Penerapan Metode *Logistic Regression* Untuk Kelayakan Perahu Karet Pada Aktivitas Arung Jeram Pekalen Probolinggo

M. Zainul Robbasi[1], Imam Marzuki[2*], Utami Ratna Swari[3]
[1,2,3] Teknik Elektro, Fakultas Teknik dan Informatika, Universitas Panca Marga
Jalan Yos Sudarso No 107 Pabean, Dringu, Probolinggo 67271
*Penulis Korespondensi :
(Imam Marzuki, Teknik Elektro, Fakultas Teknik dan Informatika, Universitas Panca Marga imam@upm.ac.id)

ABSTRAK

Kelayakan perahu karet merupakan salah satu faktor penting keselamatan pada wisata olahraga (*sport tourism*) arung jeram (*white water rafting*) di Sungai Pekalen, Kabupaten Probolinggo. Penelitian ini bertujuan menerapkan metode *Logistic Regression* untuk mengklasifikasikan kelayakan perahu karet ke dalam kategori Layak dan Tidak Layak serta mengidentifikasi variabel yang berkaitan dengan status kelayakan tersebut. Dataset berupa data primer hasil observasi lapangan sebanyak 200 record dengan tujuh variabel prediktor, yaitu material, tekanan udara, kapasitas, kondisi perahu, *safety equipment*, tingkat kesulitan, dan karakteristik sungai, dengan distribusi kelas yang tidak seimbang (93% Tidak Layak). Data dibagi menggunakan *stratified random sampling* 70:30 yang diulang 10 kali dan divalidasi dengan validasi silang 5-*fold*. Kinerja model dievaluasi menggunakan accuracy, precision, recall, F1-score, dan AUC, serta dibandingkan dengan *Naïve Bayes* dan *Random Forest*. Hasilnya, *Logistic Regression* memperoleh akurasi 0,968, precision 0,969, recall 0,968, F1-score 0,964, dan AUC 0,994 (rata-rata 10 kali pengulangan), dengan AUC tertinggi di antara ketiga model. Berdasarkan koefisien model, *safety equipment* (odds ratio 7,912), tekanan udara (7,412), dan kondisi perahu (2,793) memiliki koefisien positif terbesar terhadap peluang perahu dinyatakan layak. Kontribusi penelitian ini adalah model klasifikasi yang dapat menjadi calon komponen sistem pendukung keputusan evaluasi kelayakan perahu karet; perancangan dan pengujian sistem merupakan tahap penelitian lanjutan.

Kata kunci: *Logistic Regression*, kelayakan perahu karet, arung jeram, keselamatan wisata, Sungai Pekalen.

ABSTRACT

Inflatable boat feasibility is a key safety factor in white-water rafting sport tourism on the Pekalen River, Probolinggo Regency. This study applies the Logistic Regression method to classify inflatable boat feasibility into Feasible and Not Feasible categories and to identify the variables associated with feasibility status. The dataset consists of 200 primary records from field observations with seven predictors: material, air pressure, capacity, boat condition, safety equipment, difficulty level, and river characteristics, with imbalanced classes (93% Not Feasible). The data were split using 70:30 stratified random sampling repeated 10 times and validated using 5-fold cross-validation. Model performance was evaluated using accuracy, precision, recall, F1-score, and AUC, and compared with Naïve Bayes and Random Forest. Logistic Regression achieved an accuracy of 0.968, precision of 0.969, recall of 0.968, F1-score of 0.964, and AUC of 0.994 (mean of 10 repetitions), with the highest AUC among the three models. Based on the model coefficients, safety equipment (odds ratio 7.912), air pressure (7.412), and boat condition (2.793) had the largest positive coefficients for the probability of a boat being classified as feasible. The contribution of this study is a classification model that can serve as a candidate component of a decision support system for inflatable boat feasibility evaluation; system design and testing are left for future work.

Keywords: *Logistic Regression, inflatable boat feasibility, white-water rafting, tourism safety, Pekalen River.*

1. PENDAHULUAN

Pariwisata merupakan salah satu sektor strategis yang memiliki kontribusi besar terhadap pertumbuhan ekonomi nasional, peningkatan pendapatan masyarakat, serta pengembangan potensi daerah berbasis sumber daya alam dan budaya. Dalam perkembangannya, sektor pariwisata tidak hanya berorientasi pada wisata konvensional, tetapi juga berkembang menuju wisata berbasis pengalaman dan tantangan, salah satunya adalah *sport tourism* atau wisata olahraga [1]. Wisata olahraga menjadi salah satu sektor yang mengalami pertumbuhan cukup pesat karena mampu mengombinasikan aktivitas rekreasi, olahraga, hiburan, dan petualangan dalam satu kegiatan wisata [1]. Perkembangan tersebut sejalan dengan meningkatnya minat wisatawan terhadap aktivitas wisata berbasis alam yang memberikan pengalaman berbeda serta melibatkan interaksi langsung dengan lingkungan.

Salah satu bentuk wisata olahraga yang berkembang pesat di Indonesia adalah wisata minat khusus (*special interest tourism*) berbasis petualangan atau *adventure tourism*. Wisata minat khusus merupakan aktivitas perjalanan wisata yang dilakukan dengan motivasi tertentu dan menekankan unsur pengalaman, tantangan, rekreasi, serta pencapaian pribadi [1]. Dalam konteks tersebut, arung jeram (*rafting*) menjadi salah satu aktivitas

wisata petualangan yang cukup diminati karena memadukan unsur olahraga, rekreasi, kerja sama tim, dan tantangan alam [2]. Arung jeram merupakan aktivitas menyusuri sungai berarus deras menggunakan perahu karet dengan bantuan dayung dan perlengkapan keselamatan tertentu [2]. Aktivitas ini tidak hanya berfungsi sebagai sarana rekreasi, tetapi juga berkembang menjadi olahraga prestasi yang dipertandingkan dalam berbagai kompetisi tingkat daerah, nasional, maupun internasional [3].

Indonesia memiliki potensi besar dalam pengembangan wisata arung jeram karena didukung oleh kondisi geografis berupa banyaknya sungai berarus deras yang tersebar di berbagai wilayah [2]. Potensi tersebut menjadikan wisata arung jeram sebagai salah satu wisata unggulan berbasis alam yang dapat dikembangkan secara berkelanjutan. Selain memberikan dampak positif terhadap sektor pariwisata, aktivitas arung jeram juga berkontribusi terhadap peningkatan ekonomi masyarakat melalui sektor jasa wisata, transportasi, penginapan, kuliner, serta penyediaan lapangan pekerjaan sebagai *river guide* atau pemandu arung jeram [1]. Salah satu destinasi wisata arung jeram yang cukup terkenal di Jawa Timur adalah Sungai Pekalen yang berada di Kabupaten Probolinggo. Sungai Pekalen memiliki karakteristik arus deras, jeram yang menantang, serta panorama alam yang menarik sehingga menjadi salah satu destinasi favorit wisatawan domestik maupun mancanegara [4].

