Institut Riset dan Publikasi Indonesia (IRPI)
MALCOM: Indonesian Journal of Machine Learning and Computer Science
Journal Homepage: https://journal.irpi.or.id/index.php/malcom
Vol. 6 Iss. 1 January 2026, pp: 106-115
ISSN(P): 2797-2313 | ISSN(E): 2775-8575

Implementation of the XGBoost Classifier Algorithm for Classifying Crowds of Hajj and Umrah Pilgrims

Implementasi Algoritma XGBoost Classifier untuk Klasifikasi Kerumunan Jemaah Haji dan Umrah

Rahmattun Illaihiyah[1*], Ginanjar Abdurrahman[2], Daryanto[3]

[1,2,3]Program Studi Teknik Informatika, Universitas Muhammadiyah Jember, Indonesia

E-Mail: [1]rahmatclickdownload@gmail.com, [2]abdurrahmanginanjar@unmuhjember.ac.id, [3]daryanto@unmuhjember.ac.id

Received Aug 09th 2025; Revised Nov 15th 2025; Accepted Dec 24th 2025; Available Online Jan 16th 2026
Corresponding Author: Rahmattun Illaihiyah

Abstract

Effective crowd management is essential during the Hajj and Umrah, considering the high mobility of pilgrims at specific times and locations. Uncontrolled crowd density can pose safety risks and disrupt the smooth execution of worship activities. This study proposes a crowd density classification approach that utilizes operational data to represent movement dynamics, individual behavior, and environmental conditions. The dataset consists of 10,000 entries with 29 features. The Extreme Gradient Boosting (XGBoost) algorithm was selected for its ability to process complex and heterogeneous data. The testing results show an accuracy of 98.7%, precision of 99%, recall of 99%, and f1-score of 99% for low, medium, and high density categories, in both macro and weighted average. These findings indicate that the developed model demonstrates high accuracy and reliability in predicting crowd density.

Keywords: Crowd Density, Crowd Management, Extreme Gradient Boosting, Hajj and Umrah, Machine Learning

Abstrak

Efektivitas manajemen kerumunan menjadi aspek penting dalam pelaksanaan ibadah haji dan umrah, mengingat mobilitas jemaah yang sangat tinggi pada waktu dan lokasi tertentu. Kepadatan yang tidak terkendali berpotensi menimbulkan risiko keselamatan dan mengganggu kelancaran ibadah. Penelitian ini mengusulkan pendekatan klasifikasi tingkat kepadatan kerumunan (*crowd density*) dengan memanfaatkan data operasional yang merepresentasikan dinamika pergerakan, perilaku individu, dan kondisi lingkungan. Dataset yang digunakan terdiri dari 10.000 entri dengan 29 fitur. Algoritma Extreme Gradient Boosting (XGBoost) dipilih karena kemampuannya mengolah data kompleks dan heterogen. Hasil pengujian menunjukkan akurasi 98.7%, *precision* 99%, *recall* 99%, dan *f1-score* 99% untuk kategori rendah, sedang, dan tinggi, baik pada *macro* maupun *weighted average*. Temuan ini menunjukkan bahwa model yang dikembangkan memiliki tingkat akurasi dan keandalan yang tinggi dalam memprediksi kepadatan kerumunan.

Kata Kunci: Extreme Gradient Boosting, Haji dan Umrah, Kepadatan Kerumunan, Klasifikasi, Pembelajaran Mesin

1. PENDAHULUAN

Haji dan umrah merupakan ibadah yang dilaksanakan dengan mengunjungi Baitullah di Makkah melalui rangkaian ritual tertentu. Secara etimologis, kata 'haji' berasal dari bahasa Arab *ḥajja* (حَجَّ) yang berarti menyengaja menuju sesuatu yang agung, sedangkan 'umrah' berasal dari akar kata *i'timar* (اعتمر) yang berarti berkunjung atau berziarah [1][2]. Kedua ibadah ini dilaksanakan oleh setiap Muslim yang mampu secara fisik, mental, dan finansial [3]. Pelaksanaan ibadah haji dilakukan secara serentak oleh umat Islam dari berbagai penjuru dunia pada waktu tertentu setiap tahunnya, sedangkan ibadah umrah dapat dilaksanakan kapan saja sepanjang tahun. Keduanya tidak hanya merupakan kewajiban atau tindakan religius yang sangat dianjurkan, tetapi juga mencerminkan interaksi global yang memperkuat ukhuwah Islamiyah serta solidaritas antarbangsa. Skala dan kompleksitas pelaksanaan ibadah ini menuntut manajemen yang sistematis serta koordinasi tinggi antara negara-negara pengirim jemaah dengan Kerajaan Arab Saudi. Manajemen haji dan

umrah yang meliputi aspek transportasi, akomodasi, kerumunan, pelayanan kesehatan, dan keamanan, sangat penting untuk memastikan jemaah dapat menjalankan ibadah dengan khusyuk, nyaman, dan aman [4].

