E-ISSN : 3163-7329
Applied Engineering: Jurnal Rekayasa dan Sains Terapan
Vol 1 No 2 (2026)
Website: https://journal.emergingmdtrends.com/index.php/applied-science
©2026 GRM Academic Press

Penerapan Logistic Regression untuk Klasifikasi Keaktifan Anggota Pramuka sebagai Dasar Rekomendasi Pembinaan

M. Fathul Hidayat[1], Andy Haryoko[2], Amaludin Arifia[3]

[12]Program Studi Teknik Informatika, Fakultas Teknik, Universitas PGRI Ronggolawe Tuban

*Correspondent email: mbahkulmba@gmail.com

Received: 10 Juni 2026 **Revised**: 14 Juli 2026 **Accepted**: 19 Oktober 2026 **Published**: 19 Oktober 2026

©2026 by the authors. Licensee Applied Engineering: Jurnal Rekayasa dan Sains Terapan
This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution(CC BY NC) licensehttps://creativecommons.org/licenses/by-nc/4.0/.

Abstract. Member participation is an indicator of the success of character development through the Scout Movement. However, its assessment in educational institutions still relies on manual attendance records, which may lack objectivity. This study aims to classify Scout members' participation levels, identify the dominant factors influencing participation, and develop rule-based recommendations for member development. This quantitative study employed total sampling involving 178 Scout members at MA Syi'ar Islam Maibit in the 2025/2026 academic year. Six predictor variables were measured using a Likert-scale questionnaire, while participation labels were determined based on attendance percentages, with a threshold of 75%. The data were stratified into training and testing sets at an 80:20 ratio, standardized, and modeled using Logistic Regression. Model performance was evaluated using a Confusion Matrix, ROC-AUC, and five-fold cross-validation. The model achieved an accuracy of 61.11%, precision of 70.00%, recall of 63.64%, and an ROC-AUC of 0.5942 on the test set. Activity Attractiveness and Institutional Demands were identified as the dominant significant factors. Twelve IF–THEN rules were developed to generate personalized recommendations for member development. The classification performance was moderate, indicating that the system is suitable for use as a decision-support tool for Scout leaders.

Keywords: Scout participation, Logistic Regression, character development, classification, rule-based recommendations.

Abstrak. Keaktifan anggota merupakan indikator keberhasilan pembinaan karakter melalui Gerakan Pramuka, tetapi penilaiannya di satuan pendidikan masih bertumpu pada absensi manual yang kurang objektif. Penelitian ini bertujuan mengklasifikasikan keaktifan anggota Pramuka, mengidentifikasi faktor dominan yang memengaruhinya, dan menyusun rekomendasi pembinaan berbasis aturan. Penelitian kuantitatif ini menggunakan total sampling terhadap 178 anggota Pramuka MA Syi'ar Islam Maibit Tahun Ajaran 2025/2026. Enam variabel prediktor diukur dengan kuesioner skala Likert, sedangkan label keaktifan ditentukan dari persentase kehadiran dengan batas 75%. Data dibagi secara stratifikasi 80:20, distandarisasi, dimodelkan dengan Logistic Regression, lalu dievaluasi menggunakan Confusion Matrix, ROC-AUC, dan validasi silang lima lipatan. Model mencapai akurasi 61,11%, presisi 70,00%, recall 63,64%, dan ROC-AUC 0,5942 pada data uji. Daya Tarik Kegiatan dan Tuntutan Institusional menjadi faktor dominan yang signifikan. Dua belas aturan IF–THEN menghasilkan rekomendasi pembinaan personal. Kinerja klasifikasi tergolong sedang sehingga sistem layak dipakai sebagai pendukung keputusan pembina.

Kata Kunci: Keaktifan Anggota Pramuka, Regresi Logistik, Pendidikan Karakter, Klasifikasi, Rekomendasi Berbasis Aturan.

1. Pendahuluan

Pembentukan karakter generasi muda tidak hanya berlangsung di ruang kelas, tetapi juga melalui organisasi pemuda yang menanamkan kedisiplinan, tanggung jawab, kepemimpinan, kerja sama, dan kepedulian sosial (Karimah dkk., 2024). Di Indonesia, Gerakan Pramuka menjadi salah satu wadah utama untuk tujuan tersebut karena telah lama menjadi bagian dari sistem pendidikan nasional dan berfungsi mengembangkan karakter serta keterampilan nonakademik anggotanya (Rizal dkk., 2025). Peran strategis itu ditegaskan melalui Peraturan Menteri Pendidikan Dasar dan Menengah Nomor 13 Tahun 2025 yang menetapkan kegiatan Pramuka sebagai ekstrakurikuler wajib mulai jenjang sekolah dasar hingga sekolah menengah atas.