Meskipun memiliki potensi wisata yang besar, arung jeram termasuk dalam kategori olahraga ekstrem (*extreme sport*) yang memiliki tingkat risiko kecelakaan tinggi [1]. Risiko yang dapat terjadi dalam aktivitas arung jeram antara lain benturan dengan batu sungai, perahu terbalik, wisatawan terjatuh ke sungai, cedera fisik, hingga risiko tenggelam akibat derasnya arus sungai [4]. Faktor lingkungan seperti curah hujan, debit air, kondisi jeram, cuaca, serta kemampuan teknis peserta dan pemandu menjadi variabel yang memengaruhi tingkat keselamatan dalam aktivitas *rafting* [4]. Oleh sebab itu, aspek keselamatan dan keamanan wisatawan menjadi faktor fundamental dalam penyelenggaraan wisata arung jeram.

Pentingnya keselamatan wisatawan dalam aktivitas arung jeram telah diatur melalui berbagai regulasi dan standar nasional, salah satunya melalui Undang-Undang Republik Indonesia Nomor 10 Tahun 2009 tentang Kepariwisataan, Standar Kompetensi Kerja Nasional Indonesia (SKKNI) Arung Jeram, serta Peraturan Menteri Pariwisata dan Ekonomi Kreatif Republik Indonesia Nomor 13 Tahun 2014 tentang Standar Usaha Wisata Arung Jeram [5], [6], [7]. Regulasi tersebut mengatur berbagai aspek penting seperti kompetensi sumber daya manusia, penerapan standar operasional prosedur (SOP), kelengkapan peralatan keselamatan, prosedur evakuasi, pelayanan wisatawan, serta pengelolaan risiko selama aktivitas *rafting* berlangsung [6], [7]. Selain itu, dokumen SOP dan standar keselamatan wisata arung jeram juga menegaskan pentingnya pemeriksaan kondisi peralatan sebelum pengarungan dilakukan guna meminimalkan risiko kecelakaan [8].

Penelitian mengenai keamanan dan keselamatan wisatawan pada Songa Adventure *Rafting* Kabupaten Probolinggo menunjukkan bahwa masih ditemukan beberapa kelemahan dalam implementasi standar keselamatan, baik dari aspek sumber daya manusia, kelengkapan peralatan, maupun faktor lingkungan [4]. Selain itu, penerapan SOP wisata arung jeram yang mengacu pada Peraturan Menteri Pariwisata dan Ekonomi Kreatif Nomor 13 Tahun 2014 belum sepenuhnya berjalan optimal [4]. Kondisi tersebut menunjukkan bahwa pengelolaan keselamatan wisata arung jeram masih memerlukan evaluasi dan pengembangan sistem manajemen risiko yang lebih baik.

Dalam kegiatan arung jeram, salah satu faktor penting yang sangat menentukan tingkat keselamatan wisatawan adalah kelayakan perahu karet sebagai sarana utama aktivitas *rafting*. Perahu karet berfungsi sebagai media utama transportasi dan pengendali keseimbangan selama pengarungan sungai. Kondisi perahu yang mengalami kerusakan, penurunan kualitas material, kebocoran, maupun ketidaksesuaian kapasitas beban dapat meningkatkan risiko kecelakaan bagi wisatawan maupun operator arung jeram [8]. Selain itu, faktor usia penggunaan, intensitas operasional, paparan sinar matahari, benturan dengan batu sungai, serta pola penyimpanan juga dapat memengaruhi kualitas dan daya tahan perahu karet [9].

Pentingnya perawatan dan pemeliharaan peralatan keselamatan dijelaskan dalam penelitian mengenai perawatan alat keselamatan *liferaft*. Penelitian tersebut menunjukkan bahwa pemeriksaan dan pemeliharaan berkala sangat diperlukan untuk menjaga kualitas alat keselamatan serta meminimalkan risiko kecelakaan akibat kegagalan fungsi peralatan [9]. Prinsip tersebut relevan diterapkan pada kegiatan arung jeram, khususnya terhadap peralatan seperti helm, pelampung, dayung, dan perahu karet yang digunakan secara terus-menerus dalam aktivitas pengarungan sungai.

Selain aspek keselamatan, pengelolaan wisata arung jeram juga memerlukan standar usaha dan manajemen operasional yang baik. Penelitian mengenai Analisis Standar Usaha Wisata Arung Jeram Progo *Rafting* menunjukkan bahwa sebagian besar standar usaha wisata telah terpenuhi, terutama pada aspek keamanan, fasilitas, dan pelayanan pelanggan. Namun demikian, masih diperlukan peningkatan pada aspek administrasi, dokumen operasional, serta pelatihan sumber daya manusia agar sesuai dengan standar nasional pariwisata [1]. Hal tersebut menunjukkan bahwa pengelolaan wisata arung jeram memerlukan sistem evaluasi yang terstruktur dan berkelanjutan.

Dalam konteks olahraga prestasi, evaluasi pembinaan arung jeram di FAJI Banyumas menunjukkan bahwa keberhasilan pembinaan olahraga arung jeram dipengaruhi oleh dukungan sarana prasarana, kualitas pelatih, penerapan ilmu pengetahuan dan teknologi olahraga, serta manajemen organisasi yang profesional [3]. Penelitian tersebut juga menjelaskan bahwa kualitas peralatan menjadi salah satu faktor penting dalam mendukung keamanan dan performa olahraga arung jeram [3]. Oleh sebab itu, evaluasi terhadap kelayakan sarana dan prasarana, termasuk perahu karet, menjadi hal penting baik dalam konteks wisata maupun olahraga prestasi.

Perkembangan teknologi informasi dan ilmu data saat ini memberikan peluang besar dalam mendukung sistem pengambilan keputusan berbasis data pada berbagai bidang, termasuk sektor olahraga dan pariwisata. Salah satu pendekatan yang banyak digunakan adalah *data mining*. *Data mining* merupakan proses penggalian informasi dan pola tertentu dari sekumpulan data menggunakan metode statistik dan algoritma klasifikasi [10]. Penerapan metode klasifikasi dalam *data mining* telah banyak digunakan untuk membantu proses prediksi dan pengambilan keputusan pada berbagai bidang, salah satunya menggunakan algoritma *Naïve Bayes* dalam klasifikasi kualitas air [10]. Pendekatan berbasis data tersebut menunjukkan bahwa metode klasifikasi mampu memberikan hasil analisis yang lebih objektif dan sistematis dibandingkan metode konvensional.

Salah satu metode klasifikasi statistik yang banyak digunakan dalam proses prediksi adalah *Logistic Regression*. Metode *Logistic Regression* merupakan teknik analisis statistik yang digunakan untuk memprediksi probabilitas suatu kejadian kategorik berdasarkan hubungan antara variabel independen dan variabel dependen [10]. Metode ini memiliki kemampuan dalam menganalisis pengaruh beberapa variabel terhadap suatu kondisi tertentu sehingga banyak digunakan dalam penelitian berbasis klasifikasi dan prediksi. Dalam konteks penelitian ini, *Logistic Regression* digunakan untuk menganalisis tingkat kelayakan perahu karet berdasarkan beberapa variabel seperti kondisi fisik perahu, usia penggunaan, frekuensi perawatan, intensitas penggunaan, dan kualitas perlengkapan keselamatan.