Kepadatan kerumunan menjadi salah satu tantangan terbesar dalam manajemen ibadah haji dan umrah, disebabkan oleh keterbatasan ruang dan waktu yang mengakibatkan tingginya intensitas jemaah di lokasi-lokasi tertentu, seperti Masjidil Haram, Mina, Muzdalifah, dan Arafah. Kondisi ini menuntut adanya strategi khusus dalam menganalisis pola kerumunan jemaah guna mengidentifikasi tingkat kepadatan secara efektif [5][6]. Guna mendukung hal tersebut, dibutuhkan data yang mampu merepresentasikan kondisi di lapangan secara faktual dan komprehensif. Salah satu sumber yang dapat digunakan adalah dataset sekunder yang tersedia pada platform pusat data dan statistik internasional, seperti Kaggle, yang menyediakan beragam data dengan cakupan global serta relevansi tinggi terhadap topik yang diteliti.

Menurut Permana (2025), lebih dari 18,5 juta jamaah Muslim dari berbagai negara telah menunaikan ibadah haji dan umrah sepanjang tahun 2024, menjadikannya periode dengan partisipasi tertinggi dalam sejarah ziarah Islam. Informasi ini disampaikan oleh Menteri Haji dan Umrah Arab Saudi, Tawfiq bin Fauzan Al-Rabiah, dalam Konferensi dan Pameran Haji dan Umrah keempat di Jeddah pada 20 Januari 2025 [7]. Kondisi ini menegaskan bahwa strategi manajemen kerumunan harus mengintegrasikan observasi terhadap faktor lingkungan, perilaku jamaah, dan dukungan teknologi guna mengantisipasi dinamika pola kepadatan yang terus berubah.

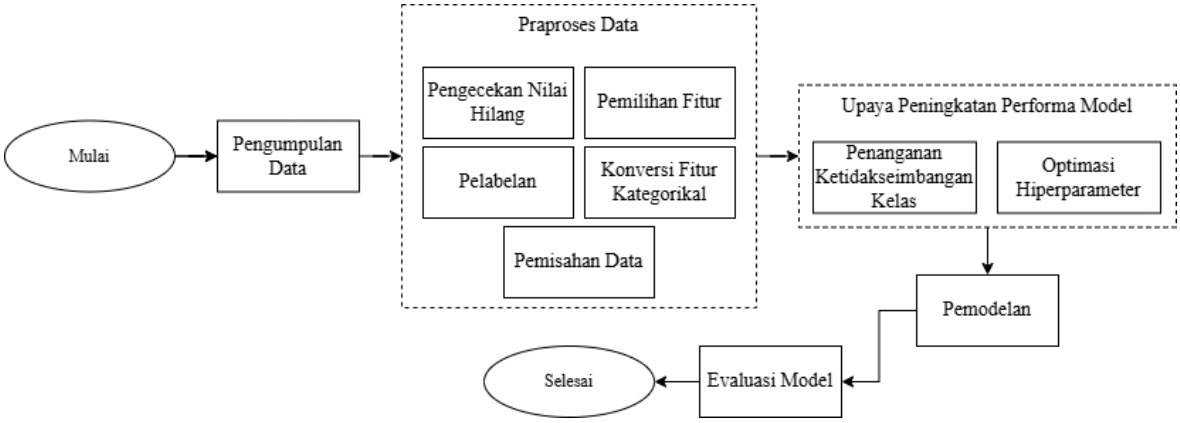
Salah satu pendekatan yang dapat digunakan untuk memperoleh pemahaman yang lebih mendalam terhadap pola kerumunan jemaah adalah dengan menerapkan algoritma prediksi kerumunan otomatis berbasis klasifikasi [5]. Algoritma ini mengklasifikasikan tingkat kepadatan menjadi tiga kategori, yaitu rendah, sedang, dan tinggi, sehingga karakteristik khas seperti kecepatan pergerakan, waktu tinggal, jarak antarindividu, frekuensi interaksi, dan tingkat kelelahan dapat diidentifikasi. Dengan demikian, pendekatan ini memungkinkan pemetaan kondisi kerumunan secara lebih sistematis dan terstruktur.

Metode klasifikasi yang digunakan dalam penelitian ini adalah Extreme Gradient Boosting (XGBoost), yaitu algoritma berbasis pohon keputusan yang dikenal unggul dalam akurasi serta efektif dalam menangani dataset berskala besar dan fitur yang beragam [8]. XGBoost bekerja dengan membangun model prediktif secara bertahap (*ensemble method*) melalui pendekatan *gradient boosting*, di mana setiap pohon keputusan yang ditambahkan berfungsi memperbaiki kesalahan prediksi dari iterasi sebelumnya [9][10]. Keunggulan utama XGBoost terletak pada skalabilitasnya, yang tercermin melalui kemampuan teknis seperti optimasi pemrosesan data *sparse* dan terbobot, penerapan algoritma *quantile sketch* untuk efisiensi komputasi, serta dukungan penuh terhadap komputasi paralel dan terdistribusi [11]. Selain itu, XGBoost menerapkan regularisasi untuk mengendalikan kompleksitas model sehingga mampu meningkatkan generalisasi dan mengurangi risiko *overfitting* [12]. Kemampuan tersebut membuat XGBoost efektif dalam menangkap hubungan non-linier dan interaksi antarfitur yang kompleks [13].