Keaktifan anggota menjadi indikator penting untuk menilai keterlibatan sekaligus keberhasilan pembinaan. Keaktifan dalam kegiatan ekstrakurikuler, termasuk kepramukaan, dilaporkan berkaitan positif dengan prestasi dan motivasi belajar siswa (Lutriani dkk., 2022). Meskipun demikian, banyak satuan pendidikan belum memiliki kerangka evaluasi yang komprehensif dan terstandar untuk mengukur keaktifan (Anton dkk., 2025). Kondisi serupa ditemukan di Gugus Depan Pramuka MA Syi'ar Islam Maibit, Tuban, yang masih mengevaluasi kegiatan melalui absensi manual pada setiap pertemuan. Absensi hanya menggambarkan kehadiran fisik, padahal keaktifan juga dipengaruhi partisipasi, motivasi, dukungan keluarga, dan lingkungan, sehingga pengukuran yang tidak komprehensif berpotensi menghasilkan evaluasi yang kurang akurat (Sutika dkk., 2023). Rekapitulasi dengan Microsoft Excel pun hanya menyajikan statistik deskriptif seperti rata-rata dan persentase kehadiran. Pendekatan itu tidak mengklasifikasikan status keaktifan dari kombinasi banyak faktor, tidak mengukur kontribusi relatif setiap faktor dengan mengendalikan faktor lain, dan tidak menghasilkan model yang dapat dipakai ulang untuk anggota baru.

Logistic Regression merupakan metode klasifikasi terawasi yang memodelkan probabilitas kelas biner, dan koefisiennya dapat ditafsirkan sebagai *Odds Ratio* sehingga cocok untuk mengidentifikasi faktor yang berpengaruh (Febrinita dkk., 2024). Metode ini menunjukkan kinerja yang baik pada berbagai kasus, yaitu akurasi 91% pada klasifikasi kepribadian (Pradana dkk., 2025), akurasi 86,67% pada klasifikasi kelulusan siswa (Harvian dkk., 2025), serta akurasi 88,01% dan AUC 84,6% pada seleksi penerima KIP Kuliah (Susetyoko dkk., 2022). Rasyad dan Maharani (2025) memperoleh akurasi 73,49% dan AUC 0,88 pada klasifikasi emosi di media sosial, sedangkan Royyani dkk.

(2023) menunjukkan bahwa minat belajar dan lingkungan keluarga berpengaruh signifikan terhadap kemandirian siswa.

Kajian mengenai keaktifan remaja dalam kegiatan terorganisasi masih menggunakan pendekatan yang berbeda. Meta dkk. (2023) menganalisis faktor keaktifan remaja di posyandu dengan uji *chi-square* tanpa klasifikasi prediktif, sedangkan Spruijtenburg dkk. (2022) menemukan motivasi sebagai faktor paling dominan dalam partisipasi remaja pada olahraga terorganisasi melalui regresi logistik multilevel. Sadouni dan Mehri (2022) membangun sistem rekomendasi berbasis prediksi keterlibatan siswa dengan *Convolutional Neural Network* (akurasi 88%), tetapi model yang kompleks itu kurang menekankan interpretabilitas kontribusi setiap faktor. Dari penelitian yang dikaji, belum teridentifikasi penelitian yang mengklasifikasikan keaktifan anggota Pramuka sekaligus menilai kontribusi relatif faktor-faktornya dan menerjemahkannya menjadi rekomendasi pembinaan.

Berdasarkan celah tersebut, penelitian ini bertujuan (1) mengimplementasikan *Logistic Regression* untuk mengklasifikasikan keaktifan anggota Pramuka ke dalam kategori aktif dan tidak aktif serta mengevaluasi kinerjanya, (2) mengidentifikasi faktor dominan keaktifan melalui koefisien regresi dan *Odds Ratio*, dan (3) merancang sistem rekomendasi pembinaan berbasis *Rule-Based System* dari hasil klasifikasi dan faktor dominan. Kontribusi penelitian ini adalah rangkaian kerja yang menghubungkan klasifikasi statistik yang dapat ditafsirkan dengan rekomendasi pembinaan yang personal bagi pembina.

2. Metode

2.1 Jenis dan Subjek Penelitian

Penelitian ini merupakan penelitian kuantitatif dengan pendekatan eksperimen dan desain *cross-sectional*, yaitu data dikumpulkan pada satu periode pengambilan. Subjek penelitian adalah seluruh anggota Pramuka MA Syi'ar Islam Maibit, Tuban, Tahun Ajaran 2025/2026 yang berjumlah 178 anggota (79 siswa kelas X dan 99 siswa kelas XI). Seluruh anggota populasi dijadikan sampel melalui teknik *total sampling* sehingga tidak ada data yang dieliminasi. Prinsip etika penelitian dipenuhi melalui persetujuan informasi, anonimitas dengan memakai nomor induk tanpa nama, dan kerahasiaan data (Gustari dkk., 2024).

2.2 Variabel dan Instrumen

Variabel bebas terdiri atas enam prediktor, yaitu Daya Tarik Kegiatan (X_1), Tuntutan Institusional (X_2), Manfaat Kegiatan (X_3), Motivasi dan Karakter Individu (X_4), Dukungan Sosial (X_5), dan Manajemen Waktu (X_6). Variabel tersebut dioperasionalkan dari tiga dimensi faktor keaktifan, yaitu daya tarik kegiatan, kewajiban institusional, dan manfaat personal (Yuliana dkk., 2020), dengan tambahan faktor internal dan eksternal anggota. Tuntutan institusional dimaknai sebagai keaktifan yang muncul karena ketentuan formal sekolah (Bimantara & Dirgantara, 2024). Variabel terikat (Y) adalah status keaktifan yang ditetapkan dari data absensi resmi sekolah. Anggota dengan persentase kehadiran sekurang-kurangnya 75% diberi label Aktif (Y = 1), sedangkan yang kurang dari 75% diberi label Tidak Aktif (Y = 0). Penentuan label dari absensi menjadikan variabel terikat objektif dan terpisah dari persepsi responden.