Meskipun penelitian mengenai keselamatan wisata *rafting*, standar usaha wisata arung jeram, dan evaluasi pembinaan olahraga arung jeram telah banyak dilakukan, penelitian yang secara khusus membahas penerapan metode *Logistic Regression* untuk analisis kelayakan perahu karet pada aktivitas arung jeram masih sangat terbatas. Sebagian besar penelitian sebelumnya lebih berfokus pada aspek keselamatan wisatawan, implementasi SOP, kualitas pelayanan wisata, dan evaluasi organisasi olahraga [1], [3], [4]. Sementara itu, penelitian mengenai metode klasifikasi berbasis *data mining* lebih banyak diterapkan pada bidang kesehatan, lingkungan, dan sistem klasifikasi umum [10]. Dengan demikian, masih terdapat *research gap* berupa belum adanya penelitian yang mengintegrasikan analisis keselamatan wisata arung jeram dengan metode klasifikasi berbasis *Logistic Regression* untuk menentukan kelayakan perahu karet.

Berdasarkan kondisi tersebut, penelitian ini memiliki *novelty* berupa penerapan metode *Logistic Regression* untuk menganalisis kelayakan perahu karet pada aktivitas arung jeram di Sungai Pekalen Probolinggo. Penelitian ini mengintegrasikan aspek keselamatan wisata dengan pendekatan *data mining* untuk membangun model klasifikasi tingkat kelayakan perahu karet secara objektif dan berbasis data. Model tersebut diharapkan dapat menjadi calon komponen sistem pendukung keputusan yang membantu operator *rafting* dalam mengevaluasi dan memelihara perahu karet secara lebih sistematis. Lingkup penelitian ini dibatasi pada pembangunan dan evaluasi model klasifikasi; perancangan sistem, implementasi, dan pengujian pengguna merupakan tahap penelitian lanjutan.

Urgensi penelitian ini semakin penting mengingat tingginya risiko kecelakaan dalam aktivitas arung jeram dan pentingnya penerapan standar keselamatan berbasis data pada sektor *sport tourism*. Evaluasi kelayakan perahu karet secara sistematis diharapkan dapat menjadi langkah preventif dalam meminimalkan risiko kecelakaan wisatawan, meningkatkan kualitas pelayanan wisata arung jeram, serta mendukung pengembangan sistem keselamatan wisata yang profesional dan berkelanjutan. Oleh karena itu, tujuan penelitian ini adalah menerapkan metode *Logistic Regression* untuk menganalisis kelayakan perahu karet pada aktivitas arung jeram di Sungai Pekalen Probolinggo berdasarkan variabel-variabel yang memengaruhi kondisi dan keamanan perahu.

2. METODOLOGI PENELITIAN

Metodologi penelitian merupakan serangkaian langkah sistematis yang digunakan untuk memperoleh data, mengolah informasi, melakukan analisis, serta menghasilkan kesimpulan penelitian secara ilmiah, objektif, dan dapat dipertanggungjawabkan. Metodologi penelitian memiliki peran yang sangat penting dalam suatu penelitian karena menjadi pedoman utama dalam pelaksanaan penelitian mulai dari tahap perencanaan hingga tahap akhir penelitian. Melalui metodologi penelitian yang tersusun dengan baik, proses penelitian dapat dilakukan secara terarah, sistematis, dan sesuai dengan tujuan penelitian yang telah ditetapkan. Selain itu, metodologi penelitian juga berfungsi untuk memastikan bahwa hasil penelitian memiliki tingkat validitas, reliabilitas, dan akurasi yang baik sehingga dapat digunakan sebagai dasar dalam pengambilan keputusan maupun pengembangan penelitian selanjutnya.

Dalam penelitian ini, metodologi penelitian disusun untuk menganalisis tingkat kelayakan perahu karet arung jeram menggunakan metode *Logistic Regression* berdasarkan variabel kualitas produk. *Logistic Regression* dipilih karena merupakan salah satu metode klasifikasi yang efektif dalam memprediksi probabilitas suatu kejadian berdasarkan variabel independen tertentu. Metode ini banyak digunakan dalam penelitian *machine learning* modern karena memiliki kemampuan klasifikasi yang baik, mudah diinterpretasikan, serta mampu menghasilkan model probabilitas yang dapat digunakan dalam proses pengambilan keputusan Sperandei (2014). Selain itu, *Logistic Regression* juga sesuai digunakan pada penelitian yang memiliki variabel dependen berbentuk kategorik, seperti klasifikasi layak dan tidak layak pada penelitian ini.

Penelitian ini menggunakan pendekatan kuantitatif dengan metode analisis klasifikasi berbasis *machine learning*. Pendekatan kuantitatif dipilih karena penelitian berfokus pada pengolahan data numerik, pengukuran

variabel penelitian, serta pengujian hubungan antarvariabel secara statistik. Pendekatan kuantitatif memungkinkan peneliti untuk memperoleh hasil analisis yang lebih objektif karena proses pengambilan keputusan dilakukan berdasarkan hasil pengolahan data dan perhitungan statistik. Menurut Creswell dan Creswell (2023), penelitian kuantitatif merupakan metode penelitian yang digunakan untuk menguji teori tertentu melalui pengukuran variabel dan analisis data numerik menggunakan prosedur statistik.

Metode *machine learning* digunakan dalam penelitian ini karena mampu membantu proses klasifikasi data secara otomatis berdasarkan pola yang diperoleh dari dataset penelitian. Dalam konteks penelitian ini, *machine learning* digunakan untuk membantu proses evaluasi kualitas perahu karet berdasarkan variabel kualitas produk seperti jenis material, ketebalan bahan, kualitas sambungan, tekanan udara, dan ketahanan produk. Dengan menggunakan pendekatan *machine learning*, proses klasifikasi kelayakan perahu karet dapat dilakukan secara lebih cepat, sistematis, dan objektif dibandingkan metode evaluasi manual yang selama ini banyak digunakan dalam proses *quality control*.

Proses penelitian dilakukan secara bertahap mulai dari identifikasi masalah, studi literatur, penentuan variabel penelitian, pengumpulan data, preprocessing data, penerapan metode *Logistic Regression*, evaluasi model, hingga penarikan kesimpulan akhir penelitian. Setiap tahapan penelitian saling berkaitan dan disusun secara sistematis agar proses penelitian berjalan sesuai dengan tujuan yang ingin dicapai. Tahapan penelitian yang terstruktur juga bertujuan untuk meminimalkan kesalahan dalam proses pengolahan data dan meningkatkan kualitas hasil analisis yang dihasilkan.

Pada tahap awal penelitian, dilakukan identifikasi masalah untuk mengetahui permasalahan utama yang berkaitan dengan kualitas dan kelayakan perahu karet arung jeram. Selanjutnya, dilakukan studi literatur untuk memperoleh dasar teori dan referensi penelitian yang relevan. Setelah itu, dilakukan penentuan variabel penelitian yang terdiri atas variabel independen dan variabel dependen. Tahapan berikutnya adalah pengumpulan data melalui observasi, dokumentasi, dan pengujian produk. Data yang diperoleh kemudian diproses melalui tahap preprocessing sebelum dianalisis menggunakan metode *Logistic Regression*.

