulus Tepat Waktu ( $Y=0$ ) : 77,40%
- b. Peluang Lulus Terlambat ( $Y=1$ ) : 12,69%
- c. Peluang Lulus Kritis ( $Y=2$ ) : 9,91%

Prediksi akhir untuk mahasiswa dengan profil tersebut adalah Tepat Waktu, karena memiliki probabilitas tertinggi.

## 5.2 Saran

Berdasarkan kesimpulan penelitian, terdapat beberapa saran penulis untuk penelitian selanjutnya dan instansi terkait dalam upaya pengembangan dan penyempurnaan penelitian sebagai berikut:

1. Meskipun kinerja model untuk kategori minoritas masih rendah, hasil penelitian ini mengonfirmasi secara statistik bahwa IPK dan Rasio SKS Lulus adalah dua faktor prediktor paling signifikan, sehingga program studi disarankan untuk menggunakan kedua indikator ini sebagai sistem peringatan dini (*early warning system*). Melalui pemantauan ketat terhadap mahasiswa yang nilai IPK atau Rasio SKS Lulusnya mulai menurun di semester-semester awal, intervensi akademik yang proaktif, seperti bimbingan intensif atau perencanaan studi yang lebih terstruktur, dapat diberikan kepada mahasiswa yang teridentifikasi berisiko tinggi sebelum mereka terlanjur masuk ke kategori Terlambat atau Kritis.
2. Kinerja model yang rendah pada kategori Terlambat dan Kritis sangat mungkin disebabkan oleh ketidakseimbangan data (*imbalanced data*). Peneliti selanjutnya disarankan untuk menerapkan teknik *resampling* (seperti SMOTE untuk *oversampling* atau *Tomek Links* untuk *undersampling*) untuk menyeimbangkan data *training* dan meningkatkan kemampuan model mengenali kelas minoritas.
3. Peneliti selanjutnya disarankan untuk mengeksplorasi variabel prediktor lain di luar data akademik (misalnya, variabel non-akademik seperti keaktifan organisasi, status bekerja, atau faktor psikologis) untuk membangun model yang lebih kaya dan berpotensi lebih akurat.

## DAFTAR PUSTAKA

- Agresti, A. (2002). *Categorical Data Analysis Second edition*. New York: John Wiley & Sons.
- BAN-PT. (2019). *Naskah Akademik IAPT 3.0*. Jakarta: Badan Akreditasi Nasional Perguruan Tinggi.
- Djudin, Tomo. (2018). Analisis Prestasi Akademik Mahasiswa Lulusan Jurusan Pendidikan Mipa Fkip Untan Ditinjau dari Jalur Masuk (SNMPTN, SBMPTN, Mandiri) dan Program Kuliah (S-1 Reguler, S-1 Percepatan APK). *Jurnal Pendidikan Matematika dan IPA*, 9 (2), 76-89, <http://dx.doi.org/10.26418/jpmipa.v9i2.25867>
- Fadilah, L. (2023). *Optimasi K-Nearest Neighbors Menggunakan Fuzzy C-Means Pada Ketepatan Waktu Kelulusan Mahasiswa*. Universitas Islam Negeri Maulana Malik Ibrahim.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). London, UK: SAGE Publications Ltd.
- Grandini, M., Bagli, E., & Visani, G. (2020). *Metrics for Multi-Class Classification: an Overview*. 1–17. <http://arxiv.org/abs/2008.05756>
- Gujarati, D. N. (2004). *Basic Econometrics, Fourth Edition*. McGraw-Hill Companies.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning.
- Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression*. Canada: John Wiley & Sons, Inc.
- Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). Hoboken, NJ: John Wiley & Sons.
- Imaslihkah, S., Ratna, M., & Ratnasari, V. (2013). Analisis regresi logistik ordinal terhadap faktor-faktor yang memengaruhi predikat kelulusan mahasiswa S1 di ITS Surabaya. *Jurnal Sains dan Seni ITS*, 2(2), 177–182.
- Kementerian Riset, Teknologi, dan Pendidikan Tinggi, Direktorat Jenderal Pembelajaran dan Kemahasiswaan. (2013). *Standar Nasional Pendidikan Tinggi (SN Dikti)*. Kementerian Riset, Teknologi, dan Pendidikan Tinggi.
- Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied linear statistical models* (5th ed.). McGraw-Hill Irwin.

