Emasains
Jurnal Edukasi Matematika dan Sains
Volume 15 Nomor 3 Bulan September Tahun 2026
DOI : 10.59672/emasains.v15i3.7142
P-ISSN 2302-2124
E-ISSN 2622 8688

Pengembangan Model Regresi Logistik Untuk Prediksi Risiko Drop Out Mahasiswa Sebagai Kerangka Deteksi Dini Akademik Berbasis Risk Scoring

Finny Aprilia Yulanda[a,*], Elang Putri Sekar Arum[b], Christiani Joice Siallagan[c] , Laura Annisa Purba[d], Naomi Febrina Sitompul[e], Hadijah[f]

[a,b,c,d,e,f] Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Medan, Medan, Indonesia
*email: finnyaprilia.4243260021@mhs.unimed.ac.id

Abstrak. Penelitian ini menganalisis tiga dimensi faktor (akademik, psikologis, dan sosial) yang berasosiasi signifikan dengan risiko drop out mahasiswa dan mengembangkan model regresi logistik sebagai kerangka deteksi dini akademik berbasis risk scoring. Data dikumpulkan dari 270 mahasiswa aktif pada sepuluh program studi di Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA), menggunakan kuesioner yang pada tahap awal mengukur empat dimensi risiko: akademik, ekonomi, psikologis, dan sosial. Namun, dimensi ekonomi tidak lolos uji validitas dan reliabilitas sehingga model akhir hanya memuat tiga dimensi, yaitu akademik, psikologis, dan sosial. Instrumen diuji validitas (r-drop $\geq$ 0,30) dan reliabilitas (Cronbach Alpha) sebelum dibentuk menjadi skor komposit. Variabel dependen merupakan proksi berbasis persepsi mahasiswa, bukan data status drop out aktual, yang dikategorikan biner melalui median split dan dianalisis dengan Maximum Likelihood Estimation (MLE), dibandingkan dengan Firth regression untuk mengantisipasi quasi-complete separation. Uji Likelihood Ratio menunjukkan model signifikan secara simultan ($p < 0,05$), dengan uji Hosmer-Lemeshow menyatakan model fit ($p = 0,316$). Model mencapai akurasi klasifikasi 70% dan AUC 0,79 pada sampel yang sama dengan yang digunakan untuk membentuk model, menunjukkan diskriminasi moderat namun belum diuji pada data independen. Uji Wald mengidentifikasi skor akademik, psikologis, dan sosial sebagai prediktor signifikan, sedangkan IPK, kehadiran, dan beban SKS tidak signifikan. Mahasiswa dikelompokkan ke dalam tiga kategori risiko yang masih eksploratif yaitu rendah, sedang, dan tinggi melalui risk scoring yang belum divalidasi terhadap data kejadian drop out aktual. Temuan ini memberikan implikasi praktis bagi fakultas dalam merancang sistem deteksi dini dan strategi retensi mahasiswa yang lebih tepat sasaran dan berbasis data, dengan catatan diperlukan validasi lanjutan sebelum diterapkan secara operasional.

Kata Kunci: Drop out mahasiswa, Regresi logistik, Skor risiko, Deteksi dini akademik, Retensi mahasiswa

PENDAHULUAN

Drop out mahasiswa merupakan salah satu permasalahan strategis dalam pendidikan tinggi karena berdampak pada efektivitas penyelenggaraan akademik, efisiensi pembiayaan, serta capaian kinerja institusi. Tingginya angka putus studi tidak hanya merugikan perguruan tinggi, tetapi juga berdampak pada mahasiswa secara individu, baik dari aspek ekonomi, sosial, maupun psikologis. Keputusan mahasiswa untuk menghentikan studi secara teoritis

Copyright ©2026 by Author.
Published by Prodi Pendidikan Matematika dan
Prodi Pendidikan Biologi Universitas PGRI Mahadewa Indonesia

JEms Emasains
Jurnal Edukasi Matematika dan Sains P-ISSN 2302-2124
Volume 15 Nomor 3 Bulan September Tahun 2026 E-ISSN 2622 8688
DOI : 10.59672/emasains.v15i3.7142

dipengaruhi oleh integrasi akademik dan sosial serta kondisi personal yang saling berinteraksi (Tinto, 1975; Aulck, 2016).

Perkembangan paradigma data-driven decision making dalam pengelolaan pendidikan tinggi mendorong pemanfaatan analisis statistik untuk mengidentifikasi mahasiswa yang memerlukan pemantauan akademik secara lebih objektif dan terukur. Regresi logistik merupakan salah satu metode yang banyak digunakan untuk memodelkan respons biner, termasuk status drop out aktual maupun kategori risiko yang dibentuk berdasarkan indikator terukur (Lubis & Armansyah, 2025; Cuji Chacha et al., 2023). Metode ini menghasilkan estimasi probabilitas dan nilai odds ratio, sehingga hubungan antara variabel prediktor dan kategori respons dapat diinterpretasikan secara lebih transparan (Hosmer & Lemeshow, 2013).

Kajian survei Lobo et al. (2023) menunjukkan bahwa penelitian prediksi drop out banyak membandingkan regresi logistik dengan berbagai pendekatan machine learning untuk meningkatkan performa prediksi. Pada penelitian berbasis regresi logistik, Lubis dan Armansyah (2025) mengklasifikasikan mahasiswa berpotensi drop out menggunakan data mahasiswa Program Studi Sistem Informasi angkatan 2015-2024, sedangkan Chacha et al. (2025)sedangkan Cuji Chacha et al. (2023) membangun model dari variabel demografis dan akademik, yaitu status pernikahan, usia, jenis kelamin, dan nilai. Studi-studi tersebut menitikberatkan pada data administratif, indikator demografis-akademik, dan performa klasifikasi, sehingga faktor psikologis dan sosial berbasis persepsi mahasiswa belum menjadi fokus yang dianalisis secara simultan bersama indikator akademik dalam satu kerangka regresi logistik yang mudah diinterpretasikan. Integrasi beberapa dimensi risiko dan pengelompokan probabilitas prediksi ke dalam kategori rendah, sedang, dan tinggi berpotensi membantu institusi menentukan prioritas pemantauan dan pendampingan mahasiswa. Meskipun demikian, kategori risiko yang dibangun dari data persepsi perlu dipahami sebagai alat skrining awal dan bukan sebagai penetapan bahwa mahasiswa pasti akan mengalami drop out.

Berdasarkan uraian tersebut, penelitian ini menjawab tiga pertanyaan. Pertama, faktor apa saja yang berasosiasi signifikan dengan kategori risiko drop out mahasiswa FMIPA Universitas Negeri Medan. Kedua, bagaimana model regresi logistik dapat digunakan untuk mengestimasi probabilitas kategori risiko tersebut. Ketiga, bagaimana probabilitas hasil model dapat disusun menjadi klasifikasi risk scoring sebagai kerangka awal deteksi dini akademik.