Hasil analisis model kemudian dievaluasi menggunakan beberapa parameter evaluasi seperti accuracy, precision, recall, dan F1-score untuk mengetahui tingkat performa model klasifikasi yang dihasilkan. Tahap terakhir penelitian adalah penarikan kesimpulan berdasarkan hasil analisis yang telah dilakukan serta penyusunan saran untuk pengembangan penelitian selanjutnya. Dengan tahapan penelitian yang sistematis tersebut, penelitian ini diharapkan mampu menghasilkan model klasifikasi kelayakan perahu karet yang memiliki tingkat akurasi yang baik dan dapat menjadi calon komponen sistem pendukung keputusan dalam evaluasi kelayakan perahu karet arung jeram.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Penelitian

Dataset yang digunakan dalam penelitian ini merupakan data primer hasil observasi lapangan terhadap perahu karet yang beroperasi pada kegiatan wisata arung jeram di Sungai Pekalen, Kabupaten Probolinggo. Dataset terdiri dari 200 record data yang mencakup tujuh variabel prediktor dan satu variabel target. Variabel prediktor yang digunakan meliputi: Material, Tekanan Udara, Kapasitas, Kondisi Perahu, *Safety Equipment*, Tingkat Kesulitan, dan Karakteristik Sungai. Variabel target adalah Status kelayakan perahu yang dikategorikan menjadi dua kelas, yaitu Layak dan Tidak Layak, mengacu pada kriteria standar usaha wisata arung jeram yang ditetapkan dalam regulasi nasional [7].

Distribusi kelas pada dataset menunjukkan ketidakseimbangan (*class imbalance*) yang besar, di mana sebanyak 186 data (93%) dikategorikan Tidak Layak dan hanya 14 data (7%) dikategorikan Layak. Kondisi ini mencerminkan realitas lapangan bahwa sebagian besar perahu karet yang dioperasikan di Sungai Pekalen belum sepenuhnya memenuhi seluruh kriteria kelayakan secara holistik, sebagaimana juga diidentifikasi dalam studi keamanan wisata *rafting* di Kabupaten Probolinggo sebelumnya [4]. Ketidakseimbangan kelas ini perlu dipertimbangkan dalam interpretasi metrik evaluasi model, khususnya pada metrik yang sensitif terhadap distribusi kelas seperti F1 Score dan Matthews Correlation Coefficient (MCC).

1. Encoding Variabel Kategorikal

Seluruh variabel kategorikal dikonversi menjadi representasi numerik (label encoding) sebelum diproses oleh algoritma *Logistic Regression*. Proses encoding dilakukan berdasarkan urutan ordinal yang mencerminkan tingkat kondisi atau risiko dari masing-masing kategori, mengacu pada standar penilaian kondisi peralatan arung jeram yang termuat dalam SKKNI Arung Jeram [6] dan pedoman teknis Kementerian Pariwisata dan Ekonomi Kreatif [8]. Tabel 1 menjelaskan skema encoding yang diterapkan pada setiap variabel.

Tabel 1. Skema Encoding Variabel Kategorikal

No	Variabel	Kategori Asli	Nilai Encoding	Keterangan
1	Material	PVC	1	Bahan lebih umum
		Hypalon	2	Bahan lebih tahan lama
2	Tekanan Udara	Rendah	1	Tekanan di bawah standar
		Normal	2	Tekanan sesuai standar
3	Kondisi Perahu	Bocor	0	Kondisi paling buruk
		Bocor Ringan	1	Ada kebocoran kecil
		Baik	2	Kondisi layak pakai
		Sangat Baik	3	Kondisi prima
4	*Safety Equipment*	Tidak Lengkap	1	Peralatan kurang
		Lengkap	2	Peralatan lengkap
5	Tingkat Kesulitan	Sangat Mudah	0	Grade 1
		Mudah	1	Grade 2
		Menengah	2	Grade 3
		Sulit	3	Grade 4
		Sangat Sulit	4	Grade 5
6	Karakteristik Sungai	Arus cepat dengan jeram kecil	0	Kondisi ringan
		Arus tenang dan riak kecil	1	Kondisi sangat mudah
		Ombak tidak beraturan dan perlu manuver	2	Kondisi menengah
		Jeram panjang, arus kuat, banyak hambatan	3	Kondisi sulit
		Arus ekstrem dan rescue sulit	4	Kondisi sangat sulit
7	Status (Target)	Tidak Layak	1	Tidak memenuhi kelayakan
		Layak	2	Memenuhi kelayakan

Berdasarkan Tabel 1, variabel Material dikodekan dengan nilai 1 untuk PVC dan 2 untuk Hypalon, di mana Hypalon memiliki nilai lebih tinggi karena karakteristik ketahanan material yang lebih unggul. Variabel Tekanan Udara dikodekan dengan nilai 1 untuk kondisi Rendah dan 2 untuk kondisi Normal, mencerminkan bahwa tekanan udara normal merupakan kondisi yang lebih aman sesuai standar operasional [7]. Variabel Kondisi Perahu dikodekan secara ordinal dari Bocor (0) sebagai kondisi terburuk hingga Sangat Baik (3) sebagai kondisi terbaik. *Safety Equipment* dikodekan 1 untuk Tidak Lengkap dan 2 untuk Lengkap sesuai persyaratan kelengkapan peralatan keselamatan [6][8]. Tingkat Kesulitan dan Karakteristik Sungai dikodekan secara ordinal berdasarkan tingkat kesulitan dan risiko.

2. Pembagian Data Latih dan Data Uji

Dataset dibagi menjadi dua subset menggunakan metode *Random sampling Stratified* dengan proporsi 70:30, yaitu 70% (140 record) sebagai data latih dan 30% (60 record) sebagai data uji. Stratifikasi dilakukan untuk memastikan proporsi kelas Layak dan Tidak Layak tetap terjaga pada kedua subset, sehingga distribusi kelas tidak menjadi sumber bias dalam proses pelatihan maupun evaluasi model. Eksperimen diulang sebanyak 10 kali dengan seed acak berbeda untuk memperoleh estimasi performa yang lebih stabil dan reliabel, sesuai dengan konfigurasi yang diterapkan pada perangkat lunak Orange *Data mining* [10].

Tabel 2. Distribusi Data Latih dan Data Uji

Subset Data	Total Record	Layak	Tidak Layak	Persentase Split
Data Latih	140	10 (7,1%)	130 (92,9%)	70%
Data Uji	60	4 (6,7%)	56 (93,3%)	30%
Total	200	14 (7%)	186 (93%)	100%

3.2 Pemodelan *Logistic Regression*

Logistic Regression merupakan algoritma klasifikasi statistik yang memodelkan probabilitas suatu observasi termasuk dalam kelas tertentu menggunakan fungsi sigmoid (logistik). Pada penelitian ini, *Logistic Regression* digunakan untuk mengklasifikasikan kelayakan perahu karet ke dalam dua kelas, yaitu Layak (kelas 2) dan Tidak Layak (kelas 1). Model dilatih menggunakan Maximum Likelihood Estimation (MLE) untuk mengestimasi koefisien regresi yang optimal. Pendekatan klasifikasi berbasis algoritma ini telah terbukti efektif dalam berbagai domain prediksi dan klasifikasi biner, termasuk prediksi kualitas berbasis data sensor [10].

Persamaan umum model *Logistic Regression* yang dibangun adalah sebagai berikut:

$$P(Y) = \frac{1}{1+e^{-(\beta_0+\beta_1X_1+\beta_2X_2+\cdots+\beta_nX_n)}} \quad (1)$$

Keterangan:
1. $P(Y)$= probabilitas kelayakan perahu,
2. β_0= konstanta,
3. $\beta_1 \ldots \beta_n$= koefisien variabel,
4. $X_1 \ldots X_n$= variabel independen.

Metode *Logistic Regression* dipilih karena mampu menghasilkan model klasifikasi biner yang efektif dalam menentukan status "Layak" dan "Tidak Layak" berdasarkan probabilitas yang dihasilkan dari proses analisis data.

