

BAB I

PENDAHULUAN

1.1. Latar Belakang

Pemilihan Umum (Pemilu) merupakan salah satu pilar utama dalam sistem demokrasi yang memungkinkan rakyat untuk memilih pemimpin serta menentukan arah kebijakan pemerintah. Kualitas pemilu tergantung pada tinggi rendahnya tingkat partisipasi, karena dari partisipasi ini akan terlihat seberapa besar masyarakat menaruh perhatiannya pada masalah negara (Muhaling, 2013). Penyelenggaraan pemilu di Indonesia merupakan wujud nyata penerapan demokrasi yang dilaksanakan setiap lima tahun sekali yang memberikan peran kepada masyarakat untuk dapat ikut secara langsung memilih pejabat pemerintahan. Tingkat partisipasi pemilih menjadi salah satu indikator kesehatan demokrasi suatu negara. Partisipasi pemilih yang merupakan salah satu komponen keberlangsungan demokrasi juga akan berdampak pada siapa yang akan memenangkan pemilu yang menguasai hajat hidup orang banyak (Miriam Budiarjo, 2015). Tingkat partisipasi yang rendah dapat mengindikasikan adanya ketidakpuasan atau ketidakpercayaan terhadap sistem politik. Pentingnya peran pemilih juga tentu dibebankan kepada generasi millennial yakni mahasiswa atau pemuda dalam memilih pejabat negara. Generasi muda atau pemuda harus mampu mengawasi pelaksanaan proses politik yang adil agar dapat secara efektif berperan sebagai agen perubahan dan mengarahkan demokrasi ke arah yang lebih baik. Pemuda dapat meningkatkan keterlibatan mereka dalam politik dengan menjadi lebih aktif sebagai organisator, peserta kegiatan, dan pengawas proses politik (Lestari et al., 2019). Oleh karena itu, perlu diteliti hal yang paling mempengaruhi partisipasi dari masyarakat terutama generasi muda pada saat pemilu 2024.

Bentuk pengaruh partisipasi politik generasi muda cenderung menunjukkan perubahan dibandingkan generasi sebelumnya. Jika dahulu partisipasi politik mendapatkan informasi melalui surat kabar, radio, dan berita melalui salah satu stasiun televisi. Pengaruh generasi muda dalam berpartisipasi dipandang sesuatu yang 'baru' yaitu melalui media sosial, kepercayaannya terhadap sistem politik yang ada, serta lingkungan organisasi yang diikuti juga mempengaruhi keinginan generasi muda untuk berpartisipasi pada pemilu 2024. Generasi muda sering kali dianggap sebagai kelompok masyarakat yang paling tidak peduli dengan persoalan politik, yang sering kali mengalami putus hubungan dengan komunitasnya, yang tidak berminat pada proses dan persoalan politik, yang memiliki tingkat kepercayaan rendah pada politisi serta sinis terhadap berbagai lembaga politik dan pemerintahan (Haste & Hogan, 2006).

Namun, sejumlah studi menunjukkan kekeliruan pandangan sebelumnya yang menganggap generasi muda tidak tertarik pada politik terutama pemilihan umum. Generasi muda adalah kelompok yang dinilai paling peduli terhadap berbagai isu politik (Harris, Wyn & Younes, 2010). Survei yang dilakukan Global Web Index

(2019) menunjukkan bahwa Indonesia memiliki pengguna internet yang paling banyak menggunakan media sosial diantara negara Asia lainnya yaitu (61,6%), sedangkan Jepang (21%), singapura (55,5%) dan Thailand (52,6%). Hal ini menunjukkan media sosial sangat mempengaruhi pandangan generasi muda dalam menilai pemilihan umum dan kampanye setiap pasangan calon pemimpin yang terlibat.