- Malay, I., Ramadhani, T. R., Azkia, W., Alamsyah, S. K., & Koansa, M. R. (2025). Peran Analisis Probalitas dalam Memperkirakan Kelulusan Tepat Waktu di Perguruan Tinggi. *Jurnal Pendidikan Tambusai*, 9(2), 25273-25278.
- Mashfia, F. R. (2022). *Prediksi Ketepatan Waktu Kelulusan Mahasiswa Menggunakan Metode Naïve Bayes Classifier*. Universitas Islam Negeri Maulana Malik Ibrahim.
- Muhendra, R. (2021). Development of Graduation Prediction Model for Industrial Engineering Students Using Decision Tree. *Jurnal Manajemen dan Supervisi Pendidikan*, 6(1), 28-37.
- Qur'an Kemenag. (2024). Diakses pada 5 Maret 2025 dari <https://quran.kemenag.go.id/>
- Ristekdikti.go.id. (2019). Pertanyaan Umum Seputar Bidikmisi. Diakses pada 18 Mei 2025, dari <https://bidikmisi.belmawa.ristekdikti.go.id/petunjuk/20>
- Santrock, J. W. (2014). *Psikologi Pendidikan*. Edisi 5 Jilid 1. (Harya Bhimasena Translator). Jakarta: Salemba Humanika.
- Shihab, M. Q. (2002). *Tafsir Al-Mishbah: Pesan, kesan dan keserasian al-Qur'an (Jilid 1)*. Lentera Hati.
- Suprpto, S., Zubaidah, S., & Corebima, A. D. (2018). Pengaruh gender terhadap keterampilan berpikir kreatif siswa pada pembelajaran biologi. *Jurnal Pendidikan: Teori, Penelitian, dan Pengembangan*, 3(3), 325-329.
- Syahrani, R. (2022). *Prediksi Kategori Kelulusan Mahasiswa Menggunakan Metode Regresi Logistik Multinomial*. Universitas Islam Negeri Maulana Malik Ibrahim.
- Tim UIN Maulana Malik Ibrahim Malang. (2019). *Pedoman Pendidikan*. Malang: UIN Malang Press.
- Undang-Undang Republik Indonesia Nomor 12. (2012). Tentang Pendidikan Tinggi.
- Varamita, A. (2017). *Analisis Regresi dan Aplikasinya pada Penyakit Anemia untuk Ibu Hamil di RSKD Ibu dan Anak Siti Fatimah Makassar*. Universitas Negeri Makassar.
- Wahyuni, T., Agoestanto, A., & Pujiastuti, E. (2018). Analisis Regresi Logistik terhadap Keputusan Penerimaan Beasiswa PPA di FMIPA Unnes Menggunakan Software Minitab. *PRISMA, Prosiding Seminar Nasional Matematika*, 1, 755-764.



Widarjono, A. (2010). *Analisis Statistika Multivariat Terapan*. Yogyakarta: UPP STIM YKPN.