Penelitian ini bertujuan untuk mengidentifikasi faktor akademik, ekonomi, psikologis, dan sosial (dimensi ekonomi diuji pada tahap instrumen, tetapi tidak masuk model akhir) yang berkaitan dengan risiko drop out mahasiswa FMIPA Universitas Negeri Medan, membangun model regresi logistik biner untuk mengestimasi probabilitas kategori risiko tersebut, serta mengembangkan klasifikasi risk scoring sebagai kerangka awal deteksi dini dan pendampingan akademik berbasis data. Dalam penelitian ini, risiko drop out dioperasionalkan sebagai proksi berbasis skor persepsi mahasiswa dan bukan sebagai status drop out aktual yang tercatat dalam sistem akademik. Kerangka integrasi faktor akademik, sosial, dan personal terhadap keputusan putus studi telah dikemukakan dalam teori Tinto (1975), sehingga kebaruan penelitian ini bukan pada gagasan integrasi dimensi tersebut. Kebaruan penelitian ini terletak pada paket metodologisnya, yaitu (1) operasionalisasi dimensi akademik, psikologis, dan sosial melalui instrumen yang diuji validitas (r-drop) dan reliabilitasnya (Cronbach Alpha) lalu dikombinasikan dengan indikator administratif (IPK, kehadiran, dan beban SKS) dalam satu model regresi logistik, (2) penggunaan regresi Firth sebagai analisis sensitivitas terhadap risiko quasi-complete separation, dan (3) penyusunan risk scoring tiga kategori dari probabilitas

Copyright ©2026 by Author.
Published by Prodi Pendidikan Matematika dan
Prodi Pendidikan Biologi Universitas PGRI Mahadewa Indonesia

JEms Emasains
Jurnal Edukasi Matematika dan Sains P-ISSN 2302-2124
Volume 15 Nomor 3 Bulan September Tahun 2026 E-ISSN 2622 8688
DOI : 10.59672/emasains.v15i3.7142

model pada konteks mahasiswa FMIPA. Berbeda dengan Lubis dan Armansyah (2025) serta Chacha et al. (2025)serta Cuji Chacha et al. (2023) yang bertumpu pada data mahasiswa dan variabel demografis-akademik, penelitian ini menambahkan dimensi psikologis dan sosial berbasis persepsi mahasiswa.

METODE PENELITIAN

Desain Penelitian, Populasi, dan Sampel

Penelitian ini menggunakan pendekatan kuantitatif dengan desain prediktif, yaitu merancang model regresi logistik untuk mengestimasi probabilitas risiko drop out mahasiswa berdasarkan faktor akademik, ekonomi, psikologis, dan sosial. Populasi penelitian adalah mahasiswa aktif Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA) Universitas Negeri Medan pada tahun akademik 2025/2026.

Penentuan jumlah sampel didasarkan pada pendekatan Events Per Variable (EPV) dalam regresi logistik, yang merekomendasikan minimal 10 kejadian (event) untuk setiap variabel independen agar estimasi parameter model stabil dan tidak bias (Peduzzi et al., 1996). Model akhir melibatkan 6 variabel independen (IPK, tingkat kehadiran, jumlah SKS, skor akademik, skor psikologis, dan skor sosial) yang menghasilkan 10 parameter regresi di luar konstanta, karena tiga variabel kategorik dikodekan sebagai variabel dummy. Dengan demikian, jumlah minimal kejadian yang diperlukan adalah $10 \times 10 = 100$ kejadian, yang berarti dibutuhkan sedikitnya 100 responden pada kategori persepsi risiko lebih tinggi. Untuk memenuhi kebutuhan tersebut sekaligus meningkatkan kestabilan model dan kekuatan statistik analisis, penelitian ini menetapkan jumlah sampel minimal sebanyak 270–300 responden. Pada 270 data akhir, terdapat 147 responden pada kategori persepsi risiko lebih tinggi (kode 1), sehingga persyaratan EPV tersebut terpenuhi.

Pengambilan sampel dilakukan dengan teknik simple random sampling terhadap mahasiswa aktif FMIPA yang memenuhi kriteria penelitian, melalui penyebaran kuesioner secara daring menggunakan Google Form. Total data valid yang terkumpul adalah 282 responden, telah melampaui batas minimal yang ditetapkan. Karena distribusi awal menunjukkan proporsi responden Program Studi Statistika relatif lebih besar dibandingkan program studi lain, dilakukan penyesuaian jumlah responden agar distribusi antarprogram studi lebih proporsional. Sebanyak 12 responden Program Studi Statistika tidak diikutsertakan dalam analisis, sehingga jumlah data akhir yang digunakan dalam penelitian ini menjadi 270 responden yang berasal dari seluruh program studi di lingkungan FMIPA Universitas Negeri Medan.

Instrumen dan Operasionalisasi Variabel

Instrumen penelitian berupa kuesioner tertutup yang pada tahap awal terdiri atas empat dimensi prediktor, yaitu akademik, ekonomi, psikologis, dan sosial, serta satu dimensi respons berupa persepsi risiko drop out. Seluruh item diukur menggunakan skala Likert 1–5, dengan skor 1 menyatakan Sangat Tidak Setuju dan skor 5 menyatakan Sangat Setuju. Data pendukung berupa IPK, tingkat kehadiran, dan jumlah SKS dikumpulkan dalam bentuk kategori. Skor yang lebih tinggi pada dimensi akademik menunjukkan persepsi kondisi akademik yang lebih positif, skor psikologis yang lebih tinggi menunjukkan beban psikologis yang lebih besar, sedangkan skor sosial yang lebih tinggi menunjukkan dukungan dan keterlibatan sosial yang lebih kuat.

Pengolahan Data dan Pembentukan Variabel Respons

Copyright ©2026 by Author.
Published by Prodi Pendidikan Matematika dan
Prodi Pendidikan Biologi Universitas PGRI Mahadewa Indonesia

JEms Emasains
Jurnal Edukasi Matematika dan Sains P-ISSN 2302-2124
Volume 15 Nomor 3 Bulan September Tahun 2026 E-ISSN 2622 8688
DOI : 10.59672/emasains.v15i3.7142

Data yang diperoleh dari hasil penyebaran kuesioner terlebih dahulu melalui proses data cleaning untuk mengidentifikasi data yang tidak lengkap, inkonsisten, maupun duplikasi. Selanjutnya dilakukan coding terhadap seluruh variabel sehingga dapat diolah menggunakan perangkat lunak statistik.

Setiap konstruk yang diukur melalui beberapa indikator dibentuk menjadi skor komposit dengan menjumlahkan atau menghitung rata-rata skor item penyusunnya sesuai dengan hasil uji validitas dan reliabilitas instrumen. Skor komposit tersebut selanjutnya digunakan sebagai variabel prediktor dalam pembentukan model regresi logistik.