Selain itu, generasi muda dalam berorganisasi juga mempengaruhi pandangan mereka dalam berperan aktif saat proses demokrasi. Keaktifan diri dalam organisasi, baik itu organisasi kemahasiswaan, kepemudaan, sosial atau politik dapat memperluas wawasan mereka tentang isu-isu politik. Melalui interaksi dengan sesama anggota, pemimpin organisasi, dan tokoh-tokoh masyarakat, individu dapat mengembangkan koneksi yang memperkuat komitmen mereka terhadap partisipasi politik. Jaringan ini juga dapat mempermudah saluran komunikasi efektif untuk menyebarkan informasi terkait pemilu dan mendorong partisipasi. Keterlibatan dalam organisasi mendidik individu tentang pentingnya nilai-nilai demokratis seperti partisipasi, tanggung jawab, dan transparansi. Nilai-nilai ini menjadi landasan bagi anggota organisasi untuk terlibat aktif dalam pemilu, baik dalam memberikan suara maupun dalam kegiatan lain seperti kampanye, pengawas pemilu, serta penentu pilihan. Di lingkungan kampus, mahasiswa berada pada fase kritis dalam membentuk pemahaman dan preferensi politik mereka. Pendidikan politik di kampus dapat membantu mahasiswa memahami konsep-konsep dasar politik, seperti demokrasi, hak asasi manusia, kebebasan berpendapat, dan partisipasi politik (Loficha Metesa, 2024). Oleh karena itu, pendidikan politik yang diterima ketika mengikuti organisasi sangat diperlukan untuk menjadi warga negara yang bertanggungjawab dan berpartisipasi aktif dalam kehidupan bermasyarakat.

Kepercayaan pada sistem politik menjadi faktor utama yang mempengaruhi berdasarkan pandangan setiap generasi muda terutama mahasiswa. Mahasiswa akan lebih selektif dalam memilih calon presiden dan wakil presiden karena lebih mempertimbangkan karir masing-masing calon dan kinerja di masa lalu seperti partai, ras, dan agama dibandingkan yang lain (Loficha Metesa, 2024) Mahasiswa yang memiliki kepercayaan tinggi pada sistem politik cenderung berpartisipasi dalam pemilu karena mereka percaya bahwa suara hasil pemilu yang dapat membawa perubahan. Pengaruh dari Kepercayaan pada sistem politik, media sosial, maupun keaktifan organisasi di dalam diri seorang mahasiswa dapat menjadi faktor keterlibatan mereka dalam memilih pemimpin. Kurangnya kesadaran politik atau rendahnya pendidikan politik bagi para pemilih pemula dikhawatirkan akan menurunkan tingkat partisipasi politik pada pemilu yang akan datang. (Eta Yuni Lestari, 2018)

Dalam menyelesaikan partisipasi anggota KM FMIPA dalam pemilu 2024 digunakan metode regresi logistik di mana terdapat dua variabel dependen (biner) yaitu kemungkinan berpartisipasi atau tidak berpartisipasi. Hasil regresi logistik dapat dengan mudah diinterpretasikan untuk memahami hubungan antara variabel

independen dan probabilitas dari variabel dependen. Dengan demikian, perlu adanya analisis lebih lanjut mengenai partisipasi anggota KM FMIPA Unhas pada pemilu 2024 agar terlihat variabel mana yang paling berpengaruh. Oleh karena itu, berdasarkan permasalahan tersebut maka penulis mengangkat tugas akhir dengan judul **“Analisis Regresi Logistik untuk Menilai Pengaruh Media Sosial, Keaktifan di Organisasi, dan Kepercayaan pada Sistem Politik terhadap Partisipasi Anggota KM FMIPA Unhas pada Pemilu 2024”**.