Instrumen berupa kuesioner 15 butir dengan skala Likert 1–5 (1 = sangat tidak setuju hingga 5 = sangat setuju). Jumlah butir untuk X_1 sampai X_6 berturut-turut adalah 3, 2, 2, 2, 4, dan 2, dan nilai setiap variabel adalah rata-rata skor butir penyusunnya. Uji coba dilakukan pada 30 responden dari populasi yang sama. Validitas diuji dengan korelasi *Product Moment* Pearson (butir valid bila r hitung lebih besar daripada r tabel 0,361) dan reliabilitas dengan *Cronbach's Alpha* (reliabel bila lebih besar daripada 0,60) (Ramadhan dkk., 2024).

2.3 Prapemrosesan dan Pembagian Data

Prapemrosesan meliputi konversi data kuesioner ke numerik, validasi rentang skala 1–5 (nilai di luar rentang dianggap hilang), imputasi median untuk nilai hilang, pembentukan variabel X_1–X_6, pembersihan data absensi (kehadiran tidak negatif dan tidak melebihi total pertemuan), pelabelan, serta penggabungan data kuesioner dan absensi berdasarkan nomor induk. Dataset dibagi menjadi data latih (80%) dan data uji (20%) dengan *stratified split* agar proporsi kelas tetap terjaga pada data yang tidak seimbang (Widagdo dkk., 2025). Standarisasi *Z-score* dipelajari hanya dari data latih, lalu diterapkan pada data uji untuk mencegah kebocoran data.

2.4 Pemodelan dan Evaluasi

Model dibangun dengan pustaka *scikit-learn* pada Python di Google Colab, menggunakan *solver* lbfgs, iterasi maksimum 1.000, dan *class weight* seimbang untuk menangani ketidakseimbangan kelas. Probabilitas anggota termasuk kelas Aktif dihitung dengan fungsi sigmoid pada Persamaan (1), dan anggota diklasifikasikan Aktif apabila probabilitasnya sekurang-kurangnya 0,5.

$$P(Y=1|X) = 1 / (1 + \exp[-(\beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5 + \beta_6X_6)]) \quad (1)$$

Kinerja model dievaluasi pada data uji dengan *Confusion Matrix* beserta turunannya, yaitu *accuracy*, *precision*, *recall*, *F1-score*, dan *specificity* (Sriyanti dkk., 2024), ditambah ROC-AUC dan *Average Precision*. Konsistensi model diuji dengan *stratified 5-fold cross-validation* pada seluruh 178 data (Wijiyanto dkk., 2024). Ambang klasifikasi optimal dicari dengan *Youden Index*, yaitu J = TPR − FPR.

2.5 Identifikasi Faktor Dominan dan Sistem Rekomendasi

Koefisien regresi (β) diubah menjadi *Odds Ratio* dengan Persamaan (2). Signifikansi setiap variabel diuji pada taraf 0,05 menggunakan *statsmodels* pada data latih yang telah distandarisasi. Korelasi Pearson bivariat antara X_1–X_6 dan Y dibandingkan dengan arah koefisien multivariat untuk mendeteksi indikasi *suppressor effect*.

$$OR = e^{\beta} \quad (2)$$

Pada tingkat individu, faktor dominan ditetapkan sebagai variabel dengan kontribusi absolut terbesar, yaitu hasil kali koefisien regresi dan nilai variabel terstandarisasi (Persamaan 3). Kombinasi status keaktifan hasil prediksi dan faktor dominan kemudian dipetakan ke aturan IF–THEN dalam *Rule-Based System*, yaitu sistem yang menyatakan pengetahuan pembinaan dalam bentuk aturan kondisi dan tindakan.

$$\text{Faktor dominan}(i) = \text{argmax}_j |\beta_j \times Z_{ij}| \quad (3)$$

3. Hasil dan Pembahasan

3.1 Karakteristik Data dan Kualitas Instrumen

Seluruh 178 anggota menjadi responden dan tidak ada data yang dieliminasi. Pemeriksaan kualitas data tidak menemukan nilai di luar skala Likert maupun nilai hilang, dan seluruh data kuesioner berhasil digabungkan dengan data absensi. Sebanyak 111 anggota (62,4%) berstatus Aktif dan 67 anggota (37,6%) berstatus Tidak Aktif. Rasio 1,66 : 1 menunjukkan ketidakseimbangan kelas tingkat sedang sehingga model memakai *class weight* seimbang. Seluruh 15 butir kuesioner valid (r hitung 0,445–0,746 > r tabel 0,361) dan seluruh variabel reliabel (α = 0,688–0,897), sebagaimana ditunjukkan pada Tabel 1.