## LAMPIRAN

### Lampiran 1. Data dan Coding



#### 1. Data Kelulusan Mahasiswa Program Studi Matematika UIN Maulana Malik Ibrahim Malang Angkatan 2018-2021

No.	Tahun Masuk	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$Y$
1	2018	3,30	Laki-laki	SMA	Bidikmisi	SNMPTN	Luar Kota Malang	0,83	Terlambat
2	2018	3,67	Perempuan	MA	Bidikmisi	SNMPTN	Luar Kota Malang	0,89	Terlambat
3	2018	3,43	Perempuan	SMA	Non Bidikmisi	SNMPTN	Luar Kota Malang	0,87	Kritis
4	2018	3,27	Perempuan	SMA	Non Bidikmisi	SNMPTN	Luar Kota Malang	0,87	Terlambat
5	2018	3,65	Laki-laki	SMA	Non Bidikmisi	SNMPTN	Luar Kota Malang	0,88	Tepat Waktu
6	2018	3,70	Perempuan	MA	Non Bidikmisi	SNMPTN	Luar Kota Malang	0,85	Tepat Waktu
7	2018	3,29	Perempuan	MA	Non Bidikmisi	SNMPTN	Luar Kota Malang	0,89	Terlambat
8	2018	3,29	Perempuan	MA	Non Bidikmisi	SNMPTN	Luar Kota Malang	0,87	Terlambat
9	2018	3,08	Perempuan	SMA	Non Bidikmisi	SNMPTN	Luar Kota Malang	0,87	Terlambat
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
323	2021	3,42	Laki-laki	SMA	Non Bidikmisi	MANDIRI	Luar Kota Malang	0,91	Tepat Waktu
324	2021	3,59	Laki-laki	MA	Non Bidikmisi	MANDIRI	Luar Kota Malang	0,91	Tepat Waktu
325	2021	3,64	Perempuan	SMA	Non Bidikmisi	MANDIRI	Luar Kota Malang	0,90	Tepat Waktu
326	2021	3,69	Perempuan	MA	Non Bidikmisi	MANDIRI	Luar Kota Malang	0,90	Tepat Waktu
327	2021	3,65	Perempuan	SMA	Non Bidikmisi	MANDIRI	Luar Kota Malang	0,90	Tepat Waktu
328	2021	3,51	Laki-laki	SMA	Non Bidikmisi	MANDIRI	Luar Kota Malang	0,91	Tepat Waktu
329	2021	3,59	Perempuan	SMA	Non Bidikmisi	MANDIRI	Kota Malang	0,91	Tepat Waktu
330	2021	3,49	Perempuan	SMA	Non Bidikmisi	MANDIRI	Luar Kota Malang	0,90	Tepat Waktu
331	2021	3,53	Perempuan	SMA	Non Bidikmisi	MANDIRI	Luar Kota Malang	0,90	Tepat Waktu

#### Keterangan:

$x_1$  : IPK

$x_2$  : Jenis Kelamin

$x_3$  : Jenis Asal Sekolah

$x_4$  : Jenis Pembiayaan

$x_5$  : Jalur Masuk

$x_6$  : Asal Daerah

$x_7$  : Rasio SKS Lulus

$Y$  : Ketepatan Waktu Kelulusan

2. Data Setelah dilakukan *Dummy Variable*

No	Tahun Masuk	$x_1$	$x_2(1)$	$x_3(1)$	$x_3(3)$	$x_4(1)$	$x_5(1)$	$x_5(3)$	$x_6(1)$	$x_7$	Y
1	2018	3,30	1	0	0	1	1	0	0	0,83	1
2	2018	3,67	0	1	0	1	1	0	0	0,89	1
3	2018	3,43	0	0	0	0	1	0	0	0,87	2
4	2018	3,27	0	0	0	0	1	0	0	0,87	1
5	2018	3,65	1	0	0	0	1	0	0	0,88	0
6	2018	3,70	0	1	0	0	1	0	0	0,85	0
7	2018	3,29	0	1	0	0	1	0	0	0,89	1
8	2018	3,29	0	1	0	0	1	0	0	0,87	1
9	2018	3,08	0	0	0	0	1	0	0	0,87	1
10	2018	3,32	0	0	0	0	1	0	0	0,88	1
11	2018	3,57	0	0	0	0	1	0	0	0,89	0
12	2018	3,50	0	0	0	0	1	0	0	0,89	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
320	2021	3,23	0	0	0	0	0	0	0	0,91	0
321	2021	3,53	0	1	0	0	0	0	0	0,91	0
322	2021	3,25	1	0	0	0	0	1	0	0,88	0
323	2021	3,42	1	0	0	0	0	1	0	0,91	0
324	2021	3,59	1	1	0	0	0	1	0	0,91	0
325	2021	3,64	0	0	0	0	0	1	0	0,90	0
326	2021	3,69	0	1	0	0	0	1	0	0,90	0
327	2021	3,65	0	0	0	0	0	1	0	0,90	0
328	2021	3,51	1	0	0	0	0	1	0	0,91	0
329	2021	3,59	0	0	0	0	0	1	1	0,91	0
330	2021	3,49	0	0	0	0	0	1	0	0,90	0
331	2021	3,53	0	0	0	0	0	1	0	0,90	0

Keterangan:

$x_1$  : IPK

$x_2(1)$  : Jenis Kelamin Kategori Laki-laki

$x_3(1)$  : Jenis Asal Sekolah Kategori MA

$x_3(3)$  : Jenis Asal Sekolah Kategori SMK

$x_4(1)$  : Jenis Pembiayaan Kategori Bidikmisi

$x_5(1)$  : Jalur Masuk Kategori SNMPTN

$x_5(3)$  : Jalur Masuk Kategori Mandiri

$x_6(1)$  : Asal Daerah Kategori Kota Malang

$x_7$  : Rasio SKS Lulus

Y : Ketepatan Waktu Kelulusan

## 3. Data Setelah dilakukan Standardisasi

No	Tahun Masuk	$x_1$	$x_2(1)$	$x_3(1)$	$x_3(3)$	$x_4(1)$	$x_5(1)$	$x_5(1)$	$x_6(1)$	$x_7$	Y
1	2018	-1,07	1	0	0	1	1	0	0	-2,05	1
2	2018	0,89	0	1	0	1	1	0	0	-0,14	1
3	2018	-0,38	0	0	0	0	1	0	0	-0,77	2
4	2018	-1,23	0	0	0	0	1	0	0	-0,77	1
5	2018	0,78	1	0	0	0	1	0	0	-0,46	0
6	2018	1,04	0	1	0	0	1	0	0	-1,41	0
7	2018	-1,12	0	1	0	0	1	0	0	-0,14	1
8	2018	-1,12	0	1	0	0	1	0	0	-0,77	1
9	2018	-2,23	0	0	0	0	1	0	0	-0,77	1
10	2018	-0,96	0	0	0	0	1	0	0	-0,46	1
11	2018	0,36	0	0	0	0	1	0	0	-0,14	0
12	2018	-0,01	0	0	0	0	1	0	0	-0,14	0
13	2018	-0,49	0	1	0	0	1	0	0	-0,46	0
14	2018	0,46	0	0	0	0	1	0	0	-0,14	0
15	2018	-1,7	1	1	0	0	1	0	0	-4,59	2
16	2018	-2,7	0	0	0	0	1	0	0	-3	1
17	2018	-1,07	0	1	0	0	1	0	0	-0,77	2
18	2018	0,73	0	0	0	0	1	0	0	-0,77	0
19	2018	-0,22	0	0	0	0	1	0	0	-0,77	0
20	2018	-0,28	1	1	0	0	1	0	0	-0,14	1
∴	∴	∴	∴	∴	∴	∴	∴	∴	∴	∴	∴
312	2021	2,15	0	1	0	0	0	0	0	-0,14	0
313	2021	2,1	1	0	0	0	0	0	0	0,5	0
314	2021	0,31	0	0	0	0	0	0	0	0,5	0
315	2021	1,36	1	0	0	0	0	0	1	0,5	0
316	2021	0,99	1	0	0	0	0	0	0	0,5	0
317	2021	-0,28	0	0	0	0	0	0	0	0,5	0
318	2021	0,41	0	0	0	0	0	0	0	0,5	0
319	2021	-0,06	0	1	0	0	0	0	0	0,5	0
320	2021	-1,44	0	0	0	0	0	0	0	0,5	0
321	2021	0,15	0	1	0	0	0	0	0	0,5	0
322	2021	-1,33	1	0	0	0	0	1	0	-0,46	0
323	2021	-0,43	1	0	0	0	0	1	0	0,5	0
324	2021	0,46	1	1	0	0	0	1	0	0,5	0
325	2021	0,73	0	0	0	0	0	1	0	0,18	0
326	2021	0,99	0	1	0	0	0	1	0	0,18	0
327	2021	0,78	0	0	0	0	0	1	0	0,18	0
328	2021	0,04	1	0	0	0	0	1	0	0,5	0
329	2021	0,46	0	0	0	0	0	1	1	0,5	0
330	2021	-0,06	0	0	0	0	0	1	0	0,18	0
331	2021	0,15	0	0	0	0	0	1	0	0,18	0