Variabel respons dalam penelitian ini bukan status drop out aktual, melainkan proksi persepsi risiko drop out yang dibentuk dari skor komposit item D1–D4. Skor komposit dihitung menggunakan rata-rata skor item yang telah memenuhi kriteria validitas. Nilai median skor komposit sebesar 2 digunakan sebagai batas kategorisasi. Responden dengan skor sama dengan atau lebih besar dari median dikategorikan sebagai kelompok dengan persepsi risiko lebih tinggi dan diberi kode 1, sedangkan responden dengan skor di bawah median dikategorikan sebagai kelompok dengan persepsi risiko lebih rendah dan diberi kode 0. Penggunaan median split dilakukan untuk memperoleh respons biner yang sesuai dengan regresi logistik, tetapi dapat mengurangi informasi yang terdapat pada skor kontinu dan karena itu dinyatakan sebagai keterbatasan penelitian.

Sebelum dilakukan pemodelan, seluruh variabel independen diperiksa terhadap kemungkinan terjadinya multikolinearitas menggunakan nilai Variance Inflation Factor (VIF). Untuk variabel kategorik yang memiliki lebih dari dua kategori digunakan Generalized Variance Inflation Factor (GVIF). Variabel dikatakan tidak mengalami multikolinearitas apabila nilai VIF berada di bawah batas yang direkomendasikan dalam literatur, sedangkan GVIF dievaluasi menggunakan nilai

(1)

dengan GVIF adalah Generalized Variance Inflation Factor dan df adalah derajat bebas (Degrees of Freedom).

Selain itu, dilakukan pemeriksaan terhadap kemungkinan terjadinya quasi-complete separation, yaitu kondisi ketika suatu atau beberapa variabel prediktor mampu memisahkan kategori respon secara hampir sempurna sehingga menyebabkan estimasi koefisien regresi logistik menjadi tidak stabil atau gagal dikonvergensikan. Deteksi dilakukan melalui proses estimasi model serta pemeriksaan peringatan (warning) selama proses fitting model.

Reverse Coding, Uji Validitas, dan Uji Reliabilitas Instrumen

Sebelum pengujian validitas dan reliabilitas dilakukan, terlebih dahulu diidentifikasi item-item bermuatan negatif (unfavorable) yang arah pengukurannya berlawanan dengan konstruk yang diukur. Berdasarkan hasil identifikasi, item E1 pada variabel ekonomi dan item P3 pada variabel psikologis merupakan item negatif sehingga perlu dilakukan reverse coding sebelum data digabungkan dengan item lain dalam variabel yang sama. Proses reverse coding pada skala Likert 1–5 dilakukan menggunakan rumus berikut:

$$X' = (k + 1) - X \tag{2}$$

Copyright ©2026 by Author.
Published by Prodi Pendidikan Matematika dan
Prodi Pendidikan Biologi Universitas PGRI Mahadewa Indonesia

JEms Emasains
Jurnal Edukasi Matematika dan Sains P-ISSN 2302-2124
Volume 15 Nomor 3 Bulan September Tahun 2026 E-ISSN 2622 8688
DOI : 10.59672/emasains.v15i3.7142

dengan X' adalah skor hasil reverse coding, X adalah skor jawaban asli responden, dan k adalah nilai skala tertinggi (k = 5 untuk skala Likert 1–5). Dengan demikian, rumus tersebut secara operasional menjadi X' = 6 − X, sehingga item E1 dan P3 ditransformasi menjadi E1_R dan P3_R agar arah pengukurannya konsisten dengan item lain dalam dimensi yang sama.

Uji validitas instrumen dilakukan menggunakan nilai corrected item-total correlation (r-drop), yaitu koefisien korelasi antara skor suatu item dengan skor total variabel setelah item tersebut dikeluarkan dari perhitungan total (Hair et al., 2019). Item dinyatakan valid apabila memiliki nilai r-drop ≥ 0,30, sedangkan item dengan nilai r-drop di bawah 0,30 dinyatakan tidak valid dan dieliminasi dari pembentukan skor variabel (Sekaran & Bougie, 2016). Kriteria ini digunakan secara konsisten pada seluruh dimensi variabel, baik sebelum maupun setelah proses reverse coding, agar hasil pengujian dapat diperbandingkan secara objektif.

Uji reliabilitas instrumen dilakukan menggunakan koefisien Cronbach's Alpha, yang mengukur konsistensi internal antaritem dalam satu variabel (Cronbach, 1951). Interpretasi nilai Cronbach's Alpha dalam penelitian ini mengacu pada kriteria berikut: nilai 0,80–1,00 dikategorikan sangat baik/tinggi, 0,60–0,79 dikategorikan baik, 0,40–0,59 dikategorikan sedang, 0,20–0,39 dikategorikan rendah, dan di bawah 0,20 dikategorikan sangat rendah (Taber, 2018). Pengujian reliabilitas dilakukan setelah proses eliminasi item yang tidak memenuhi kriteria validitas, sehingga nilai Cronbach's Alpha yang dilaporkan mencerminkan konsistensi internal dari item-item yang dipertahankan pada masing-masing variabel.

Hasil pengujian awal menunjukkan bahwa sejumlah item pada variabel ekonomi dan psikologis belum memenuhi kriteria validitas meskipun telah dilakukan reverse coding pada item E1 dan P3. Rincian hasil uji validitas untuk seluruh item disajikan pada Tabel 1, sedangkan ringkasan hasil uji reliabilitas setelah eliminasi item yang tidak valid disajikan pada Tabel 2.

Pembentukan Model Prediksi

Model prediksi risiko drop out mahasiswa dibangun menggunakan regresi logistik biner karena variabel respon terdiri atas dua kategori, yaitu mahasiswa berisiko drop out dan tidak berisiko drop out.

Parameter model diestimasi menggunakan metode Maximum Likelihood Estimation (MLE) untuk memperoleh koefisien regresi yang memberikan peluang maksimum terhadap data yang diamati. Bentuk umum model regresi logistik adalah:

(3)

dengan p merupakan probabilitas mahasiswa berada pada kategori persepsi risiko drop out yang lebih tinggi, sedangkan X1, X2, …, Xn merupakan variabel-variabel prediktor.

Apabila pada proses estimasi ditemukan kondisi quasi-complete separation atau complete separation yang menyebabkan estimasi MLE menjadi bias atau tidak konvergen, maka digunakan Firth Logistic Regression sebagai model pembanding. Metode Firth menerapkan pendekatan penalized maximum likelihood sehingga mampu mengurangi bias estimasi parameter serta menghasilkan koefisien yang tetap stabil pada kondisi ukuran sampel kecil maupun data yang mengalami separation.