1. Koefisien Model dan Odds Ratio

Setelah proses pelatihan, model menghasilkan koefisien regresi untuk setiap variabel prediktor. Tabel 3 menyajikan koefisien model beserta nilai Odds Ratio (eβ) yang menunjukkan pengaruh masing-masing variabel terhadap probabilitas kelayakan perahu.

Tabel 3. Koefisien *Logistic Regression* dan Odds Ratio

No	Variabel Prediktor	Koefisien (β)	Odds Ratio (eβ)	Pengaruh
1	*Safety Equipment*	2.068	7.912	Positif (+)
2	Tekanan Udara	2.003	7.412	Positif (+)
3	Kondisi Perahu	1.027	2.793	Positif (+)
4	Kapasitas	0.072	1.074	Positif (+)
5	Tingkat Kesulitan	-0.268	0.765	Negatif (-)
6	Karakteristik Sungai	-0.327	0.721	Negatif (-)
7	Material	-0.572	0.564	Negatif (-)
	Intercept (β_0)	-9.728	-	-

Berdasarkan Tabel 3, intercept model adalah -9,728. Variabel *Safety Equipment* memiliki koefisien tertinggi sebesar 2,068 dengan odds ratio 7,912, yang mengindikasikan bahwa kelengkapan *Safety Equipment* meningkatkan odds kelayakan perahu sebesar 7,91 kali. Hal ini selaras dengan ketentuan Permenparekraf Nomor 13 Tahun 2014 [7] dan SKKNI Arung Jeram [6] yang mewajibkan kelengkapan peralatan keselamatan sebagai syarat operasional utama wisata arung jeram. Variabel Tekanan Udara menempati posisi kedua dengan koefisien 2,003 dan odds ratio 7,412, menunjukkan bahwa perahu dengan tekanan udara Normal memiliki peluang layak 7,41 kali lebih besar. Pentingnya perawatan tekanan alat keselamatan juga dikonfirmasi dalam studi perawatan peralatan keselamatan di sektor pengelolaan sarana transportasi air [9]. Variabel Kondisi Perahu memiliki koefisien 1,027 dengan odds ratio 2,793.

Sementara itu, variabel Material, Karakteristik Sungai, dan Tingkat Kesulitan memiliki koefisien negatif. Variabel Material dengan koefisien -0,572 menunjukkan bahwa jenis material saja tidak menjadi penentu utama kelayakan apabila faktor keselamatan lainnya tidak terpenuhi. Variabel Tingkat Kesulitan (-0,268) dan Karakteristik Sungai (-0,327) menunjukkan bahwa semakin tinggi tingkat kesulitan dan kompleksitas sungai, semakin besar kemungkinan perahu diklasifikasikan Tidak Layak karena standar kelayakan yang lebih ketat, konsisten dengan standar pengelolaan wisata arung jeram [1][7].

1. *Confusion matrix*

Confusion matrix merupakan alat evaluasi utama untuk mengukur performa model klasifikasi. Tabel 4 menyajikan *confusion matrix* hasil pengujian model *Logistic Regression* pada data uji (60 record) dengan konfigurasi *random sampling* 70:30 *stratified*.

Evaluasi performa model klasifikasi dilakukan menggunakan empat metrik utama: Accuracy, Precision, Recall, dan F1-Score. Keempat metrik tersebut dihitung berdasarkan nilai True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) yang diperoleh dari *confusion matrix* sebagaimana disajikan pada Tabel 4.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (5)$$

Keterangan: *TP (True Positive)* = jumlah data Layak yang diprediksi benar sebagai Layak; *TN (True Negative)* = jumlah data Tidak Layak yang diprediksi benar sebagai Tidak Layak; *FP (False Positive)* = jumlah data Tidak Layak yang salah diprediksi sebagai Layak; *FN (False Negative)* = jumlah data Layak yang salah diprediksi sebagai Tidak Layak.

2. Validasi Silang 5-*Fold* (*Cross validation*)

Untuk memverifikasi stabilitas dan kemampuan generalisasi model, dilakukan validasi silang 5-*fold stratified* (5-*Fold Stratified Cross validation*). Metode ini membagi dataset menjadi 5 subset (*fold*) yang proporsional, melatih model pada 4 *fold*, dan mengujinya pada 1 *fold*, kemudian merotasi hingga seluruh *fold* menjadi data uji. Hasil validasi silang disajikan pada Tabel 5.

Tabel 5. Hasil Validasi Silang 5-*Fold Logistic Regression*

Metrik Evaluasi	Rata-rata	Std. Deviasi	Rentang
Akurasi (CA)	0.935	± 0.034	0.900 – 0.983
F1 Score	0.925	± 0.033	0.890 – 0.960
AUC	0.974	± 0.040	0.920 – 1.000

Hasil validasi silang pada Tabel 5 menunjukkan bahwa model *Logistic Regression* mempertahankan performa yang konsisten dengan akurasi rata-rata 93,5% (±3,4%) dan AUC rata-rata 97,4% (±4,0%). Nilai standar deviasi yang relatif kecil mengindikasikan bahwa model tidak mengalami *overfitting* dan memiliki stabilitas yang baik lintas subset data yang berbeda. Konsistensi performa ini memperkuat keyakinan bahwa model dapat digeneralisasi ke data baru di luar dataset pelatihan, sehingga layak diimplementasikan sebagai alat bantu evaluasi kelayakan perahu karet di lapangan [7][8].

3.4 Hasil Pengujian Model

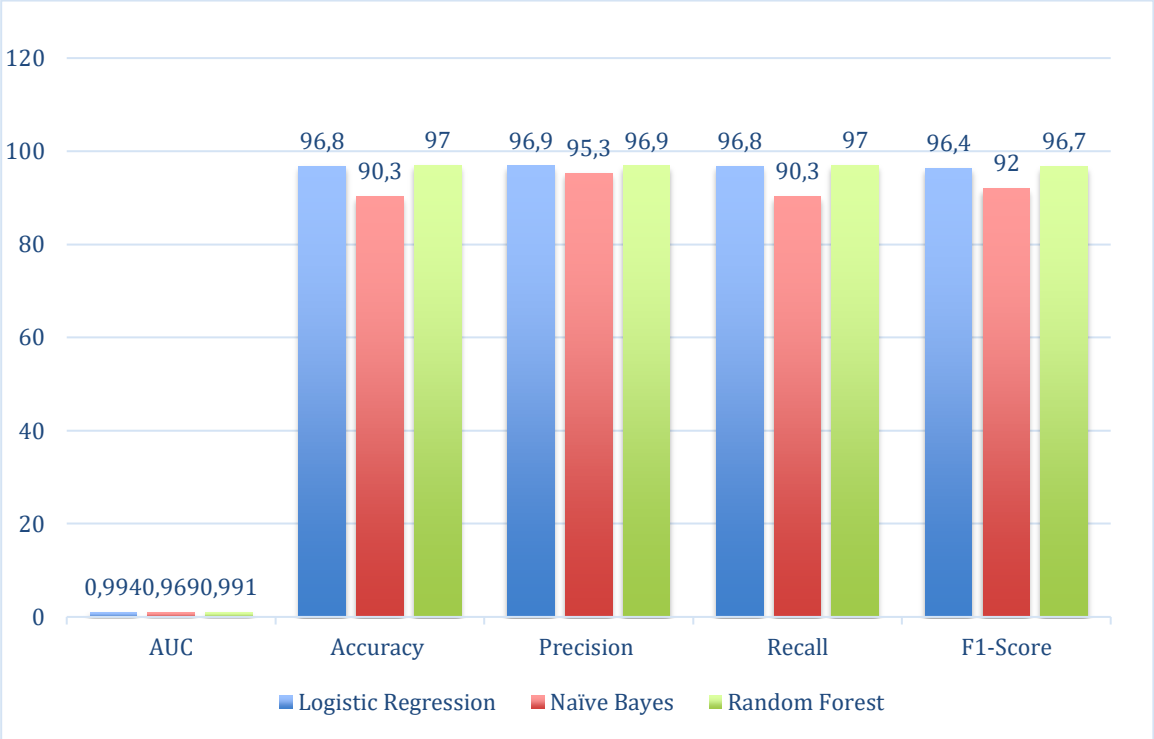
Untuk menilai keunggulan komparatif *Logistic Regression*, model dibandingkan dengan dua algoritma klasifikasi lainnya, yaitu *Naïve Bayes* dan *Random Forest*. Perbandingan dilakukan menggunakan konfigurasi identik: *random sampling* 10 kali, pembagian data 70:30, dan stratifikasi kelas. Algoritma *Naïve Bayes* dipilih sebagai pembanding karena relevansinya dengan domain klasifikasi berbasis probabilitas sebagaimana diterapkan dalam penelitian terdahulu pada prediksi kualitas berbasis data [10]. Hasil perbandingan disajikan pada Tabel 6.