Penerapan algoritma XGBoost dalam klasifikasi tingkat kepadatan kerumunan jemaah haji dan umrah diharapkan mampu menyajikan informasi berbasis data secara akurat dan sistematis. Meskipun menunjukkan kinerja klasifikasi yang baik, hasil ini tetap perlu dikaji secara kritis dengan mempertimbangkan potensi *overfitting* serta keterbatasan generalisasi pada konteks data yang berbeda. Temuan ini tidak hanya merepresentasikan bentuk validasi terhadap pendekatan yang digunakan, tetapi juga dapat menjadi dasar dalam pengembangan sistem pemantauan cerdas yang adaptif guna mendukung pengelolaan ibadah haji dan umrah secara lebih efektif di masa mendatang.

2. METODE PENELITIAN

Implementasi penelitian ini dilakukan melalui beberapa tahapan sistematis, yang dijabarkan secara komprehensif pada Gambar 1.



Gambar 1. Metodologi Penelitian

2.1. Pengumpulan Data

Penelitian ini memanfaatkan dataset sekunder yang diperoleh dari situs Kaggle. Data tersebut terdiri atas 10.000 entri dan mencakup 29 variabel yang merepresentasikan karakteristik jemaah, parameter mobilitas, serta kondisi lingkungan selama penyelenggaraan ibadah haji dan umrah pada tahun 2024. Cuplikan data yang digunakan disajikan pada Tabel 1.

Tabel 1. *Hajj Umrah Crowd Management Dataset*

index	Movement Speed	Interaction Frequency	Time Spent at Location minutes	Temperature	Sound Level dB	Distance Between People meters
0	0,9	6	77	44	82	0,94
1	0,55	8	92	39	80	2,04
2	0,94	2	16	32	84	1,85
3	0,55	2	74	41	79	0,96
4	0,36	10	20	44	67	1,05
...	...	...	...	...	...	...
9995	0,2	7	88	38	72	2,11
9996	0,41	0	96	45	86	1,28
9997	0,45	6	50	41	75	2,17
9998	1,35	1	33	34	84	0,79
9999	0,26	1	67	43	88	1,08

2.2. Praproses Data

Praproses data bertujuan untuk membersihkan dan mempersiapkan data mentah agar layak digunakan dalam pemodelan lebih lanjut. Tahap ini mencakup pengurangan *noise*, memverifikasi nilai yang hilang, serta transformasi data ke dalam format yang sesuai dengan kebutuhan algoritma [14]. Berikut merupakan penjelasan lebih spesifik mengenai tahapan praproses data yang dilakukan dalam penelitian ini.

1. Pengecekan nilai hilang.
   Pada tahap ini, dilakukan verifikasi terhadap nilai-nilai dalam dataset untuk memastikan tidak terdapat data yang kosong (*missing values*) [15]. Verifikasi ini bertujuan untuk menjaga kualitas data dan menghindari potensi kesalahan dalam proses analisis, pelatihan model, maupun interpretasi hasil.

2. Pemilihan fitur.
   Setelah proses pengecekan nilai, tahap selanjutnya adalah pemilihan fitur (*feature selection*). Tujuannya adalah memilih atribut yang paling relevan guna meningkatkan performa dan efisiensi model dengan menghilangkan fitur yang tidak signifikan maupun bersifat redundan [14]. Fitur dipilih berdasarkan faktor kepadatan kerumunan (*crowd density*) yang memiliki pengaruh nyata terhadap dinamika pergerakan manusia, seperti kecepatan pergerakan, jarak antarindividu, frekuensi interaksi, fluktuasi suhu, tren intensitas suara, kondisi cuaca, serta karateristik konvergensi dan divergensi arus.

3. Pelabelan.
   Setelah pemilihan fitur dilakukan, tahap selanjutnya adalah pelabelan (*labeling*) variabel menggunakan pendekatan *rule-based classification* berbasis skoring multivariat [16][17]. Pada tahap ini, setiap fitur yang terpilih diberikan bobot tertentu untuk menghasilkan skor densitas kerumunan. Nilai skor tersebut kemudian digunakan untuk mengklasifikasikan data ke dalam tiga kategori tingkat kepadatan, yaitu rendah, sedang, dan tinggi. Rincian skoring disajikan pada Tabel 2 dan Tabel 3.
   Label ditetapkan berdasarkan skor total, yaitu “Tinggi” untuk skor ≥ 6, “Sedang” untuk skor $3-5$, dan “Rendah” untuk skor < 3.