Tabel 1. Hasil Uji Instrumen dan Statistik Deskriptif Variabel (N = 178)

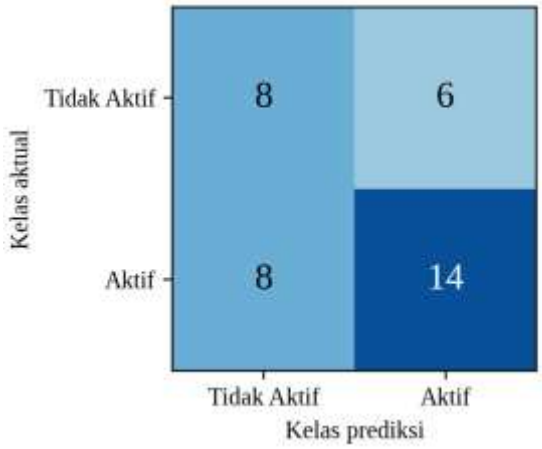
Variabel	Butir	r hitung	α	Rerata	SD
X_1 Daya Tarik Kegiatan	3	0,580–0,637	0,729	3,311	0,684
X_2 Tuntutan Institusional	2	0,627–0,744	0,688	3,640	0,734
X_3 Manfaat Kegiatan	2	0,445–0,465	0,826	3,728	0,733
X_4 Motivasi dan Karakter Individu	2	0,616–0,663	0,800	3,396	0,821
X_5 Dukungan Sosial	4	0,497–0,746	0,897	3,511	0,545
X_6 Manajemen Waktu	2	0,511–0,737	0,802	3,163	0,811

Rerata seluruh variabel berada di atas 3,00 sehingga persepsi responden cenderung sedang hingga tinggi. Manfaat Kegiatan memiliki rerata tertinggi (3,728) dan Manajemen Waktu terendah (3,163). Motivasi dan Karakter Individu memiliki variasi jawaban terbesar (SD = 0,821), sedangkan Dukungan Sosial paling seragam (SD = 0,545).

3.2 Kinerja Model Logistic Regression

3.2.1 Model Logistic Regression

Pembagian stratifikasi menghasilkan 142 data latih (89 aktif dan 53 tidak aktif) serta 36 data uji (22 aktif dan 14 tidak aktif), dengan proporsi kelas yang mendekati proporsi dataset. Pada data uji, model mengklasifikasikan 22 dari 36 anggota secara benar, yaitu 8 anggota tidak aktif (*true negative*) dan 14 anggota aktif (*true positive*), seperti terlihat pada Gambar 1. Kesalahan klasifikasi berjumlah 14 data, terdiri atas 6 anggota tidak aktif yang diprediksi aktif (*false positive*) dan 8 anggota aktif yang diprediksi tidak aktif (*false negative*).



Gambar 1. Confusion Matrix model Logistic Regression pada data uji (n = 36)

Nilai metrik evaluasi dirangkum pada Tabel 2. Akurasi 61,11% menunjukkan bahwa model telah menangkap sebagian pola hubungan antara keenam prediktor dan keaktifan. Presisi 70,00% berarti tujuh dari sepuluh anggota yang diprediksi aktif memang aktif, sedangkan *recall* 63,64% berarti sekitar dua pertiga anggota aktif berhasil dikenali. Model lebih baik pada kelas Aktif (presisi 0,70; *recall* 0,64; F1 0,67) daripada kelas Tidak Aktif (presisi 0,50; *recall* 0,57; F1 0,53), yang sejalan dengan dominasi kelas aktif dalam data. Bagi pembinaan, kesalahan ini bermakna praktis. Enam dari 14 anggota tidak aktif pada data uji diprediksi aktif sehingga berisiko menerima rekomendasi pemeliharaan, bukan pendampingan.

Tabel 2. Kinerja Model pada Data Uji dan Validasi Silang

Metrik	Nilai	Keterangan
Accuracy	0,6111	22 dari 36 data uji diprediksi benar
Precision	0,7000	Ketepatan prediksi kelas Aktif
Recall	0,6364	Anggota aktif yang berhasil dikenali
F1-Score	0,6667	Keseimbangan precision dan recall
Specificity	0,5714	Anggota tidak aktif yang berhasil dikenali
ROC-AUC	0,5942	Kemampuan membedakan kedua kelas
Average Precision	0,7278	Ringkasan kurva precision–recall
AUC validasi silang 5-fold	0,6266 ± 0,0597	Rerata dan simpangan baku lima lipatan

3.2.2 Validasi Silang dan Optimasi Ambang

ROC-AUC sebesar 0,5942 berada di atas 0,5 sehingga model lebih baik daripada tebakan acak, tetapi daya diskriminasinya masih terbatas. *Average Precision* 0,7278 lebih tinggi daripada proporsi kelas aktif pada data uji (22/36 = 0,611), yang menunjukkan model cukup membantu mengenali anggota aktif. Validasi silang lima lipatan menghasilkan AUC 0,6104; 0,5844; 0,6656; 0,7203; dan 0,5524 dengan rerata 0,6266 ± 0,0597. Seluruh lipatan berada di atas 0,5 dan rerata validasi silang mendekati hasil data uji, tetapi sebaran 0,5524 hingga 0,7203 menunjukkan bahwa kinerja model masih sensitif terhadap pembagian data, dengan satu lipatan yang hampir setara tebakan acak. Rendahnya AUC wajar mengingat prediktor berupa persepsi skala Likert dengan dua sampai empat butir per variabel, sebaran skor yang terkonsentrasi di kisaran sedang hingga tinggi, serta ukuran data uji yang kecil (36 data) sehingga nilai AUC mudah berubah oleh satu-dua prediksi.