Model terbaik dipilih berdasarkan hasil evaluasi performa menggunakan ukuran seperti nilai Akaike Information Criterion (AIC), akurasi klasifikasi, sensitivitas, spesifisitas, Area Under the Receiver Operating Characteristic Curve (AUC), serta interpretasi koefisien regresi yang signifikan. Koefisien model akhir kemudian digunakan sebagai dasar dalam penyusunan

Copyright ©2026 by Author.
Published by Prodi Pendidikan Matematika dan
Prodi Pendidikan Biologi Universitas PGRI Mahadewa Indonesia

Emasains
Jurnal Edukasi Matematika dan Sains
Volume 15 Nomor 3 Bulan September Tahun 2026
DOI : 10.59672/emasains.v15i3.7142
P-ISSN 2302-2124
E-ISSN 2622 8688

risk scoring untuk mengidentifikasi tingkat risiko drop out mahasiswa sebagai sistem deteksi dini akademik.

Tabel 1. Hasil Uji Validitas Item Instrumen (r-drop Awal dan Setelah Reverse Coding/Uji Ulang)

Variabel	Item	r-drop Awal	r-drop Setelah Reverse Coding	Keterangan	Status Akhir
Akademik	A1	0,394	0,394	–	Valid
Akademik	A2	0,336	0,336	–	Valid
Akademik	A3	0,423	0,423	–	Valid
Akademik	A4	0,459	0,459	–	Valid
Ekonomi	E1	0,195	0,195	Reverse coding → E1_R; tetap tidak valid	Dieliminasi
Ekonomi	E2	0,377	0,377	–	Valid
Ekonomi	E3	0,249	0,249	Bukan item negatif (tidak direverse)	Dieliminasi
Ekonomi	E4	0,030	0,030	Bukan item negatif (tidak direverse)	Dieliminasi
Psikologis	P1	0,050	0,050	Bukan item negatif (tidak direverse)	Dieliminasi
Psikologis	P2	0,026	0,026	Bukan item negatif (tidak direverse)	Dieliminasi
Psikologis	P3	-0,167	-0,167	Reverse coding → P3_R; tetap tidak valid	Dieliminasi
Psikologis	P4	0,609	0,696	–	Valid
Psikologis	P5	0,707	0,745	–	Valid
Psikologis	P6	0,550	0,668	–	Valid
Sosial	S1	0,532	0,612	–	Valid
Sosial	S2	0,626	0,696	–	Valid
Sosial	S3	0,536	0,624	–	Valid
Sosial	S4	0,508	0,553	–	Valid
Sosial	S5	0,067	0,067	Bukan item negatif (tidak direverse)	Dieliminasi
Risiko Drop Out	D1	0,628	0,628	–	Valid

Copyright ©2026 by Author.
Published by Prodi Pendidikan Matematika dan
Prodi Pendidikan Biologi Universitas PGRI Mahadewa Indonesia

Emasains
Jurnal Edukasi Matematika dan Sains
Volume 15 Nomor 3 Bulan September Tahun 2026
DOI : 10.59672/emasains.v15i3.7142
P-ISSN 2302-2124
E-ISSN 2622 8688

Risiko Drop Out	D2	0,652	0,652	–	Valid
Risiko Drop Out	D3	0,777	0,777	–	Valid
Risiko Drop Out	D4	0,751	0,751	–	Valid

Keterangan: item dengan r-drop ≥ 0,30 dinyatakan valid; item dengan r-drop < 0,30 dieliminasi dari pembentukan skor variabel.

Tabel 2. Hasil Uji Reliabilitas Instrumen Setelah Eliminasi Item Tidak Valid

Variabel	Item Dipertahankan	Cronbach's Alpha	Kesimpulan
Akademik	A1, A2, A3, A4	0,609	Dapat digunakan dengan catatan
Ekonomi	E2	–	Tidak direkomendasikan
Psikologis	P4, P5, P6	0,839	Reliabel
Sosial	S1, S2, S3, S4	0,805	Reliabel
Risiko Drop Out	D1, D2, D3, D4	0,854	Reliabel

* Variabel ekonomi hanya menyisakan satu item valid (E2) setelah eliminasi sehingga nilai Cronbach's Alpha tidak dapat dihitung. Oleh karena itu, variabel ekonomi tidak dibentuk sebagai konstruk komposit dan tidak dimasukkan ke dalam model regresi logistik utama.

Berdasarkan hasil pengujian akhir, dimensi akademik, psikologis, sosial, dan persepsi risiko drop out memiliki item yang memenuhi kriteria validitas dan dapat digunakan pada tahap pembentukan skor komposit. Reliabilitas dimensi akademik sebesar 0,609 menunjukkan bahwa konstruk tersebut dapat digunakan dengan kehati-hatian, sedangkan dimensi psikologis, sosial, dan persepsi risiko drop out menunjukkan konsistensi internal yang memadai. Dimensi ekonomi hanya menyisakan satu item valid sehingga tidak dapat direpresentasikan sebagai konstruk multi-item dan tidak dimasukkan ke dalam model utama.

HASIL DAN PEMBAHASAN

Hasil Penelitian

Hasil Model Regresi Logistik

Model regresi logistik biner dibangun untuk mengestimasi probabilitas mahasiswa berada pada kategori persepsi risiko drop out yang lebih tinggi. Variabel respons diberi nama risiko_dropout_biner, dengan kode 1 menunjukkan kategori persepsi risiko lebih tinggi dan kode 0 menunjukkan kategori persepsi risiko lebih rendah. Model menggunakan prediktor IPK, tingkat kehadiran, jumlah SKS, skor akademik, skor psikologis, dan skor sosial. Uji Likelihood Ratio menunjukkan bahwa model secara keseluruhan signifikan dibandingkan model tanpa prediktor ($G^2 = 81{,}854$; $df = 10$; $p < 0{,}001$). Hasil ini mengindikasikan bahwa prediktor yang dimasukkan ke dalam model, secara bersama-sama, berasosiasi signifikan terhadap risiko drop out.

Copyright ©2026 by Author.
Published by Prodi Pendidikan Matematika dan
Prodi Pendidikan Biologi Universitas PGRI Mahadewa Indonesia

JEms Emasains
Jurnal Edukasi Matematika dan Sains P-ISSN 2302-2124
Volume 15 Nomor 3 Bulan September Tahun 2026 E-ISSN 2622 8688
DOI : 10.59672/emasains.v15i3.7142

Tabel 3 menyajikan koefisien regresi (B), rasio odds (Odds Ratio/OR), selang kepercayaan 95%, nilai uji Wald, dan nilai p untuk masing-masing prediktor.