Tabel 6. Perbandingan Performa Model Klasifikasi (Rata-rata 10 kali *Random sampling*)

Model	AUC	Akurasi (CA)	F1 Score	Precision	Recall	MCC
Logistic Regression	0.994	0.968	0.964	0.969	0.968	0.713
Naïve Bayes	0.969	0.903	0.920	0.953	0.903	0.561
Random Forest	0.991	0.970	0.967	0.969	0.970	0.731

Berdasarkan Tabel 6, *Logistic Regression* menghasilkan AUC tertinggi sebesar 0,994 dibandingkan *Naïve Bayes* (0,969) dan sedikit di atas *Random Forest* (0,991). Nilai akurasi *Logistic Regression* (0,968) setara dengan *Random Forest* (0,970) dan jauh melampaui *Naïve Bayes* (0,903). Nilai F1 Score *Logistic Regression* (0,964) juga kompetitif dengan *Random Forest* (0,967). Nilai-nilai ini konsisten dengan tampilan hasil evaluasi pada perangkat lunak Orange *Data mining*.

Meskipun *Random Forest* menunjukkan akurasi yang sedikit lebih tinggi, *Logistic Regression* dipilih sebagai model utama dengan beberapa justifikasi. Pertama, *Logistic Regression* menghasilkan koefisien yang dapat diinterpretasikan secara langsung dalam bentuk odds ratio, memberikan wawasan *actionable* bagi pengelola wisata arung jeram untuk mengidentifikasi faktor risiko dominan sesuai amanat Permenparekraf Nomor 13 Tahun 2014 [7]. Kedua, model lebih sederhana dan praktis untuk dikembangkan lebih lanjut sebagai komponen sistem pendukung keputusan di lapangan, sesuai pedoman teknis pengelolaan wisata arung jeram [8]. Ketiga, AUC *Logistic Regression* (0,994) yang lebih tinggi dari *Random Forest* (0,991) menunjukkan kemampuan diskriminasi probabilistik yang lebih baik secara keseluruhan.



Gambar 1. Diagram perbandingan model klasifikasi

Berdasarkan Gambar 1 dan Tabel 6, ketiga algoritma klasifikasi yang diuji, yaitu *Logistic Regression*, *Naïve Bayes*, dan *Random Forest*, menunjukkan performa yang baik dalam melakukan klasifikasi kelayakan perahu karet. Namun demikian, terdapat perbedaan tingkat kinerja pada setiap metrik evaluasi yang digunakan.

Pada metrik Accuracy, *Random Forest* memperoleh nilai tertinggi sebesar 97,0%, diikuti *Logistic Regression* sebesar 96,8%, dan *Naïve Bayes* sebesar 90,3%. Hasil ini menunjukkan bahwa *Random Forest* memiliki kemampuan terbaik dalam mengklasifikasikan data secara keseluruhan dengan tingkat kesalahan yang paling rendah. Selisih Accuracy antara *Random Forest* dan *Logistic Regression* relatif kecil, yaitu hanya sebesar 0,2%, yang menunjukkan bahwa kedua model memiliki performa yang hampir setara dalam melakukan klasifikasi.

Pada metrik Precision, *Logistic Regression* dan *Random Forest* menghasilkan nilai yang sama, yaitu 96,9%, sedangkan *Naïve Bayes* memperoleh nilai 95,3%. Nilai precision yang tinggi menunjukkan bahwa sebagian besar prediksi yang dikategorikan sebagai "Layak" benar-benar sesuai dengan kondisi aktual. Dengan demikian, *Logistic Regression* dan *Random Forest* memiliki kemampuan yang sangat baik dalam meminimalkan kesalahan prediksi positif (False Positive).

Pada metrik Recall, *Random Forest* kembali menunjukkan performa terbaik dengan nilai 97,0%, sedikit lebih tinggi dibandingkan *Logistic Regression* sebesar 96,8%, sedangkan *Naïve Bayes* memperoleh nilai 90,3%. Nilai recall yang tinggi mengindikasikan bahwa model mampu mendeteksi hampir seluruh data yang benar-benar termasuk dalam kategori yang sesuai. Dalam konteks penelitian ini, kemampuan recall yang tinggi sangat penting karena berkaitan dengan kemampuan sistem dalam mengidentifikasi kondisi kelayakan perahu secara tepat sehingga dapat mendukung aspek keselamatan operasional arung jeram.

Pada metrik F1-Score, *Random Forest* memperoleh nilai tertinggi sebesar 96,7%, diikuti *Logistic Regression* sebesar 96,4%, dan *Naïve Bayes* sebesar 92,0%. F1-Score merupakan kombinasi antara precision dan recall sehingga dapat memberikan gambaran yang lebih seimbang mengenai performa model. Tingginya nilai F1-Score pada *Random Forest* menunjukkan bahwa model tersebut memiliki keseimbangan yang sangat baik antara kemampuan mendeteksi kelas secara benar dan meminimalkan kesalahan klasifikasi.

Selain itu, berdasarkan nilai AUC (Area Under Curve), *Logistic Regression* memperoleh nilai tertinggi sebesar 0,994, dibandingkan *Random Forest* sebesar 0,991 dan *Naïve Bayes* sebesar 0,969. Nilai AUC yang mendekati 1 menunjukkan kemampuan model yang sangat baik dalam membedakan kelas "Layak" dan "Tidak Layak". Hasil ini mengindikasikan bahwa *Logistic Regression* memiliki kemampuan diskriminatif yang sedikit lebih baik dibandingkan model lainnya dalam memisahkan kedua kelas tersebut.