Optimasi dengan *Youden Index* memberikan ambang 0,4835, yang menaikkan F1-Score dari 0,6667 menjadi 0,7273. Evaluasi utama tetap memakai ambang baku 0,5 agar hasil mengikuti implementasi standar, tetapi temuan ini menunjukkan bahwa penyesuaian ambang dapat diuji pada penerapan nyata.

3.3 Faktor Dominan Keaktifan

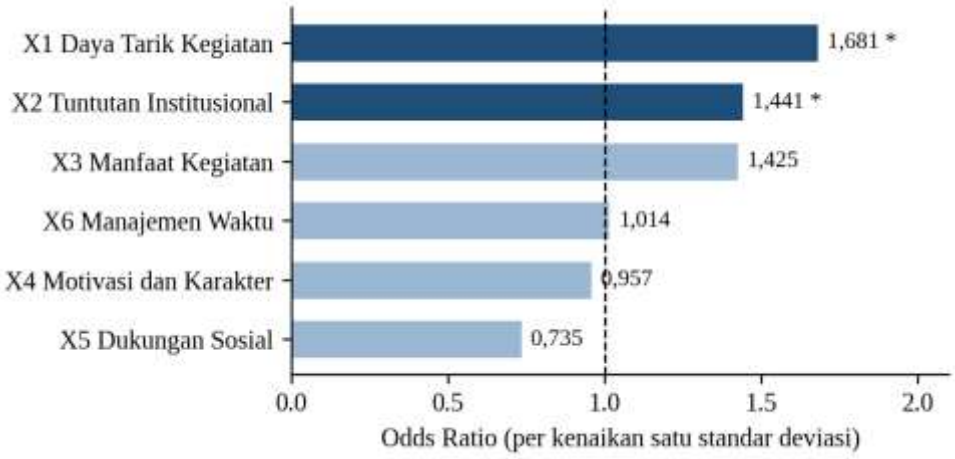
3.3.1 Koefisien Regresi dan Odds Ratio

Karena prediktor distandarisasi, setiap *Odds Ratio* menggambarkan perubahan odds keaktifan untuk kenaikan satu standar deviasi pada variabel terkait, dengan variabel lain tetap. Hasilnya disajikan pada Tabel 3 dan Gambar 2. Hanya dua variabel yang signifikan pada taraf 0,05, yaitu Daya Tarik Kegiatan (β = 0,5197; OR = 1,681; p = 0,035) dan Tuntutan Institusional (β = 0,3655; OR = 1,441; p = 0,047). Kenaikan satu standar deviasi pada Daya Tarik Kegiatan dikaitkan dengan kenaikan odds keaktifan sekitar 68,1%, sehingga variabel ini merupakan faktor paling dominan. Tuntutan Institusional menempati urutan kedua, tetapi nilai p-nya yang berada tepat di bawah 0,05 menandakan bukti yang lebih lemah dan perlu dikonfirmasi pada data yang lebih besar. Manfaat Kegiatan (OR = 1,425; p = 0,155), Motivasi dan Karakter Individu, Dukungan Sosial, dan Manajemen Waktu tidak signifikan.

Tabel 3. Korelasi Bivariat, Koefisien Regresi, Odds Ratio, dan Signifikansi

Variabel	r	β	OR	p	Keterangan

Variabel	r	β	OR	p	Keterangan
X_1 Daya Tarik Kegiatan	0,2238	0,5197	1,681	0,035	Signifikan; arah konsisten
X_2 Tuntutan Institusional	0,2044	0,3655	1,441	0,047	Signifikan; arah konsisten
X_3 Manfaat Kegiatan	0,2578	0,3544	1,425	0,155	Tidak signifikan; konsisten
X_4 Motivasi dan Karakter	0,1209	−0,0435	0,957	0,811	Tidak signifikan; arah berbeda
X_5 Dukungan Sosial	0,1014	−0,3077	0,735	0,200	Tidak signifikan; arah berbeda
X_6 Manajemen Waktu	0,0849	0,0135	1,014	0,879	Tidak signifikan; konsisten



Gambar 2. Odds Ratio setiap variabel prediktor (tanda * menunjukkan p < 0,05)

3.3.2 Korelasi Bivariat dan Indikasi Suppressor Effect

Seluruh korelasi bivariat antara prediktor dan keaktifan bernilai positif tetapi lemah (r = 0,0849–0,2578). Pada Motivasi dan Karakter Individu serta Dukungan Sosial, arah koefisien multivariat berbalik menjadi negatif (β = −0,0435 dan −0,3077). Perubahan arah ini merupakan indikasi *suppressor effect*, yaitu kondisi ketika hubungan suatu variabel dengan variabel target tertutupi oleh tumpang-tindih informasi dengan prediktor lain. Dugaan paling masuk akal adalah tumpang-tindih dengan Daya Tarik Kegiatan dan Manfaat Kegiatan yang secara konseptual berkaitan erat dengan motivasi dan dukungan sosial. Dengan demikian, ketidaksignifikanan Dukungan Sosial ditafsirkan sebagai tidak adanya kontribusi unik yang terukur setelah variabel lain dikendalikan, bukan sebagai bukti bahwa dukungan sosial tidak relevan. Karena uji

multikolinearitas formal (*Variance Inflation Factor*) belum dilakukan, penafsiran ini tetap berstatus indikasi dan bukan pembuktian.