Tabel 3. Koefisien Regresi dan Odds Ratio Model Logistik (MLE)

Variabel	B	OR (Exp(B))	95% CI	Wald z	p-value	Ket.
Konstanta	9,6962	16255,0	124,0 – 2,96×10⁶	3,822	0,0001	Sig.
IPK ≥ 3,50 (vs <2,50)	-1,9199	0,147	0,0042 – 3,48	-1,236	0,2163	ns
IPK 2,50–2,99 (vs <2,50)	0,4976	1,64	0,0308 – 100,0	0,259	0,7955	ns
IPK 3,00–3,49 (vs <2,50)	-1,8632	0,155	0,0045 – 3,72	-1,194	0,2325	ns
Kehadiran >85% (vs <75%)	1,3407	3,82	0,388 – 51,6	1,098	0,2724	ns
Kehadiran 75–85% (vs <75%)	1,3628	3,91	0,314 – 62,5	1,024	0,3059	ns
SKS >21 (vs <18)	-0,0091	0,991	0,456 – 2,15	-0,023	0,9817	ns
SKS 18–21 (vs <18)	-0,4361	0,647	0,294 – 1,40	-1,101	0,2708	ns
Skor Akademik	-1,1912	0,304	0,168 – 0,528	-4,079	<0,0001	Sig.**
Skor Psikologis	0,3982	1,49	1,06 – 2,12	2,262	0,0237	Sig.*
Skor Sosial	-1,2862	0,276	0,156 – 0,469	-4,595	<0,0001	Sig.**

Berdasarkan Tabel 3, dari enam kelompok prediktor yang diuji, hanya tiga variabel komposit yang menunjukkan pengaruh signifikan terhadap risiko drop out, yaitu skor akademik, skor psikologis, dan skor sosial, sedangkan IPK, kehadiran, dan SKS kategorikal tidak signifikan secara statistik ($p > 0,05$) meskipun ditemukan pada uji tabulasi silang awal memiliki keterkaitan deskriptif dengan risiko drop out.

Skor akademik berasosiasi negatif dan signifikan dengan kategori persepsi risiko drop out ($B = -1,191$; OR = 0,304; 95% CI: 0,168 - 0,528; $p < 0,001$). Dengan variabel lain dianggap konstan, setiap peningkatan satu satuan skor akademik berasosiasi dengan penurunan odds berada pada kategori persepsi risiko yang lebih tinggi sebesar sekitar 69,6%.

Skor psikologis berasosiasi positif dan signifikan dengan kategori persepsi risiko drop out ($B = 0,398$; OR = 1,49; 95% CI: 1,06–2,12; $p = 0,024$). Setiap peningkatan satu satuan skor psikologis berasosiasi dengan peningkatan odds berada pada kategori persepsi risiko yang lebih tinggi sebesar sekitar 49%.

Skor sosial berasosiasi negatif dan signifikan dengan kategori persepsi risiko drop out ($B = -1,286$; OR = 0,276; 95% CI: 0,156 - 0,469; $p < 0,001$). Setiap peningkatan satu satuan skor sosial berasosiasi dengan penurunan odds berada pada kategori persepsi risiko yang lebih tinggi sebesar sekitar 72,4%.

Variabel IPK, kehadiran, dan jumlah SKS tidak menunjukkan asosiasi yang signifikan secara individual. Interval kepercayaan pada beberapa kategori IPK dan kehadiran terlihat

Copyright ©2026 by Author.
Published by Prodi Pendidikan Matematika dan
Prodi Pendidikan Biologi Universitas PGRI Mahadewa Indonesia

JEms Emasains
Jurnal Edukasi Matematika dan Sains P-ISSN 2302-2124
Volume 15 Nomor 3 Bulan September Tahun 2026 E-ISSN 2622 8688
DOI : 10.59672/emasains.v15i3.7142

sangat lebar, yang menunjukkan rendahnya presisi estimasi dan kemungkinan adanya sel dengan frekuensi sangat kecil. Kondisi tersebut dapat berkaitan dengan quasi-complete separation, tetapi interval kepercayaan yang lebar saja belum cukup untuk membuktikan terjadinya separation. Oleh karena itu, dilakukan pemeriksaan lebih lanjut dan analisis sensitivitas menggunakan regresi Firth.

Evaluasi Kelayakan dan Ketepatan Klasifikasi Model

Kelayakan model dievaluasi melalui beberapa pendekatan, yaitu uji kalibrasi Hosmer-Lemeshow, confusion matrix, dan kurva ROC beserta nilai AUC. Ringkasan hasil evaluasi disajikan pada Tabel 4.

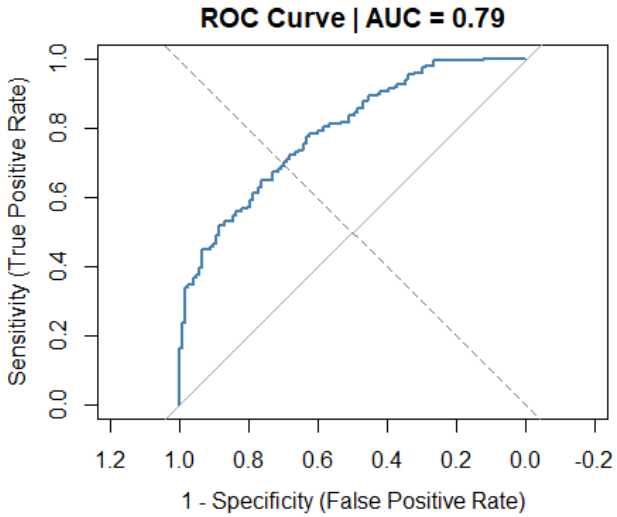
Tabel 4. Ringkasan Evaluasi Kelayakan Model

Uji / Metrik	Nilai	Interpretasi
Likelihood Ratio Test (Omnibus)	$G^2 = 81{,}854$; df = 10; $p < 0{,}001$	Model signifikan secara simultan
Hosmer-Lemeshow GOF	$\chi^2 = 9{,}321$; df = 8; $p = 0{,}316$	Tidak terdapat bukti lack of fit berdasarkan uji Hosmer–Lemeshow
AIC	312,31	-
Akurasi klasifikasi	70,00%	-
Sensitivitas	72,11%	-
Spesifisitas	67,48%	-
AUC (ROC)	0,790	Kemampuan diskriminasi moderat

Uji Hosmer–Lemeshow menghasilkan nilai p sebesar 0,316 sehingga tidak terdapat bukti statistik yang cukup untuk menyatakan adanya ketidaksesuaian model dengan data. Hasil tersebut tidak secara langsung membuktikan bahwa kalibrasi model telah sempurna, tetapi menunjukkan bahwa uji tersebut tidak mendeteksi lack of fit. Dengan ambang klasifikasi 0,50, model menghasilkan akurasi sebesar 70,00%, sensitivitas 72,11%, dan spesifisitas 67,48%. Nilai AUC sebesar 0,790 menunjukkan kemampuan diskriminasi moderat dalam membedakan kategori persepsi risiko lebih tinggi dan lebih rendah. Karena metrik tersebut dihitung pada sampel yang juga digunakan untuk membentuk model, hasilnya merupakan performa pada sampel pengembangan dan masih memerlukan validasi internal serta eksternal.