## 4. Data Setelah dilakukan Pembagian Data

DATA TRAINING											
No	Tahun Masuk	$x_1$	$x_2(1)$	$x_3(1)$	$x_3(3)$	$x_4(1)$	$x_5(1)$	$x_5(1)$	$x_6(1)$	$x_7$	Y
319	2021	-0,06	0	1	0	0	0	0	0	0,5	0
155	2019	0,04	0	0	0	0	0	0	1	0,82	1
248	2020	0,09	0	0	0	0	0	0	0	0,82	0
288	2021	1,47	0	0	0	0	0	0	0	0,18	0
132	2019	0,83	0	0	0	0	0	0	0	0,5	0
260	2020	-0,33	0	0	0	0	0	0	0	0,82	1
197	2020	-0,17	0	0	0	0	1	0	0	0,82	1
169	2019	0,09	0	1	0	0	0	1	0	0,5	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
214	2020	-0,22	0	0	0	0	0	0	0	0,82	1
9	2018	-2,23	0	0	0	0	1	0	0	-0,77	1
237	2020	1,04	0	1	0	0	0	0	0	0,82	0
307	2021	1,89	0	0	0	0	0	0	0	0,18	0
254	2020	1,04	1	0	0	0	0	1	0	0,82	1
311	2021	-0,65	0	0	0	0	0	0	1	0,5	0
145	2019	0,67	0	0	0	0	0	0	0	0,5	0
131	2019	0,36	0	1	0	0	0	0	0	0,5	0

DATA TRAINING											
No	Tahun Masuk	$x_1$	$x_2(1)$	$x_3(1)$	$x_3(3)$	$x_4(1)$	$x_5(1)$	$x_5(1)$	$x_6(1)$	$x_7$	Y
143	2019	-0,06	0	0	0	0	0	0	0	0,18	1
111	2019	-0,75	0	0	0	0	1	0	0	0,5	1
6	2018	1,04	0	1	0	0	1	0	0	-1,41	0
91	2018	-1,33	0	1	0	0	0	1	0	-0,46	2
218	2020	0,31	0	1	0	0	0	0	0	0,82	1
211	2020	0,46	1	1	0	0	0	0	0	0,82	1
208	2020	0,31	0	0	0	0	0	0	0	0,18	1
293	2021	-0,7	0	0	0	0	0	0	0	0,18	0
143	2019	-0,06	0	0	0	0	0	0	0	0,18	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
191	2020	0,04	0	1	0	0	1	0	0	0,82	1
271	2021	-0,54	0	0	0	0	1	0	0	0,18	0
203	2020	0,25	0	1	0	0	1	0	0	1,14	1
96	2018	-0,96	1	1	0	0	0	1	0	-0,77	1
330	2021	-0,06	0	0	0	0	0	1	0	0,18	0
152	2019	1,2	0	1	0	0	0	0	0	0,5	1
205	2020	0,99	0	0	0	0	1	0	0	0,82	0
130	2019	2,21	1	1	0	1	0	0	0	0,5	0
191	2020	0,04	0	1	0	0	1	0	0	0,82	1

## Lampiran 2. Output untuk Uji Multikolinearitas

### Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized	t	Sig.	Collinearity Statistics	
		B	Std. Error	Coefficients Beta			Tolerance	VIF
1	(Constant)	7,645	,944		8,098	,000		
	IPK (X1)	-1,336	,237	-,385	-5,649	,000	,508	1,970
	Jenis Kelamin (X2)_LAKI-LAKI	,025	,076	,017	,332	,740	,933	1,072
	Jenis Asal Sekolah (X3)_MA	-,018	,069	-,013	-,261	,794	,913	1,096
	Jenis Asal Sekolah (X3)_SMK	-,316	,191	-,082	-1,654	,099	,954	1,048
	Jenis Pembiayaan (X4)_BIDIKMISI	-,162	,137	-,059	-1,180	,239	,951	1,052
	Jalur Masuk (X5)_SNMPTN	-,044	,076	-,031	-,576	,565	,815	1,227
	Jalur Masuk (X5)_MANDIRI	-,019	,090	-,012	-,211	,833	,789	1,267
	Asal Daerah (X6)_KOTA MALANG	-,191	,143	-,066	-1,335	,183	,972	1,029
	Rasio SKS Lulus (X7)	-2,685	1,406	-,129	-1,909	,057	,522	1,917

a. Dependent Variable: Ketepatan Waktu Kelulusan (Y)