4. Konversi data kategorikal.
   Tahap berikutnya adalah mengonversi fitur kategorikal ke dalam format numerik biner agar fitur-fitur tersebut dapat diproses oleh algoritma pembelajaran mesin. Transformasi ini diperlukan karena Extreme Gradient Boosting (XGBoost) secara intrinsik hanya memproses masukan berupa data kuantitatif. Sebagai konsekuensi, penelitian ini menerapkan metode One-Hot Encoding karena mampu merepresentasikan setiap kategori secara eksplisit tanpa mengasumsikan adanya tingkatan atau urutan antar kategori, sehingga sesuai untuk data nominal tanpa hierarki nilai [18].

5. Pemisahan data.
   Untuk menjaga konsistensi distribusi label, data dibagi menggunakan teknik stratifikasi dengan proporsi 80% untuk pelatihan dan 20% untuk pengujian [19]. Pemilihan teknik ini bertujuan untuk mempertahankan rasio kelas minoritas pada setiap partisi, sehingga dapat menekan potensi bias terhadap kelas mayoritas [20].

2.3. Upaya Peningkatan Performa Model

Guna meningkatkan sensitivitas model selama proses pelatihan serta memperkuat kemampuan generalisasi model pada tahap pengujian, penelitian ini menerapkan dua strategi sebagai berikut.

1. Penanganan ketidakseimbangan kelas.
   Selain pemisahan data secara stratifikasi, penelitian ini juga menerapkan pembobotan kelas (*class weighting*) untuk menangani disparitas proporsi antarkelas. Pembobotan kelas merupakan metode penyeimbangan data dengan memberikan bobot lebih besar pada kelas minoritas dan bobot lebih kecil pada kelas mayoritas selama proses pelatihan model. Pendekatan umum dalam metode ini ialah *inverse class frequency*, yaitu bobot kelas w_k ditentukan berdasarkan perbandingan antara jumlah total observasi (N) dan jumlah sampel pada kelas ke-k, yang kemudian disesuaikan dengan jumlah kelas (K). Hasil pembobotan kemudian diintegrasikan dengan fungsi kerugian (*loss function*) pada model [21].

2. Optimasi hiperparameter.
   Dalam rangka memperoleh konfigurasi hiperparameter secara efisien, penelitian ini menerapkan Randomized Search CV sebagai metode optimasi. Teknik ini mengeksplorasi kombinasi hiperparameter secara acak dari ruang pencarian yang telah ditentukan sehingga waktu komputasi menjadi lebih terkendali tanpa harus menguji seluruh kemungkinan dengan mengevaluasi setiap kolaborasi parameter [22].

Tabel 2. Indikasi Kepadatan Tinggi

Kondisi	Skor
Movement_speed < 0.6	+1
Distance between people meters < 1.2	+1
Sound level dB > 80	+1
Waiting time for transport > 10	+1
Security checkpoint wait time > 10	+1
Time spent at location minutes > 60	+1
Interaction frequency >= 7	+1
Jenis aktivitas padat: "tawaf" atau "sa'i"	+1
Event terkait kemacetan kerumunan: "*crowd congestion*"	+1

Tabel 3. Indikasi Kepadatan Rendah

Kondisi	Skor
Distance_between_people_meters > 2.5	-1
Movement speed > 1.0	-1
Sound level dB < 70	-1
Event kurang berkaitan dengan kepadatan: "*religious activity*" atau "*transport delay*"	-1

2.4. Pemodelan

Penerapan model XGBoost pada studi kasus klasifikasi multikelas dibangun melalui serangkaian iterasi *boosting* yang berlandaskan pada fungsi kerugian *multiclass softmax logistic* [23]. Pada kasus klasifikasi dengan target non-hierarki, model dilatih menggunakan himpunan data yang memiliki label saling eksklusif, sehingga setiap sampel hanya mewakili satu kelas tertentu secara independent [11]. Algoritma pembelajaran mesin ini mempelajari skor *logit* (*log-odds*) sebagai dasar dalam menentukan probabilitas prediksi. Setiap kelas dipelajari secara terpisah, tetapi dioptimalkan secara simultan. Pada tahap permulaan pelatihan, skor *logit* diinisialisasi berdasarkan nilai *log-prior* dari masing-masing kelas atau menggunakan nilai nol sebagai baseline skor. Selanjutnya, pada setiap iterasi, skor *logit* dikonversi menjadi probabilitas melalui fungsi *softmax* [24]. Berdasarkan hasil konversi tersebut, turunan pertama (*gradient*) dan turunan kedua (*hessian*) dari fungsi kerugian dihitung untuk setiap sampel dan kelas. Nilai-nilai tersebut kemudian diagregasikan pada setiap *region* kandidat daun untuk membentuk sejumlah pohon keputusan berbasis CART (*Classification and Regression Tree*) yang masing-masing merepresentasikan pembaruan skor *logit* [10]. Dengan memanfaatkan pendekatan Taylor orde dua, bobot optimal pada setiap daun diperoleh untuk meminimalkan kesalahan prediksi [25][26]. Sementara itu, kualitas pemisahan *node* dievaluasi melalui *gain* yang memperhitungkan kontribusi *gradient*, *hessian*, serta parameter regularisasi yang berfungsi sebagai penalti guna mengendalikan kompleksitas struktur pohon [13]. Adapun fungsi kerugian murni yang menjadi objek minimasi pada metode ini diformulasikan pada Persamaan 1 [12].