3.4 Sistem Rekomendasi Pembinaan

Sistem rekomendasi memadukan dua informasi untuk setiap anggota, yaitu status keaktifan hasil prediksi dan faktor dominan menurut Persamaan (3). Pola aturannya berbentuk *IF* (status = Tidak Aktif) *AND* (faktor dominan = X_2) *THEN* rekomendasi = pendekatan persuasif dan pendampingan kehadiran. Kombinasi dua status dan enam faktor menghasilkan 12 aturan, yang ringkasannya disajikan pada Tabel 4. Anggota berstatus Aktif memperoleh rekomendasi pemeliharaan, sedangkan anggota berstatus Tidak Aktif memperoleh rekomendasi perbaikan.

Tabel 4. Ringkasan Dua Belas Aturan Rekomendasi Pembinaan

Faktor dominan	Status Aktif (mempertahankan)	Status Tidak Aktif (meningkatkan)
X_1 Daya Tarik Kegiatan	Pertahankan variasi dan inovasi kegiatan, misalnya wide game, pioneering, dan jelajah alam	Hadirkan program yang lebih kreatif, interaktif, dan sesuai kebutuhan anggota
X_2 Tuntutan Institusional	Jaga konsistensi jadwal latihan dan komitmen pada aturan; beri apresiasi atas kehadiran baik	Lakukan pendekatan persuasif, jelaskan aturan secara edukatif, dan dampingi anggota berkehadiran rendah
X_3 Manfaat Kegiatan	Tekankan manfaat nyata melalui pelatihan keterampilan, kepemimpinan, dan kerja sama tim	Tunjukkan manfaat melalui praktik keterampilan, pengalaman lapangan, dan testimoni anggota aktif
X_4 Motivasi dan Karakter	Beri peran sebagai pengurus, pemimpin regu, atau panitia kegiatan	Lakukan pembinaan personal melalui konseling, pendampingan, dan target bertahap
X_5 Dukungan Sosial	Jaga komunikasi dan kolaborasi pembina, teman sebaya, orang tua, dan sekolah	Perkuat dukungan melalui komunikasi intensif, sistem pendampingan sebaya, dan pemantauan berkala
X_6 Manajemen Waktu	Dorong jadwal yang seimbang antara latihan, kegiatan akademik, dan aktivitas pribadi	Bantu menyusun jadwal fleksibel dan realistis agar tidak berbenturan dengan kegiatan akademik

Penerapan pada 36 data uji menghasilkan 20 anggota berstatus Aktif dan 16 anggota berstatus Tidak Aktif. Faktor dominan terbanyak adalah Tuntutan Institusional (14 anggota), diikuti Daya Tarik Kegiatan (10), Dukungan Sosial (8), dan Manfaat Kegiatan (4). Motivasi dan Karakter Individu serta Manajemen Waktu tidak muncul sebagai faktor dominan bagi anggota mana pun karena koefisiennya mendekati nol. Sebagai contoh, anggota bernomor induk 2331 (probabilitas aktif 0,3620) berstatus Tidak Aktif dengan faktor dominan Tuntutan Institusional sehingga menerima saran pendekatan persuasif dan pendampingan, sedangkan anggota 2374 (probabilitas 0,5800) berstatus Aktif dengan faktor dominan Daya Tarik Kegiatan sehingga menerima saran untuk terus menghadirkan kegiatan yang menarik. Seluruh keluaran diekspor ke berkas Excel agar dapat dipakai pembina.

Perlu ditegaskan bahwa isi rekomendasi untuk Dukungan Sosial disusun positif berdasarkan landasan teori keterlibatan sosial meskipun koefisiennya pada tingkat populasi negatif dan tidak signifikan. Sistem ini dengan demikian bersifat *theory-driven* pada isi rekomendasi, sedangkan penentuan anggota penerimanya murni *data-driven* berdasarkan kontribusi individual $|\beta \times Z|$.