### Lampiran 3. Output untuk Estimasi Parameter, Uji Parsial

		Parameter Estimates						95% Confidence Interval for Exp (B)	
y <sup>a</sup>		B	Std. Error	Wald	df	Sig.	Exp(B)	Lower Bound	Upper Bound
1,00	Intercept	-,678	,267	6,454	1	,011			
	x1 (IPK_std)	-1,143	,239	22,830	1	,000	,319	,200	,510
	x2_lk	,144	,349	,170	1	,680	1,155	,583	2,287
	x3_MA	-,111	,317	,123	1	,725	,895	,481	1,664
	x3_SMK	-1,087	,935	1,351	1	,245	,337	,054	2,107
	x4_Bidikmisi	-1,134	,729	2,421	1	,120	,322	,077	1,342
	x5_SNMPTN	,423	,343	1,523	1	,217	1,527	,780	2,990
	x5_Mandiri	,437	,411	1,133	1	,287	1,549	,692	3,466
	x6_Kota Malang	-,232	,593	,153	1	,695	,793	,248	2,535
	x7 (SKS_std)	-,014	,263	,003	1	,958	,986	,589	1,651
2,00	Intercept	-1,779	,413	18,564	1	,000			
	x1 (IPK_std)	-,967	,392	6,089	1	,014	,380	,177	,820
	x2_lk	,343	,571	,361	1	,548	1,409	,460	4,311
	x3_MA	-,725	,597	1,474	1	,225	,485	,150	1,561
	x3_SMK	-1,774	1,576	1,267	1	,260	,170	,008	3,726
	x4_Bidikmisi	-1,038	1,320	,619	1	,431	,354	,027	4,704
	x5_SNMPTN	-,354	,657	,290	1	,590	,702	,194	2,544
	x5_Mandiri	,171	,706	,059	1	,808	1,187	,297	4,738
	x6_Kota Malang	-20,052	,000	.	1	.	1,956E-9	1,956E-9	1,956E-9
	x7 (SKS_std)	-,804	,366	4,815	1	,028	,448	,218	,918

a. The reference category is: .00.

**Lampiran 4.** *Output* untuk Estimasi Parameter Model Revisi, Uji Parsial, dan *Odds Ratio*

		Parameter Estimates							95% Confidence Interval for Exp (B)	
y <sup>a</sup>		B	Std. Error	Wald	df	Sig.	Exp(B)	Lower Bound	Upper Bound	
1,00	Intercept	-,575	,152	14,355	1	,000				
	x1 (IPK_std)	-1,186	,232	26,122	1	,000	,306	,194	,481	
	x7 (SKS_std)	,063	,252	,063	1	,802	1,065	,650	1,746	
2,00	Intercept	-2,172	,280	60,099	1	,000				
	x1 (IPK_std)	-,945	,367	6,620	1	,010	,389	,189	,798	
	x7 (SKS_std)	-,780	,359	4,711	1	,030	,459	,227	,927	

a. The reference category is: ,00.



**Lampiran 5. Output** untuk Uji Signifikansi Parameter Secara Simultan**Model Fitting Information**

Model	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	473,511			
Final	387,369	86,142	18	,000

**Lampiran 6.** *Output* untuk Uji Signifikansi Parameter Secara Simultan Model Revisi

**Model Fitting Information**

Model	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	473,511			
Final	400,489	73,022	4	,000

**Lampiran 7. Output** untuk Uji Kesesuaian Model**Goodness-of-Fit**

	Chi-Square	df	Sig.
Pearson	516,499	498	,274
Deviance	397,716	498	1,000

**Lampiran 8.** *Output* untuk Pengukuran Keباikan Model**Pseudo R-Square**

Cox and Snell	,242
Nagelkerke	,289
McFadden	,153

**Lampiran 9. Output** untuk Evaluasi Kinerja Model

**y \* Predicted Response Category \* datasplit Crosstabulation**

datasplit			Predicted Response Category			Total	
			,00	1,00	2,00		
,00	y	,00	Count	34	4	0	38
			% within y	89,5%	10,5%	0,0%	100,0%
	1,00	,00	Count	15	8	0	23
			% within y	65,2%	34,8%	0,0%	100,0%
	2,00	,00	Count	3	2	1	6
			% within y	50,0%	33,3%	16,7%	100,0%
Total			Count	52	14	1	67
			% within y	77,6%	20,9%	1,5%	100,0%