Copyright ©2026 by Author.
Published by Prodi Pendidikan Matematika dan
Prodi Pendidikan Biologi Universitas PGRI Mahadewa Indonesia

Emasains
Jurnal Edukasi Matematika dan Sains
Volume 15 Nomor 3 Bulan September Tahun 2026
DOI : 10.59672/emasains.v15i3.7142
P-ISSN 2302-2124
E-ISSN 2622 8688



Gambar 1. Kurva ROC Model Regresi Logistik (AUC = 0,7899)

Perbandingan Model MLE dengan Firth Regression (Robustness Check)

Selang kepercayaan yang sangat lebar pada kategori IPK dan kehadiran mengindikasikan potensi quasi-complete separation, yaitu kondisi ketika salah satu sel pada tabulasi silang antara prediktor kategorikal dan variabel dependen bernilai kosong atau sangat kecil, yang dapat menyebabkan estimasi maximum likelihood menjadi tidak stabil. Sebagai langkah robustness check, model diestimasi ulang menggunakan Firth-type penalized likelihood regression (Kosmidis & Firth, 2021), yang secara teoritis lebih tahan terhadap masalah separation dibandingkan estimasi maximum likelihood konvensional (Puhr et al., 2017). Perbandingan koefisien kedua model disajikan pada Tabel 5.

Tabel 5. Perbandingan Koefisien Model MLE dan Firth Regression

Variabel	Koefisien Firth	OR Firth	95% CI
Konstanta	9,1679	9584,476	—
IPK ≥ 3,50 (vs <2,50)	-1,7236	0,178	—
IPK 2,50–2,99 (vs <2,50)	0,2871	1,333	—
IPK 3,00–3,49 (vs <2,50)	-1,6683	0,189	—
Kehadiran >85% (vs <75%)	1,1551	3,174	—
Kehadiran 75–85% (vs <75%)	1,1642	3,203	—
SKS >21 (vs <18)	0,0008	1,001	—
SKS 18–21 (vs <18)	-0,4055	0,667	—
Skor Akademik	-1,1283	0,324	—
Skor Psikologis	0,3800	1,462	—

Copyright ©2026 by Author.
Published by Prodi Pendidikan Matematika dan
Prodi Pendidikan Biologi Universitas PGRI Mahadewa Indonesia

Emasains
Jurnal Edukasi Matematika dan Sains P-ISSN 2302-2124
Volume 15 Nomor 3 Bulan September Tahun 2026 E-ISSN 2622 8688
DOI : 10.59672/emasains.v15i3.7142

Skor Sosial	-1,2145	0,297	—

Koefisien model MLE dan regresi Firth menunjukkan arah hubungan yang konsisten. Perbedaan koefisien pada prediktor utama juga relatif kecil. Hasil ini menunjukkan bahwa kesimpulan mengenai arah asosiasi skor akademik, psikologis, dan sosial relatif tidak berubah setelah diterapkan estimasi penalized. Regresi Firth dalam penelitian ini diposisikan sebagai analisis sensitivitas terhadap kemungkinan ketidakstabilan estimasi, bukan sebagai validasi eksternal model.

Klasifikasi Kategori Risiko

Probabilitas prediksi setiap mahasiswa diperoleh dari model regresi logistik MLE, kemudian dikelompokkan secara eksploratif menjadi tiga kategori, yaitu rendah untuk probabilitas < 0,33, sedang untuk probabilitas 0,33–0,66, dan tinggi untuk probabilitas > 0,66. Ambang tersebut dibentuk menggunakan pendekatan pembagian interval dengan lebar yang sama dan belum merupakan batas diagnostik atau batas intervensi akademik yang telah divalidasi terhadap kejadian drop out aktual (Akçapınar et al., 2019).

Hasil klasifikasi terhadap 270 mahasiswa menunjukkan distribusi pada Tabel 6 berikut.

Tabel 6. Distribusi Responden Berdasarkan Kategori Skor Prediksi

Kategori Risiko	Jumlah (n)	Persentase (%)
Rendah	61	22,59
Sedang	119	44,07
Tinggi	90	33,33

Dalam sampel penelitian, sebanyak 61 responden atau 22,59% diklasifikasikan pada kategori skor prediksi rendah, 119 responden atau 44,07% pada kategori sedang, dan 90 responden atau 33,33% pada kategori tinggi. Distribusi tersebut menggambarkan hasil kategorisasi dalam sampel berdasarkan ambang eksploratif yang digunakan dan tidak dapat langsung ditafsirkan sebagai prevalensi mahasiswa FMIPA yang akan mengalami drop out.

Profil Karakteristik per Kategori Risiko

Untuk memahami karakteristik setiap kelompok, dilakukan perbandingan rata-rata probabilitas risiko dan rata-rata skor komposit tiga dimensi yang signifikan dalam model (akademik, psikologis, dan sosial):

Tabel 7. Profil Karakteristik Berdasarkan Kategori Skor Prediksi

Kategori	n	Rata-rata Probabilitas	Skor Akademik	Skor Psikologis	Skor Sosial
Rendah	61	0,204	4,39	3,10	4,90
Sedang	119	0,492	3,94	3,50	4,46
Tinggi	90	0,844	3,34	3,94	3,69

Pola yang konsisten terlihat pada ketiga dimensi: semakin tinggi kategori risiko, skor akademik dan skor dukungan sosial semakin menurun, sedangkan skor beban psikologis semakin meningkat. Mahasiswa pada kategori risiko rendah menunjukkan persepsi akademik

Copyright ©2026 by Author.
Published by Prodi Pendidikan Matematika dan
Prodi Pendidikan Biologi Universitas PGRI Mahadewa Indonesia

JEms Emasains
Jurnal Edukasi Matematika dan Sains P-ISSN 2302-2124
Volume 15 Nomor 3 Bulan September Tahun 2026 E-ISSN 2622 8688
DOI : 10.59672/emasains.v15i3.7142

yang paling positif (4,39) dan dukungan sosial yang paling kuat (4,90), sementara mahasiswa pada kategori risiko tinggi menunjukkan pola sebaliknya, dengan skor akademik terendah (3,34) dan beban psikologis tertinggi (3,94). Pola ini konsisten dengan hasil uji Wald pada model regresi logistik, di mana ketiga variabel tersebut terbukti berpengaruh signifikan terhadap risiko drop out (skor akademik: OR = 0,304, $p < 0,001$; skor psikologis: OR = 1,49, $p = 0,024$; skor sosial: OR = 0,276, $p < 0,001$).