$$\mathcal{L} = -\sum_{i=1}^{n}\sum_{k=1}^{k} 1(yi = k)\, log(p_{ik}) \quad (1)$$

Persamaan 1 merepresentasikan nilai rata-rata kesalahan prediksi. Secara konseptual, persamaan tersebut mengukur deviasi antara probabilitas prediksi yang dihasilkan melalui fungsi *softmax* (p_{ik}) dan label sebenarnya (yi). Di sisi lain, notasi yang memvisualisasikan fungsi kerugian, parameter regularisasi serta kontribusi bobot daun dapat dinyatakan sebagai fungsi objektif, ditunjukkan pada Persamaan 2 [27].

$$\mathcal{O}bj^{(t)} = \sum_{i=1}^{n}\sum_{k=1}^{K} l(yi, \hat{y}i_k^{(t-1)} + f_{t,k}(x_i)) + \Omega(f_t) \quad (2)$$

Menguraikan Persamaan 2, fungsi kerugian diekspresikan melalui notasi $l(yi, \hat{y}i)$ secara format sederhana, sedangkan $f_{t,k}(x_i)$ digunakan untuk merepresentasikan skor *logit* yang dihasilkan oleh pohon keputusan pada iterasi ke-t, yang berkontribusi terhadap pembaruan prediksi kelas ke-k untuk sampel ke-i. Sementara itu, notasi $\Omega(f_t)$ merupakan fungsi regularisasi yang memasukkan parameter lambda (λ) dan gamma (γ) untuk memberikan penalti terhadap jumlah daun dan besarnya bobot daun pada pohon keputusan yang dibangun. Fungsi objektif merupakan komponen fundamental yang mendasari mekanisme pembelajaran pada algoritma ini. Secara umum, fungsi objektif berperan menentukan arah pembaruan model pada setiap iterasi agar hasil prediksi semakin mendekati nilai sebenarnya, sekaligus mengontrol kompleksitas model agar tidak terjadi *overfitting* [10].

2.5. Prediksi dan Evaluasi

Setelah proses pemodelan, tahap selanjutnya adalah melakukan generalisasi terhadap data uji guna mengevaluasi kinerja model [28]. Evaluasi dilakukan menggunakan *confusion matrix*, yang menjadi dasar perhitungan sejumlah metrik evaluasi seperti akurasi, presisi, *recall*, dan *f1-score*. Nilai-nilai metrik tersebut dihitung menggunakan rumus sebagaimana ditunjukkan pada Persamaan 3−6 [29].

$$\text{Akurasi} = \frac{TP + TN}{\text{TP + TN + FP + FN}} \quad (3)$$

$$\text{Presisi} = \frac{\text{TP}}{\text{TP + FP}} \quad (4)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP + FN}} \quad (5)$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} = \frac{2TP}{2TP + FP + FN} \quad (6)$$

Penelitian ini juga menerapkan Stratified K-Fold Cross Validation untuk memvalidasi konsistensi performa model pada setiap pembagian data. Dalam metode ini, dataset dipisahkan menjadi beberapa *k-fold* (lipatan) dengan tetap menjaga proporsi label pada setiap subset [30]. Untuk memperkuat validitas empiris terhadap XGBoost secara kritis dan berimbang, peneliti melakukan perbandingan dengan berbagai pendekatan pembelajaran mesin, mencakup *baseline* klasik (*non-ensemble*), *tree-based bagging*, *margin-based classifier* serta *gradient boosting* generasi lanjutan, dengan menerapkan prosedur optimasi performa yang seragam serta cakupan kofigurasi parameter secara ekstensif yang disesuaikan pada masing-masing metode. Perbandingan dilakukan menggunakan metrik akurasi dan makro *f1-score* [19][22].