3.5 Pembahasan

Akurasi 61,11% dan AUC 0,5942 lebih rendah daripada kinerja *Logistic Regression* pada penelitian terdahulu, yaitu akurasi 86,67%–91% pada klasifikasi kelulusan, kepribadian, dan seleksi penerima beasiswa (Harvian dkk., 2025; Pradana dkk., 2025; Susetyoko dkk., 2022), serta AUC 0,88 pada klasifikasi emosi (Rasyad & Maharani, 2025). Perbedaan itu dapat dipahami karena penelitian terdahulu memakai data yang lebih besar atau prediktor yang lebih dekat dengan label, sedangkan penelitian ini memprediksi label berbasis absensi dari persepsi anggota pada satu sekolah dengan 178 data. Kehadiran juga dapat dipengaruhi hal di luar sikap, seperti kesehatan dan bentrok jadwal, yang tidak tertangkap kuesioner. Karena itu model ini lebih tepat dipandang sebagai alat bantu penyaringan dan penentuan prioritas pembinaan, bukan penentu tunggal status keaktifan. Kelebihannya terletak pada interpretabilitas yang tidak dimiliki model kompleks seperti *Convolutional Neural Network* pada sistem rekomendasi keterlibatan siswa (Sadouni & Mehri, 2022).

Dominasi Daya Tarik Kegiatan menguatkan pandangan bahwa desain kegiatan yang menarik, bervariasi, dan dilaksanakan di alam terbuka mendorong keaktifan anggota (Yuliana dkk., 2020). Temuan ini berbeda dengan Spruijtenburg dkk. (2022)

yang menempatkan motivasi sebagai faktor paling dominan pada olahraga terorganisasi. Motivasi tidak signifikan dalam penelitian ini, kemungkinan karena konteks yang berbeda serta tumpang-tindih dengan Daya Tarik Kegiatan, sebagaimana ditunjukkan oleh indikasi *suppressor effect*. Signifikansi Tuntutan Institusional sesuai dengan kondisi sekolah yang menjadikan Pramuka bagian pembinaan karakter. Namun, karena butir pengukurnya berkaitan dengan kewajiban mengikuti kegiatan, hubungannya dengan kehadiran perlu dibaca hati-hati dan tidak ditafsirkan sebagai hubungan sebab-akibat.

Penelitian ini memiliki beberapa keterbatasan. Data berasal dari satu sekolah sehingga generalisasi terbatas, dan data uji hanya 36 anggota. Uji multikolinearitas belum dilakukan. Nilai p diperoleh dari estimasi *statsmodels* yang terpisah dari model *scikit-learn* yang memakai bobot kelas sehingga keduanya tidak sepenuhnya identik. Label keaktifan hanya didasarkan pada kehadiran selama satu semester. Kesesuaian dan efektivitas rekomendasi juga belum divalidasi oleh pembina maupun diuji pada program pembinaan yang sesungguhnya.

4. Kesimpulan

Logistic Regression dapat diterapkan untuk mengklasifikasikan keaktifan anggota Pramuka dari enam persepsi anggota, tetapi kemampuannya membedakan anggota aktif dan tidak aktif baru berada pada tingkat sedang. Model karena itu layak menjadi pendukung keputusan pembina dan belum layak menjadi penentu tunggal status keaktifan. Daya Tarik Kegiatan dan Tuntutan Institusional merupakan faktor yang signifikan memengaruhi keaktifan, sehingga pembinaan perlu memprioritaskan kualitas dan variasi kegiatan serta konsistensi aturan dan pendampingan kehadiran.

Rule-Based System dengan 12 aturan IF–THEN mengubah keluaran statistik menjadi saran pembinaan yang personal, sehingga gugus depan memperoleh dasar pembinaan berbasis data tanpa perangkat yang rumit. Penelitian selanjutnya perlu memperluas objek ke beberapa gugus depan, melakukan uji multikolinearitas, membandingkan algoritma lain seperti *Random Forest* dan *XGBoost*, menambah prediktor seperti prestasi dan partisipasi nonrutin, serta memvalidasi rekomendasi bersama pembina sebelum dikembangkan menjadi aplikasi web atau seluler.

Referensi

Anton, Nurazizah, A., Fitriana, I., & Maudi. (2025). Sistem evaluasi penilaian kegiatan ekstrakurikuler PAI di sekolah tingkat pendidikan menengah. *JIIC: Jurnal Intelek Insan Cendekia, 2*, 336–342.

Bimantara, G. A., & Dirgantara, G. A. (2024). Ekstrakurikuler Pramuka tidak wajib diikuti peserta didik: Perspektif filsafat pendidikan humanisme. *Jurnal Kiprah Pendidikan, 3*, 73–78. https://doi.org/10.33578/kpd.v3i2.222

Febrinita, F., Zaman, W. I., & Puspitasari, W. D. (2024). Faktor-faktor yang mempengaruhi hasil belajar statistika mahasiswa melalui pemodelan regresi logistik biner. *Jurnal Riset HOTS Pendidikan Matematika, 4*, 523–534.

Gustari, L. A., Riswanto, & Kholis, N. (2024). Prinsip dasar dan etika dalam penelitian ilmiah. *Pendas: Jurnal Ilmiah Pendidikan Dasar, 9*(4), 486–494.

Harvian, Handayani, R., Awu, M. Y., & Apdhal, M. (2025). Klasifikasi kelulusan siswa menggunakan metode Logistic Regression. *Jurnal Cakrawala dan Integrasi Pendidikan, 1*(1), 18–23.