Jika ditinjau dari seluruh metrik evaluasi, *Random Forest* memang menghasilkan nilai Accuracy, Recall, dan F1-Score tertinggi. Akan tetapi, perbedaan performanya terhadap *Logistic Regression* sangat kecil. Di sisi lain, *Logistic Regression* memiliki keunggulan dalam aspek interpretabilitas model karena mampu menjelaskan pengaruh masing-masing variabel melalui nilai koefisien regresi dan odds ratio. Keunggulan ini menjadikan *Logistic Regression* lebih mudah digunakan sebagai dasar pengambilan keputusan dalam evaluasi kelayakan perahu karet dibandingkan model *Random Forest* yang bersifat lebih kompleks.

Oleh karena itu, meskipun *Random Forest* menunjukkan performa klasifikasi sedikit lebih tinggi, *Logistic Regression* tetap dipilih sebagai model utama dalam penelitian ini karena mampu memberikan keseimbangan antara tingkat akurasi yang tinggi dan kemampuan interpretasi hasil yang lebih baik. Karakteristik tersebut sangat penting apabila model ini kelak dikembangkan menjadi komponen sistem pendukung keputusan keselamatan wisata arung jeram, di mana hasil klasifikasi tidak hanya harus akurat tetapi juga dapat dijelaskan secara ilmiah kepada pengelola maupun pihak terkait.

3.5 Pembahasan

Hasil penelitian menunjukkan bahwa metode *Logistic Regression* mampu mengklasifikasikan kelayakan perahu karet pada aktivitas arung jeram di Sungai Pekalen, Kabupaten Probolinggo dengan tingkat akurasi 96,8% dan AUC 0,994, yang termasuk dalam kategori excellent. Temuan ini relevan dengan penelitian terdahulu yang menggunakan pendekatan *data mining* berbasis klasifikasi untuk prediksi biner pada domain kualitas dan keselamatan [10], serta konsisten dengan studi pengelolaan standar usaha wisata arung jeram di Indonesia [1][2].

Temuan yang paling menonjol dari analisis koefisien adalah bahwa *Safety Equipment* merupakan variabel dengan koefisien terbesar terhadap kelayakan perahu (β = 2,068, OR = 7,912). Hal ini mengindikasikan bahwa ketersediaan perlengkapan keselamatan yang lengkap merupakan prasyarat utama kelayakan yang tidak dapat dikompensasi oleh kondisi fisik perahu yang baik sekalipun. Temuan ini konsisten dengan ketentuan SKKNI Arung Jeram [6] yang menetapkan kelengkapan *Safety Equipment* sebagai kompetensi wajib, serta Permenparekraf Nomor 13 Tahun 2014 [7] yang mewajibkan pemenuhan standar peralatan keselamatan sebagai syarat sertifikasi usaha arung jeram. Studi sosialisasi olahraga arung jeram di Indonesia [2] juga menegaskan pentingnya pemenuhan standar keselamatan dalam pengembangan wisata arung jeram yang berkelanjutan.

Variabel Tekanan Udara menempati posisi kedua dalam tingkat pengaruh (β = 2,003, OR = 7,412). Temuan ini memiliki implikasi praktis penting: tekanan udara perahu karet yang tidak sesuai standar secara langsung meningkatkan risiko deflasi saat navigasi jeram. Pentingnya pemeriksaan dan perawatan tekanan peralatan keselamatan secara berkala sebelum operasional juga didukung oleh studi perawatan alat keselamatan di sektor transportasi air [9]. Pengelola wisata arung jeram Sungai Pekalen disarankan melakukan pengecekan tekanan udara setiap unit perahu sebelum sesi pengarungan sebagai bagian dari SOP keselamatan sesuai pedoman Kemenparekraf [8].

Kondisi fisik perahu (β = 1,027, OR = 2,793) memiliki koefisien positif, di mana setiap kenaikan satu tingkat kondisi fisik pada skala encoding (Tabel 1) meningkatkan odds kelayakan hampir tiga kali lipat. Temuan ini mendukung perlunya jadwal inspeksi fisik perahu yang terstruktur dan terdokumentasi, termasuk pemeriksaan kebocoran, kondisi lapisan material, dan integritas sambungan, selaras dengan studi evaluasi pembinaan arung jeram [3] dan analisis keselamatan wisata *rafting* di Probolinggo [4] yang mengidentifikasi kondisi fisik peralatan sebagai determinan utama keselamatan aktivitas arung jeram.

Dibandingkan dengan penelitian terdahulu, kontribusi spesifik penelitian ini dapat diuraikan pada tiga hal. Pertama, studi mengenai standar usaha wisata arung jeram [1] dan keamanan wisatawan pada wisata *rafting* di Probolinggo [4] melaporkan pemenuhan standar usaha serta kelemahan implementasi keselamatan, antara lain pada kelengkapan peralatan, SOP, dan sumber daya manusia; penelitian ini menambahkan estimasi kuantitatif keterkaitan faktor-faktor peralatan dengan status kelayakan perahu melalui koefisien dan odds ratio, sehingga *Safety Equipment*, tekanan udara, dan kondisi perahu dapat diurutkan menurut besar koefisiennya. Kedua, penelitian terdahulu pada klasifikasi kualitas air [10] menerapkan *Naïve Bayes*; penelitian ini memindahkan pendekatan klasifikasi tersebut ke domain kelayakan perahu karet dan menunjukkan bahwa pada konfigurasi pengujian yang identik, *Logistic Regression* (akurasi 0,968; AUC 0,994) lebih baik daripada *Naïve Bayes* (akurasi 0,903; AUC 0,969) dan setara dengan *Random Forest* dengan keunggulan interpretabilitas. Ketiga, studi pembinaan arung jeram [3] dan perawatan alat keselamatan [9] menekankan pentingnya kualitas serta pemeliharaan peralatan; temuan bahwa tekanan udara dan kondisi perahu memiliki koefisien positif yang besar sejalan dengan penekanan tersebut. Meskipun demikian, variabel administrasi, pelatihan sumber daya manusia, dan kepatuhan SOP yang dibahas pada studi terdahulu [1], [4] tidak termasuk dalam dataset penelitian ini, sehingga pengaruhnya belum dapat dibandingkan secara langsung.

Dari perspektif manajemen risiko wisata, model *Logistic Regression* yang dibangun berpotensi menjadi calon komponen *decision support system* berbasis data untuk membantu penilaian kelayakan perahu karet sebelum pengarungan, sejalan dengan arah kebijakan Undang-Undang Nomor 10 Tahun 2009 tentang Kepariwisataan [5] yang menekankan keselamatan wisatawan sebagai tanggung jawab penyelenggara wisata. Namun, penelitian ini melaporkan pelatihan dan evaluasi model klasifikasi, bukan implementasi atau pengujian suatu sistem informasi, sehingga klaim kontribusi dibatasi pada model klasifikasi. Kemampuan model untuk menginterpretasikan pengaruh setiap faktor melalui odds ratio memberikan kuantifikasi risiko yang terukur, tetapi keunggulannya dibandingkan inspeksi konvensional belum diuji dalam penelitian ini.