3. HASIL DAN PEMBAHASAN

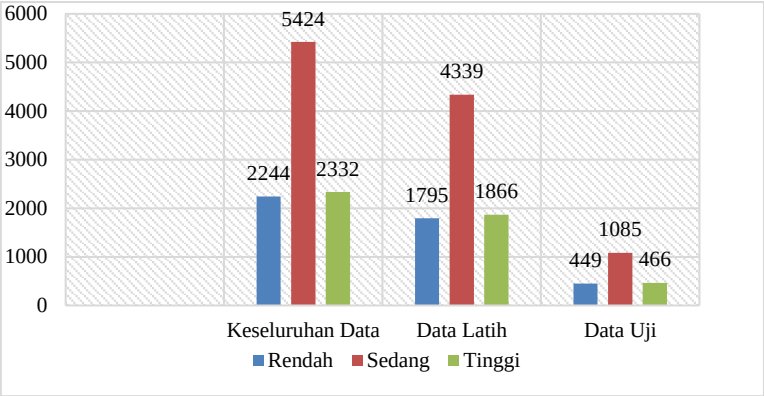
3.1. Praproses Data

Bagian ini memaparkan hasil pengecekan nilai, distribusi label berbasis *rule-based classification*, serta pemilihan fitur sebagai tahapan pra-pemrosesan data dalam penelitian ini.

1. Pengecekan nilai hilang.
   Hasil pengecekan mengonfirmasi bahwa dataset tidak memiliki nilai hilang, yang menandakan keutuhan observasi di setiap variabel serta mengurangi kebutuhan akan proses imputasi.

2. Distribusi label.
   Gambar 2 menunjukkan kelas pada dataset dengan sedang adalah kelas paling dominan. Distribusi ini tetap konsisten setelah pembagian data, yang membuktikan bahwa teknik stratifikasi berhasil mempertahankan proporsi label secara seimbang.

3. Pemilihan fitur.
   Berdasarkan target penelitian, peneliti menetapkan 18 fitur yang relevan dan mengeliminasi beberapa fitur redundan dari proses konversi. Daftar fitur tersebut disajikan pada Tabel 4.



Gambar 2. Hasil Pelabelan

Tabel 4. Fitur Model

Fitur	Tipe Data	Fitur	Tipe Data
Distance between people meters	Float	Activity type tawaf	Binary
Movement speed	Float	Activity type sa'i	Binary
Time spent at location minutes	Integer	Activity type walking	Binary
Interaction frequency	Integer	Activity type resting	Binary
Sound level dB	Integer	Weather condition rainy	Binary
Queue time minutes	Integer	Weather condition cloudy	Binary
Security checkpoint wait time	Integer	Event type medical emergency	Binary
Waiting time for Transport	Integer	Event type religious activity	Binary
Temperature	Integer	Event type transport delay	Binary

3.2. Hasil Pemodelan

Komponen-komponen pada bagian ini meliputi parameter optimal, *confusion matrix*, *classification report*, kurva pembelajaran, *stratified k-fold cross-validation*, serta perbandingan model. Berikut adalah penjelasan dan pemaparan hasil tersebut.

1. Parameter model.

   Berdasarkan karakteristik dan kompleksitas dataset, kombinasi parameter yang diperoleh melalui penerapan Randomized Search CV ditunjukkan pada Tabel 5.

Tabel 5. Parameter Optimal

Hyperparameter	Value
Subset sample	0.8
Regularisasi L2 (*reg_lambda*)	5.0
Regularisasi L1 (*reg_alpha*)	1.0
Number of estimators	500
Maximal depth	5
Gamma (γ)	0.3
Minimal child weight	7
Learning rate (η)	0.2
Column sample by tree	1.0

2. *Confusion matrix.*

   Berikut merupakan hasil prediksi model terhadap data uji yang disajikan dalam Tabel 6.

Tabel 6. *Confusion Matrix*

	Prediksi Rendah	Prediksi Sedang	Prediksi Tinggi
Aktual Rendah	442	7	0
Aktual Sedang	4	1070	11
Aktual Tinggi	0	4	462

Tabel 6 menggambarkan sebagian besar data tersegmentasi dengan tepat. Kesalahan generalisasi relatif kecil, seperti 4 label tinggi, 7 label rendah, dan 15 label sedang. Kesalahan prediksi tertinggi terletak pada kelas sedang, yang menegaskan bahwa penerapan teknik stratifikasi dan pembobotan kelas efektif dalam menekan bias terhadap kelas mayoritas, sehingga model tidak memberikan prioritas berlebih pada kelas sedang.