1.2. Rumusan Masalah

Berdasarkan latar belakang di atas, maka rumusan yang akan dikaji dalam penelitian ini sebagai berikut:

1. Bagaimana pengaruh penggunaan media sosial terhadap partisipasi politik anggota KM FMIPA Unhas pada Pemilu 2024
2. Bagaimana keaktifan dalam organisasi mempengaruhi partisipasi politik anggota KM FMIPA Unhas pada Pemilu 2024
3. bagaimana kepercayaan pada sistem politik berpengaruh terhadap partisipasi politik anggota KM FMIPA Unhas pada Pemilu 2024
4. Apa variabel yang paling mempengaruhi anggota KM FMIPA Unhas menentukan pilihan saat Pemilu 2024
5. Apakah model regresi logistik dapat secara akurat memprediksi partisipasi politik berdasarkan variabel-variabel independen yang diteliti.

1.3. Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk menganalisis pengaruh media sosial, keaktifan di organisasi, dan kepercayaan pada sistem politik terhadap partisipasi anggota KM FMIPA Unhas pada Pemilu 2024 dengan Algoritma Regresi Logistik

1.4. Manfaat Penelitian

Manfaat penelitian ini adalah sebagai berikut:

1. Menilai faktor-faktor yang mempengaruhi partisipasi pemilih, Mengidentifikasi kebutuhan informasi politik di kalangan mahasiswa, serta menilai dampak yang paling berpengaruh anggota KM FMIPA terhadap menentukan pilihan.
2. Meningkatkan pemahaman akademik tentang partisipasi politik dan wawasan untuk strategi kampanye politik.

1.5. Landasan Teori

1.5.1 Pengertian Pemilihan Umum (Pemilu)

Menurut Undang-Undang Nomor 15 Tahun 2011 tentang Penyelenggara Pemilihan Umum Pasal 1 Ayat 1 berbunyi “Pemilihan Umum, selanjutnya disingkat Pemilu, adalah sarana pelaksanaan kedaulatan rakyat yang diselenggarakan secara langsung, umum, bebas, rahasia, jujur, dan adil dalam Negara Kesatuan Republik

Indonesia berdasarkan Pancasila dan Undang-Undang Dasar Negara Republik Indonesia Tahun 1945. Pemilu adalah arena kompetisi untuk mengisi jabatan-jabatan politik di pemerintahan yang didasarkan pada pilihan formal dari warga negara yang memenuhi syarat. Dilansir dari situs resmi Komisi Pemilihan Umum (KPU), pengertian pemilu adalah salah satu pilar utama dari proses akumulasi kehendak masyarakat. Pemilu sekaligus merupakan proses demokrasi untuk memilih pemimpin. proses pemilu melibatkan berbagai tahapan, termasuk kampanye politik, pemungutan suara, dan perhitungan suara. Melalui pemilu, masyarakat dapat menyampaikan preferensi politik mereka dan berpartisipasi dalam pembentukan arah kebijakan negara.

Berdasarkan Undang - Undang Republik Indonesia No. 7 tahun 2017 tentang Pemilihan Umum Bab 2 Pasal 4 yaitu pengaturan penyelenggaraan, pemilu memiliki beberapa tujuan, yaitu sebagai berikut:

1. Tujuan pemilu adalah memperkuat sistem ketatanegaraan yang demokratis
2. Tujuan pemilu adalah mewujudkan pemilu yang adil dan berintegritas
3. Tujuan pemilu adalah menjamin konsistensi pengaturan sistem pemilu
4. Tujuan pemilu adalah memberikan kepastian hukum dan mencegah duplikasi dalam pengaturan pemilu, dan:
5. Tujuan pemilu adalah mewujudkan pemilu yang efektif dan efisien.

1.5.2 Partisipasi Generasi Z

Partisipasi generasi Z merujuk pada pemahaman dan keterampilan politik yang dimiliki oleh individu atau kelompok generasi Z. Generasi Z adalah generasi kerja terbaru, yang lahir antara tahun 1995 sampai 2012 yang bisa disebut juga sebagai generasi net atau generasi internet (Stillman, 2017). Partisipasi politik memiliki definisi sebagai kemampuan individu dalam memahami, menganalisis dan berpartisipasi dalam politik, Upaya untuk dapat memahami literasi politik dibutuhkan pemikiran yang kritis, analisis, serta memahami suatu hal dari perspektif orang lain. Generasi z cenderung mengenal teknologi dan internet sedini mungkin. Dengan memahami pola bermedia dan preferensi generasi Z, partai politik maupun dapat memberikan informasi politik dengan lebih baik dan memastikan kesadaran politik yang lebih luas di kalangan mereka.