Agar model dapat dimanfaatkan sebagai komponen sistem pendukung keputusan, diperlukan rancangan sistem yang belum dicakup dalam penelitian ini. Alur konseptual yang diusulkan adalah sebagai berikut. Data inspeksi perahu dicatat oleh petugas untuk ketujuh variabel prediktor dan dikodekan sesuai Tabel 1, dengan mengacu pada pemeriksaan kondisi peralatan sebelum pengarungan [8]. Model menghitung probabilitas kelayakan perahu berdasarkan Persamaan (1) dan mengubahnya menjadi keputusan Layak atau Tidak Layak menggunakan

ambang probabilitas yang perlu ditetapkan melalui validasi. Hasil ditampilkan kepada petugas bersama nilai probabilitas dan kontribusi tiap variabel berdasarkan odds ratio. Kasus yang memerlukan pemeriksaan ahli, misalnya probabilitas yang mendekati ambang atau perahu yang diprediksi Layak tetapi mengalami kebocoran, dirujuk kepada inspektur yang berkompetensi. Keputusan model perlu dibandingkan dengan keputusan petugas inspeksi, misalnya melalui tingkat kesesuaian, sebelum digunakan, dan model dilatih ulang serta dievaluasi kembali secara berkala ketika data inspeksi baru tersedia. Seluruh alur tersebut belum diimplementasikan maupun diuji dalam penelitian ini dan menjadi agenda penelitian lanjutan.

Terkait keterbatasan penelitian, adanya *class imbalance* (93% Tidak Layak vs. 7% Layak) berpotensi memengaruhi kemampuan model dalam mendeteksi kasus Layak secara akurat, sebagaimana terlihat dari nilai FN = 3 pada *confusion matrix*. Selain itu, penelitian ini belum melaporkan uji signifikansi statistik terhadap koefisien model (misalnya uji Wald beserta nilai p atau selang kepercayaan odds ratio), sehingga pengaruh antarvariabel diinterpretasikan berdasarkan besar koefisien dan odds ratio dan belum dapat dinyatakan signifikan secara statistik. Lingkup penelitian juga dibatasi pada pembangunan dan evaluasi model, tanpa implementasi sistem dan pengujian pengguna. Penelitian lanjutan disarankan untuk mengeksplorasi teknik penanganan imbalanced data seperti SMOTE (Synthetic Minority Oversampling Technique) guna meningkatkan sensitivitas deteksi kelas Layak, memperluas cakupan data ke destinasi arung jeram lain di Indonesia untuk meningkatkan generalisabilitas model [1][2], serta merancang prototipe sistem pendukung keputusan dan mengujinya bersama petugas inspeksi.

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa metode *Logistic Regression* mampu diterapkan untuk menganalisis kelayakan perahu karet pada aktivitas arung jeram di Sungai Pekalen Probolinggo secara efektif dan berbasis data. Metode ini dapat digunakan untuk mengklasifikasikan kondisi perahu ke dalam kategori "Layak" dan "Tidak Layak" berdasarkan beberapa variabel yang memengaruhi kondisi perahu, yaitu material, tekanan udara, kapasitas, kondisi perahu, *Safety Equipment*, tingkat kesulitan, dan karakteristik sungai.

Hasil pengujian model menunjukkan bahwa metode *Logistic Regression* memiliki performa klasifikasi yang baik dengan rata-rata akurasi sebesar 96,8%, precision 96,9%, recall 96,8%, F1-score 96,4%, dan AUC 0,994 pada 10 kali pengulangan pembagian data 70:30, serta akurasi rata-rata 93,5% dan AUC rata-rata 97,4% pada validasi silang 5-*fold*. Nilai tersebut menunjukkan bahwa model mampu memprediksi kelayakan perahu dengan baik pada data penelitian ini, sehingga berpotensi menjadi calon komponen sistem pendukung keputusan dalam evaluasi keselamatan wisata arung jeram setelah dirancang dan diuji lebih lanjut. Distribusi kelas yang tidak seimbang (93% Tidak Layak) perlu dipertimbangkan dalam menafsirkan nilai tersebut.

Berdasarkan koefisien model, variabel dengan koefisien positif terbesar terhadap peluang perahu dinyatakan layak adalah *Safety Equipment* (odds ratio 7,912), tekanan udara (7,412), dan kondisi perahu (2,793). Perahu dengan perlengkapan keselamatan tidak lengkap, tekanan udara rendah, dan kondisi fisik yang buruk cenderung diklasifikasikan Tidak Layak. Pengaruh variabel tersebut belum diuji signifikansi statistiknya sehingga perlu dikonfirmasi pada penelitian lanjutan.

Penerapan metode *Logistic Regression* dalam penelitian ini memberikan kontribusi berupa model klasifikasi kelayakan perahu karet berbasis *data mining* pada sektor wisata olahraga (*sport tourism*), khususnya wisata arung jeram, yang dapat menjadi calon komponen sistem pendukung keputusan. Model klasifikasi yang dibangun berpotensi membantu operator *rafting* melakukan pemeriksaan kelayakan perahu secara lebih objektif dan sistematis. Perbandingan dengan keputusan petugas inspeksi, perancangan sistem, serta pengujian pengguna belum dilakukan dan menjadi agenda penelitian lanjutan.

Penelitian ini diharapkan dapat menjadi referensi bagi pengelola wisata arung jeram dalam meningkatkan standar keselamatan operasional *rafting* melalui penerapan teknologi analisis data. Selain itu, penelitian ini juga dapat menjadi dasar pengembangan penelitian selanjutnya, yaitu penggunaan dataset yang lebih besar, penanganan ketidakseimbangan kelas, serta perancangan dan pengujian prototipe sistem pendukung keputusan bersama petugas inspeksi.

DAFTAR PUSTAKA

[1] T. N. Wahyudi and Supriyono, "Analisis Standar Usaha Wisata Arung Jeram Progo *Rafting*," *J. Edukasi Citra Olahraga*, vol. 5, no. 3, pp. 468–477, 2025.

[2] Tofikin, Suhermon, F. Hendryanto, Siska, and Arisman, "Sosialiasi Cabang Olahraga Arung Jeram Pada Pemuda Desa Babusalam," *J. Masy. Negeri Rokania*, vol. 6, no. 2, 2025.

[3] K. Leksono, "Studi Evaluasi Program Pembinaan Prestasi Arung Jeram di FAJI Banyumas Tahun 2017–2022," Universitas Negeri Yogyakarta, 2023.

[4] Z. M. Rohmah, "Analisis Keamanan dan Keselamatan Wisatawan pada Wisata *Rafting* (Studi pada Songa Adventure *Rafting* di Kabupaten Probolinggo)," Universitas Brawijaya, 2019.

[5] R. Indonesia, "Undang-Undang Republik Indonesia Nomor 10 Tahun 2009 tentang Kepariwisataan," 2009.

[6] K. K. R. Indonesia, "Standar Kompetensi Kerja Nasional Indonesia (SKKNI) Arung Jeram," 2024.

[7] R. Indonesia, "Peraturan Menteri Pariwisata dan Ekonomi Kreatif Republik Indonesia Nomor 13 Tahun 2014 tentang Standar Usaha Wisata Arung Jeram," 2014.
[8] K. P. dan E. K. R. Indonesia, *BERARUNG*. Jakarta, 2022. [Online]. Available: https://kemenparekraf.go.id/
[9] A. A. Alfisar, F. M. Pongoh, P. D. Darmawan, and D. Anwar, "Perawatan Alat Keselamatan Liferaft Tug Boat di PT. Kartika Samudra Adijaya Site Berau," *Mar. Transp. Manag. Logist. J.*, vol. 1, no. 1, pp. 46–53, 2025.
[10] A. W. Saputra, A. I. Purnamasari, and I. Ali, "Implementasi Algoritma Naïve Bayes untuk Memprediksi Kualitas Air yang Dapat Dikonsumsi," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 1, pp. 133–140, 2024.