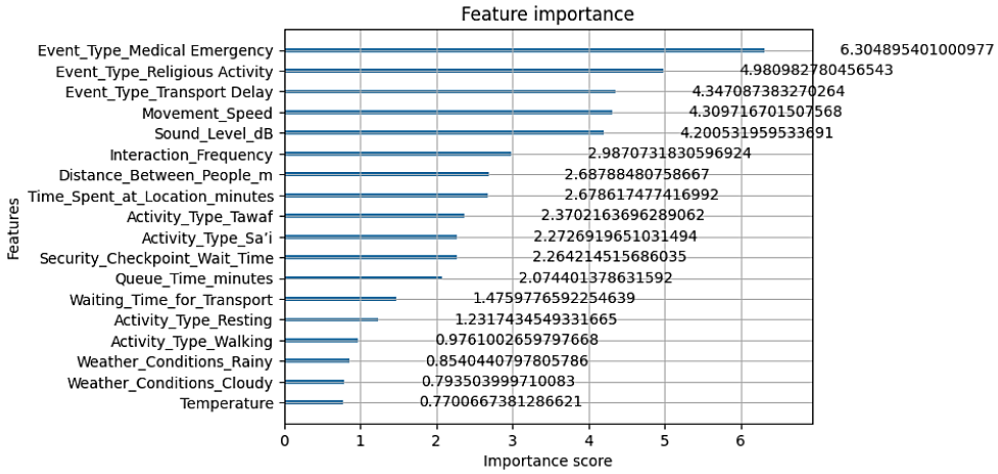
3. *Classification report.*

Tabel 7. *Classification Report*

	precision	*recall*	*f1-score*	*Support*
Rendah	0.99	0.98	0.99	449
Sedang	0.99	0.99	0.99	1085
Tinggi	0.98	0.99	0.98	466
Akurasi			0.99	2000
Macro Avg	0.99	0.99	0.99	2000
Weighted Avg	0.99	0.99	0.99	2000

Tabel 7 mendeskripsikan performa model dengan presisi, *recall*, dan *f1-score* di atas 98% pada seluruh kelas serta akurasi total mencapai 99% (0.987). Nilai *macro* dan *weighted average* yang stabil semakin menguatkan hasil tersebut.

4. Fitur terbaik model.

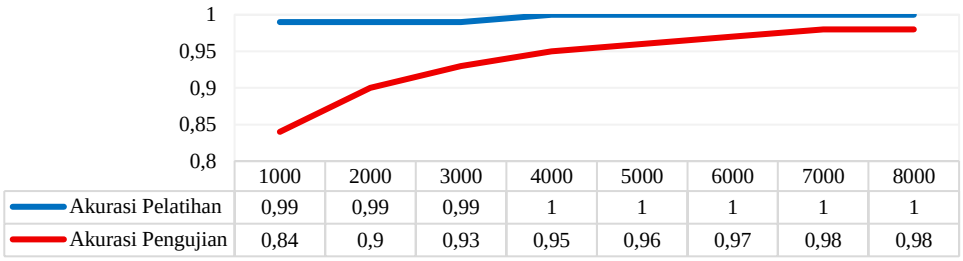


Gambar 3. Fitur Terbaik Model

Penentuan fitur terbaik dalam model pada Gambar 3 dilakukan menggunakan nilai total *gain*, yaitu total peningkatan nilai fungsi objektif yang dihasilkan ketika suatu fitur digunakan sebagai kriteria pemisahan *node* pada seluruh pohon keputusan untuk semua iterasi. Semakin besar nilai *gain* yang diperoleh suatu fitur, semakin besar kontribusinya dalam menurunkan nilai fungsi kerugian secara kumulatif. Dengan demikian, fitur yang memiliki nilai total *gain* tertinggi dianggap sebagai fitur paling signifikan dalam menjelaskan variabilitas target, karena berperan dominan dalam proses pembentukan struktur pohon secara keseluruhan. Berdasarkan hal tersebut, fitur dengan kontribusi terbesar meliputi *event type medical emergency* dengan total *gain* mencapai 6.304895, diikuti oleh *event type religious activity* (4.980983), *event type transport delay* (4.347087), *movement speed* (4.309717), dan *sound level dB* (4.200532), sedangkan fitur lingkungan seperti *temperature* dan *weather conditions* memberikan pengaruh yang relatif lebih kecil.

5. Kurva pembelajaran model.
Kurva pembelajaran (*learning curve*) memvisualisasikan perkembangan akurasi model seiring bertambahnya data pelatihan. Grafik ini berguna untuk mendeteksi *overfitting* atau *underfitting*, sekaligus menilai sejauh mana penambahan data dapat meningkatkan kinerja model. Gambar 4 memperlihatkan pola konvergensi, dengan akurasi pelatihan mencapai 100% dan akurasi pengujian sebesar 98,7%, menghasilkan selisih sekitar 1,3%.

6. *Stratified k-fold cross validation.*
Dalam penelitian ini, *stratified k-fold cross-validation* diterapkan menggunakan dua skema, yaitu 5-*fold* dan 10-*fold*, sebagaimana disajikan pada Gambar 5. Validasi dengan dua skema menunjukkan konsistensi performa model pada berbagai pembagian data. Stabilitas ini tercermin dari rata-rata akurasi pada kedua skema yang konsisten berada di kisaran 99%, dengan standar deviasi yang sangat kecil.