Menurut Sukardi (2022), Media sosial, Keluarga, Pendidikan, dan pertemanan adalah beberapa sumber lahirnya persepsi politik pada generasi Z. Generasi Z sebagai pemilih pemula dinyatakan sebagai kelompok yang mendominasi pemilik suara pemilih muda dengan rentang usianya antara 17 - 37 tahun. Secara spesifik, peran yang dapat dilakukan gen z selaku generasi yang dikatakan aktif dalam internet sangat banyak dan tentu akan sangat mempengaruhi masa menjelang hingga keberlangsungan pemilihan yang akan diadakan.

Bagi generasi Z untuk memikirkan kepentingan sendiri sudah menjadi beban, apalagi jika ditambah dengan memikirkan kepentingan politik negara. Inilah

beberapa alasan mereka untuk tidak memilih (golput) karena tidak yakin suara yang mereka berikan akan didengar ketika membela keadilan (Irma, Alya, dkk, 2024). Banyak yang masih bingung mencari informasi yang tepat, valid dan terpercaya tentang berbagai calon presiden mendatang. Sehingga pengaruh kampanye politik yang dilakukan setiap pasangan calon presiden dari media sosial sangat mempengaruhi tingkat partisipasi generasi Z.

Partisipasi pemilih muda harus dibangun karena suara generasi muda merupakan bentuk tanggungjawab terhadap proses keberlanjutan kepemimpinan. Pemilu bukan hanya sekadar upaya memperoleh suara pemilih, namun lebih substansial yaitu peningkatan pemahaman dan kesadaran sebagai masyarakat khususnya generasi Z yang sudah mulai apatis dalam proses pemilu yang merupakan proses legal pergantian pemimpin negara. Pemuda sebagai generasi selanjutnya kepada yang lebih paham tentang pemilu karena suara mereka harus digunakan dengan semurni mungkin agar terhindar dari kecurangan oleh orang yang tidak bertanggungjawab.

1.5.3 Regresi Logistik

Regresi logistik adalah metode analisis statistika yang digunakan untuk mengklasifikasikan data dengan tipe kategori. Regresi logistik bermanfaat untuk data biner atau dua kemungkinan antara peubah respon (*dependent variable*) yang memiliki dua kategori atau lebih dengan satu atau lebih peubah penjelas (*independent variable*) berskala interval (Hosmer dan Lemeshow, 2000). Terdapat dua model dalam regresi logistik, yaitu regresi logistik biner dan regresi logistik multinomial. Regresi logistik biner digunakan apabila variabel dependen dari data bersifat dikotomi. Sedangkan apabila variabel dependen yang digunakan terdiri lebih dari dua kategori, maka model regresi logistik yang tepat adalah regresi logistik multinomial (Ae, 2013).

Metode regresi logistik biner merupakan metode yang digunakan untuk menggambarkan hubungan satu atau lebih variabel independen terhadap variabel dependen. Variabel dependen yang digunakan berkategori diskrit dengan dua kemungkinan, yaitu sukses dan gagal. Kejadian sukses biasanya dinotasikan $Y = 1$, sedangkan kejadian gagal dinotasikan dengan $Y = 0$ (Hosmer, 2000).