Selain itu, distribusi IPK dan tingkat kehadiran juga menunjukkan kecenderungan yang sejalan dengan tingkat risiko. Pada kategori risiko rendah, sebanyak 77,05% mahasiswa memiliki IPK ≥ 3,50 dan 95,08% memiliki kehadiran > 85%. Sebaliknya, pada kategori risiko tinggi, proporsi mahasiswa dengan IPK ≥ 3,50 menurun menjadi 42,22%, sementara proporsi mahasiswa dengan kehadiran di bawah 85% meningkat menjadi 20% (gabungan kategori < 75% dan 75–85%), dibandingkan hanya 4,92% pada kategori risiko rendah. Pola deskriptif tersebut menunjukkan bahwa kategori skor prediksi memiliki kecenderungan yang searah dengan IPK dan tingkat kehadiran mahasiswa. Namun, pola ini tidak dapat dianggap sebagai validasi independen karena IPK dan kehadiran juga digunakan sebagai prediktor dalam pembentukan model.

Analisis Asosiasi Tambahan antara Kategori Skor Prediksi, IPK, dan Kehadiran

Analisis tambahan dilakukan untuk menggambarkan hubungan antara kategori skor prediksi dengan IPK dan tingkat kehadiran. Fisher's Exact Test digunakan karena beberapa sel pada tabel kontingensi memiliki frekuensi harapan kurang dari lima. Untuk tabel dengan lebih dari dua baris atau kolom, nilai p diperoleh menggunakan simulasi Monte Carlo dengan 10.000 replikasi (Fisher, 1922).

Tabel 8. Analisis Asosiasi Tambahan Kategori Skor Prediksi dengan IPK dan Kehadiran

Uji Asosiasi	Nilai-p (simulasi, B = 10.000)	Keterangan
Kategori Risiko × IPK	$p < 0,001$	Signifikan ($p < 0,05$)
Kategori Risiko × Kehadiran	0,0211	Signifikan ($p < 0,05$)

Kedua hasil uji menunjukkan nilai-p di bawah 0,05, yang berarti terdapat asosiasi yang signifikan secara statistik antara kategori risiko hasil klasifikasi model dengan IPK maupun tingkat kehadiran mahasiswa. Kategori risiko yang dibentuk dari probabilitas model regresi logistik tidak hanya merupakan hasil matematis semata, tetapi juga selaras dan konsisten dengan indikator akademik riil yang selama ini digunakan institusi untuk memantau performa mahasiswa. Dengan demikian, sistem risk scoring menunjukkan potensi sebagai kerangka awal skrining akademik, tetapi memerlukan validasi lebih lanjut sebelum diterapkan secara operasional.

Pembahasan

Hasil penelitian menunjukkan bahwa skor akademik, skor psikologis, dan skor sosial berasosiasi signifikan dengan kategori persepsi risiko drop out mahasiswa FMIPA Universitas Negeri Medan. Skor akademik dan sosial yang lebih tinggi berasosiasi dengan odds yang lebih rendah untuk berada pada kategori persepsi risiko yang lebih tinggi, sedangkan skor psikologis yang lebih tinggi berasosiasi dengan odds yang lebih tinggi. IPK, tingkat kehadiran, dan jumlah

Copyright ©2026 by Author.
Published by Prodi Pendidikan Matematika dan
Prodi Pendidikan Biologi Universitas PGRI Mahadewa Indonesia

Emasains
Jurnal Edukasi Matematika dan Sains P-ISSN 2302-2124
Volume 15 Nomor 3 Bulan September Tahun 2026 E-ISSN 2622 8688
DOI : 10.59672/emasains.v15i3.7142

SKS tidak menunjukkan asosiasi yang signifikan secara individual setelah seluruh prediktor dimasukkan secara simultan ke dalam model.

Model regresi logistik MLE secara keseluruhan signifikan berdasarkan uji Likelihood Ratio. Uji Hosmer–Lemeshow tidak menunjukkan bukti ketidaksesuaian model dengan data. Pada sampel pembentukan model, model menghasilkan akurasi sebesar 70,00%, sensitivitas 72,11%, spesifisitas 67,48%, dan AUC sebesar 0,790. Analisis sensitivitas menggunakan regresi Firth menghasilkan arah dan besaran koefisien yang relatif konsisten dengan estimasi MLE, khususnya pada skor akademik, psikologis, dan sosial.

Probabilitas hasil model dikelompokkan secara eksploratif menjadi kategori rendah, sedang, dan tinggi. Dalam sampel penelitian, sebanyak 22,59% responden berada pada kategori skor prediksi rendah, 44,07% pada kategori sedang, dan 33,33% pada kategori tinggi. Kategori tersebut menunjukkan asosiasi dengan IPK dan tingkat kehadiran. Namun, hasil asosiasi tersebut tidak merupakan validasi eksternal karena kedua indikator juga digunakan sebagai prediktor dalam proses pembentukan model.

Penelitian ini memiliki beberapa keterbatasan. Pertama, variabel respons merupakan proksi berbasis skor persepsi risiko drop out yang dikategorikan menggunakan median split dan bukan status drop out aktual. Oleh karena itu, probabilitas hasil model tidak dapat ditafsirkan sebagai probabilitas pasti bahwa seorang mahasiswa akan menghentikan studinya. Kedua, dimensi ekonomi hanya menyisakan satu item valid sehingga tidak dapat dibentuk sebagai konstruk komposit dan tidak dimasukkan ke dalam model utama. Ketiga, interval kepercayaan yang lebar pada beberapa kategori IPK dan kehadiran menunjukkan estimasi yang kurang presisi dan kemungkinan adanya sel dengan frekuensi kecil. Keempat, performa model masih dihitung pada sampel pembentukan model sehingga berpotensi memberikan estimasi yang optimistis. Kelima, desain cross-sectional hanya menunjukkan hubungan asosiatif dan tidak dapat digunakan untuk menyimpulkan hubungan sebab-akibat. Selain itu, prosedur pengambilan dan pengurangan responden perlu dipertimbangkan dalam menilai generalisasi hasil penelitian.

Penelitian selanjutnya perlu melakukan validasi internal menggunakan bootstrap atau cross-validation serta validasi eksternal pada kelompok mahasiswa dan periode akademik yang berbeda. Penggunaan status drop out aktual dari sistem informasi akademik dan desain longitudinal juga diperlukan untuk mengevaluasi kemampuan model dalam memprediksi kejadian putus studi yang sebenarnya. Instrumen ekonomi perlu dikembangkan dan diuji ulang, sedangkan ambang kategori rendah, sedang, dan tinggi perlu ditentukan berdasarkan kebutuhan intervensi, sensitivitas, spesifisitas, serta konsekuensi kesalahan klasifikasi. Dengan pengembangan tersebut, model dapat diarahkan menjadi sistem deteksi dini akademik yang lebih kuat, teruji, dan dapat diterapkan secara bertanggung jawab.