## RIWAYAT HIDUP



Muhamad Bayu Ramdhani, lahir di Jakarta pada 5 November 2003, biasa dipanggil Bayu. Penulis merupakan anak ketiga dari Bapak Muhamad Ali dan Ibu Nia Sriyati. Kakak-kakak penulis bernama Muhamad Resno Setiadi dan Muhamad Novianto Isra Rahmani. Penulis telah menempuh pendidikan mulai dari TK Trikari yang lulus pada tahun 2008, Penulis melanjutkan Pendidikan sekolah dasar di SDS Kresna yang lulus pada tahun 2015 dan kemudian melanjutkan pendidikan sekolah menengah pertama di MTsN 14 Jakarta yang lulus pada tahun 2018. Penulis kemudian menempuh sekolah menengah atas selama 3 tahun di SMA Muhammadiyah 4 Jakarta dan lulus pada tahun 2021. Pada tahun yang sama, penulis melanjutkan pendidikan perguruan tinggi di Universitas Islam Negeri Maulana Malik Ibrahim Malang pada Program Studi Matematika di Fakultas Sains dan Teknologi.



**BUKTI KONSULTASI SKRIPSI**

Nama : Muhamad Bayu Ramdhani  
NIM : 210601110027  
Fakultas / Jurusan : Sains dan Teknologi / Matematika  
Judul Skripsi : Pemodelan Regresi Logistik Multinomial untuk Ketepatan Waktu Kelulusan Mahasiswa (Studi Kasus: Program Studi Matematika UIN Maulana Malik Ibrahim Malang)  
Pembimbing I : Dr. Fachrur Rozi, M.Si.  
Pembimbing II : Erna Herawati, M.Pd.

No	Tanggal	Hal	Tanda Tangan
1.	18 November 2024	Konsultasi Topik dan Data	1.
2.	13 Maret 2025	Konsultasi Bab I, II, dan III	2.
3.	19 Maret 2025	Konsultasi Bab I, II, dan III	3.
4.	28 Mei 2025	Konsultasi Bab I, II, dan III	4.
5.	10 September 2025	ACC Bab I, II, dan III	5.
6.	12 September 2025	Konsultasi Kajian Agama Bab I dan II	
7.	15 September 2025	Konsultasi Kajian Agama Bab I dan II	7.
8.	17 September 2025	Konsultasi Kajian Agama Bab I dan II	8.
9.	17 September 2025	ACC Kajian Agama Bab I dan II	9.
10.	18 September 2025	ACC Seminar Proposal	10.
11.	1 Oktober 2025	Konsultasi Revisi Seminar Proposal	11.
12.	6 Oktober 2025	Konsultasi Bab IV	12.
13.	13 November 2025	Konsultasi Bab IV	13.
14.	24 November 2025	Konsultasi Bab IV dan V	14.
15.	25 November 2025	ACC Bab IV dan V	15.



**KEMENTERIAN AGAMA RI  
UNIVERSITAS ISLAM NEGERI  
MAULANA MALIK IBRAHIM MALANG  
FAKULTAS SAINS DAN TEKNOLOGI**

Jl. Gajayana No.50 Dinoyo Malang Telp. / Fax. (0341)558933

16.	25 November 2025	Konsultasi Kajian Agama Bab IV	16. <i>FR</i>
17.	26 November 2025	Konsultasi Kajian Agama Bab IV	17. <i>FR</i>
18.	26 November 2025	ACC Kajian Agama Bab IV	18. <i>FR</i>
19.	27 November 2025	ACC Seminar Hasil	19. <i>FR</i>
20.	4 Desember 2025	Konsultasi Revisi Seminar Hasil	20. <i>FR</i>
21.	15 Desember 2025	Sidang Skripsi	21. <i>FR</i>
22.	22 Desember 2025	ACC Keseluruhan	22. <i>FR</i>

Malang, 22 Desember 2025

Mengetahui,

Ketua Program Studi Matematika



Dr. Fachrud Rozi, M.Si.

NIP. 19800527 200801 1 012