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_j x_j + \varepsilon \quad (1)$$

dengan:

Y = Variabel respon

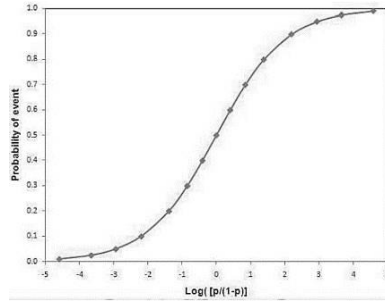
x = Variabel prediktor

β_0 = Konstanta

β_j = Koefisien regresi untuk parameter ke-j

ε = Galat acak

Dalam pemodelan, diasumsikan bahwa variabel biner ini saling bebas satu dengan yang lainnya, sehingga jumlah dari variabel biner akan memiliki sebaran binom. Model regresi logistik diturunkan dari fungsi sigmoid seperti gambar berikut (Moh.Salah,2024):



Gambar 1. Kurva model regresi logistik

dengan variabel prediktor dinyatakan pada persamaan berikut :

$$\pi(x) = \frac{e^{(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}}{1 + e^{(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}} \quad (2)$$

atau

$$\pi(x) = \hat{y}_i = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}} \quad (3)$$

Keterangan:

$\pi(x)$: Peluang kejadian sukses dengan nilai probabilitas $0 \leq \pi(x) \leq 1$

k : Banyaknya variabel prediktor

i : Data ke-i

$\pi(x)$ merupakan fungsi yang non linear yang perlu ditransformasikan ke dalam bentuk logit untuk memperoleh fungsi yang linear agar hubungan antara variabel independen dan variabel dependen dapat terlihat (Rizki dkk, 2015). Dapat diuraikan dengan menggunakan transformasi logit dari $\pi(x)$ (Nirwana, 2015).

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

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

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

$$\pi(x) = [1 - \pi(x)] e^{(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}$$

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

$$\ln \left[\frac{\pi(x)}{1-\pi(x)} \right] = \ln e^{(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}$$

$$\ln \left[\frac{\pi(x)}{1-\pi(x)} \right] = (\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p) \quad (4)$$

Sehingga setelah di logit maka diperoleh persamaan

$$\text{logit}(\pi) = \ln \left[\frac{\pi(x)}{1-\pi(x)} \right] = (\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p) \quad (5)$$

Model tersebut merupakan fungsi linier dari parameter-parameternya. Dalam model regresi linier, diasumsikan bahwa amatan dari variabel respon diekspresikan sebagai $y = E(Y|x) + \varepsilon$ sehingga

$$E(Y|x) = (\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p) = \beta_0 + \sum_{i=1}^p \beta_i x_i \quad (6)$$

Merupakan rata-rata dari populasi dan ε merupakan komponen acak yang menunjukkan penyimpangan amatan dari rata-ratanya dan ε diasumsikan mengikuti sebaran normal dengan rata-rata nol dan varians konstan.

1.5.4 Uji Tiap-Tiap Parameter Model (Uji Wald)

Uji wald digunakan untuk mengetahui signifikan atau tidak tiap-tiap parameter variabel bebas, maka akan dilakukan pengujian hipotesis:

$H_0: \beta_k = 0$ (Variabel X tidak signifikan mempengaruhi variabel Y)

$H_1: \beta_k \neq 0$ (Variabel X signifikan mempengaruhi variabel Y)

Dimana $k = 1, 2, \dots, p$ dan p merupakan jumlah variabel prediktor dalam model. Statistik uji yang digunakan pada pengujian ini dipaparkan dalam formula terdapat pada persamaan berikut:

$$W = \frac{\beta_k}{SE(\beta_k)} \quad (7)$$

β_k merupakan estimasi parameter dan $SE(\beta_k)$ merupakan taksiran standard error. Tolak H_0 pada taraf nyata α bila nilai dari $|W| > Z_{\frac{\alpha}{2}}$ dapat dilihat pada persamaan berikut:

$$W = \frac{\beta_k^2}{SE(\beta_k)^2} \quad (8)$$

Statistika uji tersebut mengikuti distribusi Chi-Squared sehingga H_0 ditolak jika $W > \chi^2_{(v, \alpha)}$ dengan w adalah banyak prediktor.