SIMPULAN DAN SARAN

Simpulan

Penelitian ini menunjukkan bahwa risiko *drop out* mahasiswa FMIPA Universitas Negeri Medan berkaitan dengan aspek akademik, psikologis, dan sosial. Skor akademik dan dukungan sosial yang lebih baik cenderung berkaitan dengan risiko yang lebih rendah,

Copyright ©2026 by Author.
Published by Prodi Pendidikan Matematika dan
Prodi Pendidikan Biologi Universitas PGRI Mahadewa Indonesia

JEms Emasains
Jurnal Edukasi Matematika dan Sains P-ISSN 2302-2124
Volume 15 Nomor 3 Bulan September Tahun 2026 E-ISSN 2622 8688
DOI : 10.59672/emasains.v15i3.7142

sedangkan beban psikologis yang lebih tinggi berkaitan dengan risiko yang lebih tinggi. Sementara itu, IPK, tingkat kehadiran, dan jumlah SKS tidak menunjukkan asosiasi yang signifikan secara individual setelah seluruh prediktor dianalisis secara simultan. Model regresi logistik yang dikembangkan dapat digunakan untuk mengestimasi probabilitas kategori persepsi risiko *drop out* dan menunjukkan kemampuan diskriminasi yang moderat pada data penelitian. Probabilitas tersebut selanjutnya dapat digunakan sebagai dasar penyusunan *risk scoring* dalam tiga kategori risiko sebagai kerangka awal deteksi dini akademik. Temuan ini menjawab celah penelitian pada Pendahuluan, yaitu bahwa dimensi psikologis dan sosial berbasis persepsi mahasiswa dapat dianalisis secara simultan bersama indikator akademik dalam satu kerangka regresi logistik yang mudah diinterpretasikan, dengan skor akademik, psikologis, dan sosial sebagai prediktor yang signifikan. Kontribusi penelitian ini terletak pada paket metodologis berupa instrumen yang divalidasi, analisis sensitivitas Firth, dan risk scoring tiga kategori yang bersifat eksploratif, bukan pada klaim integrasi dimensi yang telah dikemukakan teori Tinto (1975). Namun, karena respons yang digunakan masih berupa proksi persepsi risiko dan model belum divalidasi pada data independen maupun status *drop out* aktual, *risk scoring* yang dihasilkan belum dapat digunakan sebagai alat prediksi operasional.

Saran

Penelitian selanjutnya disarankan menggunakan data status *drop out* aktual dari sistem informasi akademik serta desain longitudinal agar kemampuan model dalam memprediksi kejadian *drop out* yang sebenarnya dapat dievaluasi dengan lebih tepat. Validasi internal melalui *bootstrap* atau *cross-validation* dan validasi eksternal pada kelompok mahasiswa maupun periode akademik yang berbeda juga diperlukan untuk menilai kestabilan dan kemampuan generalisasi model. Selain itu, instrumen pada dimensi ekonomi perlu dikembangkan dan diuji kembali agar dapat direpresentasikan secara lebih baik dalam model. Batas kategori risiko rendah, sedang, dan tinggi juga perlu ditentukan berdasarkan hasil validasi serta mempertimbangkan kebutuhan intervensi dan konsekuensi kesalahan klasifikasi. Dengan demikian, model *risk scoring* dapat dikembangkan menjadi sistem deteksi dini akademik yang lebih akurat, teruji, dan dapat digunakan secara bertanggung jawab dalam mendukung upaya pendampingan dan retensi mahasiswa..

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Ibu Hadijah selaku dosen pembimbing atas bimbingan, arahan, dan masukan yang sangat berharga selama proses penelitian ini berlangsung. Ucapan terima kasih juga disampaikan kepada Universitas Negeri Medan atas dukungan fasilitas dan lingkungan akademik yang kondusif bagi pelaksanaan penelitian ini. Penelitian ini juga didukung oleh pendanaan dari Student Grant, yang sangat membantu kelancaran proses penelitian. Tidak lupa, penulis mengucapkan terima kasih kepada seluruh responden yang telah bersedia meluangkan waktu untuk berpartisipasi dalam penelitian ini.

DAFTAR PUSTAKA

Akçapınar, G., Hasnine, M. N., Majumdar, R., Flanagan, B., & Ogata, H. (2019). Developing an early-warning system for spotting at-risk students by using eBook interaction logs. Smart Learning Environments, 6(4).

Copyright ©2026 by Author.
Published by Prodi Pendidikan Matematika dan
Prodi Pendidikan Biologi Universitas PGRI Mahadewa Indonesia

Emasains
Jurnal Edukasi Matematika dan Sains P-ISSN 2302-2124
Volume 15 Nomor 3 Bulan September Tahun 2026 E-ISSN 2622 8688
DOI : 10.59672/emasains.v15i3.7142

Aulck, L. (2016). Predicting student dropout in higher education. PLoS ONE, 11(2).

Christou, V., Kermanidis, K. L., Karydis, K., Kanellopoulos, D., & Chatzikonstantinou, P. (2023). Performance and early drop prediction for higher education students using machine learning. Expert Systems with Applications, 225, Article 120079.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16(3), 297–334.

Cuji Chacha, B. R., Gavilanes López, W. L., & Pérez Constante, M. B. (2023). Predictive model of student dropout based on logistic regression. ESPOCH Congresses: The Ecuadorian Journal of S.T.E.A.M., 3(1), 630–656. https://doi.org/10.18502/espoch.v3i1.14477

Fisher, R. A. (1922). On the interpretation of χ^2 from contingency tables, and the calculation of P. Journal of the Royal Statistical Society, 85(1), 87–94.

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. (2013). Applied logistic regression. Wiley.

Kosmidis, I., & Firth, D. (2021). Jeffreys-prior penalty, finiteness and shrinkage in binomial-response generalized linear models. Biometrika, 108(1), 71–82.

Lobo, J. L., Del Ser, J., Bifet, A., & Kasabov, N. (2023). A survey on machine learning for student dropout prediction. Applied Sciences.

Lubis, D. R. B., & Armansyah, A. (2025). Klasifikasi mahasiswa berpotensi dropout menggunakan metode regresi logistik. Techno.Com.

Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., & Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. Journal of Clinical Epidemiology, 49(12), 1373–1379.

Puhr, R., Heinze, G., Nold, M., Lusa, L., & Geroldinger, A. (2017). Firth's logistic regression with rare events: Accurate effect estimates and predictions? Statistics in Medicine, 36, 2302–2317.

Sekaran, U., & Bougie, R. (2016). Research methods for business: A skill-building approach (7th ed.). Wiley.

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. Research in Science Education, 48, 1273–1296.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. Review of Educational Research, 45(1), 89–125.

Vaarma, M., & Li, H. (2024). Predicting student dropouts with machine learning: An empirical study in Finnish higher education. Technology in Society, 76, Article 102474.

Copyright ©2026 by Author.
Published by Prodi Pendidikan Matematika dan
Prodi Pendidikan Biologi Universitas PGRI Mahadewa Indonesia