Karimah, L. N., Sunaisah, Alfina, S. V. M., Fitriana, S. N., Ardianti, S. D., & Ismaya, E. A. (2024). Korelasi keaktifan siswa dalam kegiatan ekstrakurikuler terhadap prestasi belajar siswa kelas V SD 2 Karangbener. *Protasis: Jurnal Bahasa, Sastra, Budaya, dan Pengajarannya, 3*(1), 93–100. https://doi.org/10.55606/protasis.v3i1.143

Lutriani, Lamada, M., & Massikki. (2022). Pengaruh keaktifan siswa dalam ekstrakurikuler terhadap prestasi belajar siswa kelas XI RPL SMKN 2 Wajo. *Information Technology Education Journal, 1*(1), 1–9. https://doi.org/10.59562/intec.v1i1.151

Meta, A., Abdul Majid, Y., & Ardianti, S. (2023). Faktor-faktor yang berhubungan dengan keaktifan remaja mengikuti posyandu. *Jurnal Kesehatan Jurnal Ilmiah Multi Sciences, XIII*(2), 70–79.

Pradana, M. L. A., Budiman, F., & Kurniawan, D. (2025). Pendekatan machine learning untuk klasifikasi kepribadian: Studi Logistic Regression dan Naive Bayes. *Edumatic: Jurnal Pendidikan Informatika, 9*(3), 689–697. https://doi.org/10.29408/edumatic.v9i3.32211

Ramadhan, M. F., Siroj, R. A., & Afgani, M. W. (2024). Validitas dan reliabilitas. *Journal on Education, 6*(2), 10967–10975. https://doi.org/10.31004/joe.v6i2.4885

Rasyad, G. S., & Maharani, W. (2025). Logistic Regression and Naïve Bayes comparison in classifying emotions on Indonesian X social media. *Edumatic: Jurnal Pendidikan Informatika, 9*(1), 31–40. https://doi.org/10.29408/edumatic.v9i1.29120

Rizal, S., Putra, D. A., Syafitri, R. A., Rahayu, N. C., & Zahri, Z. I. (2025). Analisis pengaruh program ekstrakurikuler terhadap pembentukan karakter siswa di SMA Negeri 2 Kota Bengkulu. *JKIP: Jurnal Kajian Ilmu Pendidikan, 5*(4), 1009–1014.

Royyani, A., Harini, S., Susanti, E., Sabara, I. M., Mulyani, E. S., & Salsabila, A. (2023). Pengujian model regresi logistik pengaruh minat belajar dan lingkungan keluarga terhadap kemandirian siswa. *JSMS: Jurnal Sains Matematika dan Statistika, 9*(2), 89–96. https://doi.org/10.24014/jsms.v9i2.23304

Sadouni, O., & Mehri, A. (2022). New recommendation system based on students' engagement prediction using CNN to optimize e-learning. *International Journal of Organizational and Collective Intelligence, 12*(4), 1–27. https://doi.org/10.4018/ijoci.312225

Spruijtenburg, G. E., van Abswoude, F., Platvoet, S., de Niet, M., Bekhuis, H., & Steenbergen, B. (2022). Factors related to adolescents' participation in organized sports. *International Journal of Environmental Research and Public Health, 19*(23). https://doi.org/10.3390/ijerph192315872

Sriyanti, Z. A., Kartika, D. S. Y., & Najaf, A. R. E. (2024). Implementasi model BERT pada analisis sentimen pengguna Twitter terhadap aksi boikot produk Israel. *JITET (Jurnal Informatika dan Teknik Elektro Terapan), 12*(3), 2335–2342. https://doi.org/10.23960/jitet.v12i3.4743

Susetyoko, R., Yuwono, W., & Purwantini, E. (2022). Model klasifikasi pada seleksi mahasiswa baru penerima KIP Kuliah menggunakan regresi logistik biner. *JIP (Jurnal Informatika Polinema), 8*, 31–40.

Sutika, R., Gunawan, H., & Hodsay, Z. (2023). Pengaruh keaktifan siswa berorganisasi terhadap peningkatan soft skill dan hasil belajar kewirausahaan di SMA Negeri 1 Pedamaran. *Jurnal Pendidikan Tambusai, 7*(2), 17047–17055. https://doi.org/10.31004/jptam.v7i2.9071

Widagdo, B. W., Fadillah, M. R., Adiguna, M. A., Wiharjo, S., & Handayani, M. (2025). Implementation of split sampling for Decision Tree and K-Nearest Neighbor algorithms in the DKI Jakarta legislative election. *Jurnal Penelitian Inovatif, 5*(4), 2979–2986. https://doi.org/10.54082/jupin.1704

Wijiyanto, Pradana, A. I., Sopingi, S., & Atina, V. (2024). Teknik K-fold cross validation untuk mengevaluasi kinerja mahasiswa. *Jurnal Algoritma, 21*(1), 239–248. https://doi.org/10.33364/algoritma/v.21-1.1618

Yuliana, Putra, M. J. A., & Antosa, Z. (2020). Faktor-faktor yang mempengaruhi motivasi siswa sekolah dasar dalam mengikuti aktivitas Pramuka Penggalang. *Tunjuk Ajar: Jurnal Penelitian Ilmu Pendidikan, 3*, 210–226. https