1.5.5 Uji Kecocokan Model

Uji kecocokan model digunakan untuk mengevaluasi cocok tidaknya model dengan data, nilai observasi yang diperoleh sama atau mendekati dengan yang diharapkan dalam model. Hipotesis yang digunakan dalam uji kecocokan model adalah sebagai berikut.

$H_0 : \hat{\pi}_i = y_i$ (model sesuai atau tidak ada perbedaan nyata antara hasil observasi dengan kemungkinan prediksi model)

$H_1 : \hat{\pi}_i \neq y_i$ (model tidak sesuai atau ada perbedaan nyata antara hasil observasi dengan kemungkinan prediksi model)

Alat yang digunakan untuk menguji kecocokan dalam regresi logistik ordinal adalah uji *deviance*. Statistik *residual deviance* dirumuskan sebagai

$$D = -2 L \quad (9)$$

dimana:

L : nilai log-likelihood untuk model

Statistik *residual deviance* adalah ratio likelihood untuk membandingkan model M dengan model Saturated. Statistik *residual deviance* menyebar menurut distribusi Chisquare dengan derajat bebas sama dengan jumlah observasi (penelitian) dikurangi dengan jumlah parameter yang tercantum dalam model. Semakin besar nilai D akan memperkecil nilai p-value, hal ini menguaktkan dugaan bahwa model yang terbentuk tidak sesuai dengan data yang ada.

1.5.6 Interpretasi Koefisien Parameter

Interpretasi koefisien-koefisien dalam model regresi logistik dilakukan dalam bentuk odds ratio (perbandingan risiko).

Odds didefinisikan sebagai: $\frac{p}{1-p}$

Keterangan:

p : probabilitas sukses (terjadinya peristiwa $y = 1$)

$1-p$: probabilitas gagal (terjadinya peristiwa $y = 0$)

- Odds ratio (perbandingan risiko), ψ adalah perbandingan nilai odds pada dua individu; misalkan individu A dan individu B.
- Odds ratio dituliskan seperti persamaan

$$\psi = \left[\frac{\frac{p(x_A)}{1-p(x_A)}}{\frac{p(x_B)}{1-p(x_B)}} \right] \quad (10)$$

Keterangan:

X_A : karakter individu A

X_B : karakter individu B

Dengan demikian odds rasio adalah rasio (perbandingan) antara peluang kejadian $y = 1$ dengan peluang kejadian untuk $y = 0$. Misalkan pada data variabel respon y yang berisi kejadian sukses dan gagal. Apabila pengamatan y ke- i merupakan sukses dan dilambangkan dengan $y_i = 1$, maka peluang untuk $y_i = 1$ (sukses) adalah π_i . Sehingga odds rasio menentukan rasio antara peluang sukses dengan peluang gagal.

1.5.7 Confusion Matrix

Matrix yang menginformasikan hasil prediksi secara keseluruhan dari nilai akurasi dan untuk melihat kinerja pengklasifikasi, yaitu seberapa sering kasus class X yang benar diklasifikasikan sebagai class X atau kesalahan. Ketika dataset hanya memiliki dua kelas, yaitu *class* positif dan *class* negative (Bramer 2007):

Tabel 1. Confusion matrix

Kelas Sebenarnya	Kelas Hasil Prediksi	
	Positif = 1	Negatif = 0
Positif = 1	TP	FN
Negatif = 0	FP	TN

Keterangan:

- True Positif (TP), yaitu jumlah dokumen dari kelas 1 yang benar dan diklasifikasikan sebagai kelas 1.
- True Negatif (TN), yaitu jumlah dokumen dari kelas 0 yang benar diklasifikasikan sebagai kelas 0.
- False Positif (FP), yaitu jumlah dokumen dari kelas 0 yang salah diklasifikasikan sebagai kelas 1.
- False Negatif (FN) yaitu jumlah dokumen dari kelas 1 yang salah diklasifikasikan sebagai kelas 0.

Dirumuskan sebagai berikut:

a. Accuration

Akurasi merupakan prediksi true atau tingkat dari nilai kebenaran yang dihasilkan dari model algoritma dalam melakukan klasifikasi, yang mencerminkan sejauh mana prediksi model cocok dengan nilai sebenarnya

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (11)$$

b. Precision

Precision adalah ukuran keakuratan suatu model algoritma yang dapat dihitung dengan membandingkan jumlah True Positive dengan total prediksi positif yang dilakukan oleh model

$$precision = \frac{TP}{TP + FP} \quad (12)$$

c. Recall

Recall adalah ukuran kelengkapan suatu model algoritma yang dapat dihitung dengan membandingkan jumlah data yang benar-benar positif dengan total data yang sebenarnya positif

$$recall = \frac{TP}{TP + FN} \quad (13)$$

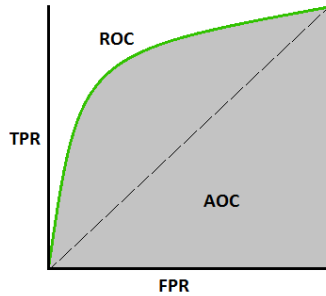
d. F1-Score

F1 score adalah ukuran kinerja model klasifikasi yang memperhitungkan keseimbangan antara precision dan recall. F1 score berguna ketika data tidak seimbang (imbalanced).

$$F1 = 2 \times \frac{precision \times recall}{precision + recall}$$

1.5.8 Kurva AUC-ROC

Kurva AUC-ROC merupakan pengukuran kinerja untuk masalah klasifikasi pada berbagai pengaturan ambang batas. ROC merupakan kurva probabilitas dan AUC merupakan derajat atau ukuran keterpisahan. Kurva ini menunjukkan akurasi dan membandingkan klasifikasi secara visual. Semakin tinggi AUC, semakin baik model tersebut dalam memprediksi kelas 0 dan kelas 1. Kurva ROC mengekspresikan confusion matrix. ROC adalah grafik dua dimensi dengan *false positif rate* (FPR) sebagai garis horizontal dan *true positif rate* (TPR) sebagai garis vertikal. Kurva ROC diplot dengan TPR terhadap FPR dimana TPR berada pada sumbu y dan FPR berada pada sumbu x. Berikut gambaran kurva AUC-ROC:



Gambar 2. Kurva AOC-ROC

Sumber: Narkhede, 2018. <https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5>

Pedoman umum untuk mengklasifikasikan keakuratan pengujian menggunakan AUC menurut (Gorunescu 2011):

0.90 - 1.00 = Excellent Classification;

0.80 - 0.90 = Good Classification;

0.70 - 0.80 = Fair Classification;

0.60 - 0.70 = Poor Classification;

0.50 - 0.60 = Failure.

1.5.9 Akaike Information Criterion (AIC)

Metode AIC adalah metode yang dapat digunakan untuk memilih model regresi terbaik yang ditemukan oleh Akaike (Grasa, 1989). Metode ini didasarkan pada metode *maximum likelihood estimation* (MLE).

Untuk menghitung nilai AIC digunakan rumus sebagai berikut:

$$AIC = e^{\frac{2k}{n} \frac{\sum_{i=1}^n \hat{u}_i^2}{n}} \quad (14)$$

Dengan:

k = Jumlah parameter yang diestimasi dalam model regresi

n = Jumlah observasi

u = Sisa

Persamaan (14) dapat juga ditulis sebagai:

$$\ln AIC = \frac{2k}{n} + \ln \left[\frac{\sum_{i=1}^n \hat{u}_i^2}{n} \right] \quad (15)$$

Menurut metode AIC, model regresi terbaik adalah model regresi yang mempunyai nilai AIC terkecil (Widarjono, 2007).

BAB II

METODOLOGI PENELITIAN

2.1 Sumber Data

Data yang digunakan pada penelitian ini merupakan data primer yang diperoleh dari kuisisioner yang disebarakan kepada anggota organisasi KM FMIPA UNHAS yang masih berstatus mahasiswa aktif. Kuesioner ini disebarakan pada tanggal 13 Juli 2024 hingga 21 Juli 2024 dengan jumlah responden sebesar 121 (seratus dua puluh satu) mahasiswa.

2.2 Pengolahan Data Awal

1. Pengumpulan dan Pembersihan Data

Langkah pertama dalam pengolahan data kuesioner adalah memastikan bahwa semua kuesioner telah dikumpulkan dengan lengkap. Periksa setiap jawaban untuk memastikan tidak ada yang kosong atau tidak sesuai format. Jika ada nilai yang hilang, isi dengan nilai yang sesuai, seperti rata-rata atau modus. Selain itu, pastikan untuk memvalidasi jawaban yang tidak konsisten dan koreksi kesalahan yang mungkin terjadi selama pengisian kuesioner.

2. Pengkodean dan Input Data

Langkah kedua melibatkan pengkodean jawaban menjadi bentuk numerik atau kategori yang dapat dianalisis. Misalnya, ubah jawaban ya/tidak menjadi 1/0 atau kategori lain yang relevan. Jika ada variabel independen kategorikal, lakukan one-hot encoding atau label encoding sesuai kebutuhan. Setelah data dikodekan, input data ke dalam perangkat lunak statistik seperti Python atau Excel. Pastikan untuk memeriksa kesalahan input data dan lakukan koreksi jika diperlukan. Hal ini penting untuk memastikan bahwa data siap untuk analisis lebih lanjut.

3. Analisis Regresi Logistik

Langkah terakhir adalah melakukan analisis regresi logistik. Pertama, standarisasi variabel independen jika diperlukan, kemudian bangun model regresi logistik menggunakan perangkat lunak statistik. Prediksi menggunakan set pengujian dan evaluasi model menggunakan metrik *confusion matrix* dan *classification report* untuk mengukur kinerja model. Langkah ini bertujuan untuk memahami hubungan antara variabel dependen biner dan variabel independen, serta memprediksi probabilitas hasil berdasarkan model yang telah dibangun.

2.3 Eksperimen dan Pengujian Model

Tahapan eksperimen dalam penelitian ini adalah:

1. Pengujian Beberapa Model

Pertama, dilakukan uji coba beberapa model sederhana dan model penuh dalam bentuk model regresi logistik menggunakan. Dalam hal ini, variable yang digunakan adalah setiap variabel independen sebagai model sederhana dengan membentuk model menggunakan Generalized Linear Model lalu dilakukan uji kecocokan model menggunakan uji deviance lalu terakhir dilakukan uji wald.

2. Penyesuaian Model

Jika hasil evaluasi menunjukkan bahwa model perlu ditingkatkan, lakukan penyesuaian model. Ini bisa melibatkan penambahan atau pengurangan variabel independen, transformasi variabel, atau penggunaan teknik regularisasi seperti L1 (*Lasso*) atau L2 (*Ridge*) untuk mencegah *overfitting*.

3. Pengujian Model Akhir

Tahapan terakhir ini akan melakukan evaluasi terhadap data yang telah diuji sebelumnya. Semakin kecil nilai AIC dari model yang diperoleh dari Analisis Regresi Logistik, maka akan semakin baik daya prediksi dari model dugaan yang diperoleh.

2.4 Evaluasi dan Validasi Hasil.

Hitung metrik evaluasi seperti akurasi, precision, recall, dan F1-score. Untuk menunjukkan distribusi prediksi model antara true positives, true negatives, false positives, dan false negatives, maka perlu untuk menghitung confusion matrix. Ini memberikan gambaran mengenai seberapa baik model dalam memprediksi kedua kelas dan sekaligus validasi hasil dari model yang telah dibuat.