	1000	2000	3000	4000	5000	6000	7000	8000
Akurasi Pelatihan	0,99	0,99	0,99	1	1	1	1	1
Akurasi Pengujian	0,84	0,9	0,93	0,95	0,96	0,97	0,98	0,98

Gambar 4. Kurva Pembelajaran

```
5-Fold CV - Mean Accuracy: 0.9904 Std: 0.0014966629547095735
10-Fold CV - Mean Accuracy: 0.9920999999999998 Std: 0.003207802986469091
5-Fold CV - Mean F1 Macro: 0.9899837257843188 Std: 0.0015738668733032753
10-Fold CV - Mean F1 Macro: 0.9917493932339182 Std: 0.0033464203402738825

5-Fold Accuracy per fold:
Fold 1: 0.988
Fold 2: 0.992
Fold 3: 0.991
Fold 4: 0.992
Fold 5: 0.991

10-Fold Accuracy per fold:
Fold 1: 0.994
Fold 2: 0.990
Fold 3: 0.990
Fold 4: 0.993
Fold 5: 0.987
Fold 6: 0.998
Fold 7: 0.988
Fold 8: 0.994
Fold 9: 0.992
Fold 10: 0.995
```

Gambar 5. *Cross Validation*

7. Perbandingan model.

```
                    EVALUATION METRICS
======================================================================
Logistic Regression  | Accuracy: 0.7935 | F1-Macro: 0.7945
SVC                  | Accuracy: 0.8410 | F1-Macro: 0.8396
Random Forest        | Accuracy: 0.8835 | F1-Macro: 0.8784
LightGBM             | Accuracy: 0.9845 | F1-Macro: 0.9840
XGBoost              | Accuracy: 0.9870 | F1-Macro: 0.9866
======================================================================
```

Gambar 6. Perbandingan Model

Hasil pengujian Gambar 6 menunjukkan bahwa performa model bervariasi sejalan dengan kapasitas dan mekanisme pembelajaran masing-masing algoritma pada studi kasus klasifikasi multikelas kepadatan kerumunan. Regresi logistik multinomial memiliki keterbatasan dalam memodelkan hubungan non-linear sehingga menghasilkan kinerja yang relatif lebih rendah, berdasarkan rasio performa berada pada rentang 9-19,35%. Support Vector Classification (SVC) dengan kernel *radial basis function* (RBF), yang memiliki kapasitas non-linearitas dan didukung proses standarisasi fitur menunjukkan performa di bawah pendekatan *ensemble* pada kedua metrik evaluasi sebesar 4%. Random Forest mampu meningkatkan kinerja melalui reduksi varians menggunakan pendekatan *bagging*, namun masih menunjukkan kesenjangan performa yang signifikan dengan selisih kinerja berada pada kisaran 10-10,82%. Sejalan dengan karakteristik tersebut, metode *gradient boosting*, khususnya LightGBM dan XGBoost, menghasilkan performa tertinggi, dengan perbedaan yang bersifat marginal.

4. KESIMPULAN

Algoritma Extreme Gradient Boosting (XGBoost) menunjukkan kinerja klasifikasi yang sangat unggul dalam memprediksi tingkat kepadatan kerumunan jemaah haji dan umrah, dengan capaian akurasi sebesar 98,7% serta nilai presisi, recall, dan *f1-score* yang konstan berada pada rentang 98–99%. Kinerja optimal tersebut diperoleh melalui proses optimasi hiperparameter menggunakan Randomized Search CV, yang dikombinasikan dengan penerapan pembobotan kelas (*class weighting*) guna meningkatkan sensitivitas

model terhadap masing-masing kelas target serta mengatasi permasalahan ketidakseimbangan distribusi kelas. Pendekatan ini terbukti mampu menekan kecenderungan model terhadap dominasi label mayoritas dalam proses pembelajaran, sehingga berdampak positif terhadap peningkatan kinerja prediksi. Keandalan model selanjutnya dikonfirmasi melalui penerapan Stratified K-Fold Cross-Validation, yang menghasilkan rata-rata akurasi sekitar 99%, sehingga mencerminkan konsistensi kinerja model tanpa indikasi bias yang relatif tinggi. Hal ini menegaskan bahwa pola non-linear dan interaksi antarfitur pada target kepadatan kerumunan dalam konteks ibadah massal dapat digeneralisasikan dengan baik oleh teknik *ensemble*, sebagaimana diimplementasikan pada pendekatan *gradient boosting*. Penelitian selanjutnya dapat mempertimbangkan eksplorasi lebih lanjut berbagai teknik penyeimbangan dan optimasi hiperparameter, seperti teknik *synthetic oversampling* (misalnya SMOTE dan turunanya), fungsi kerugian adaptif seperti *focal loss*, pendekatan hibrid, *baesiyan optimization*, serta evaluasi performa XGBoost dengan uji signifikansi statistik.

REFERENCES

[1] N. Syamsiyah, "Strategi Pemasaran Produk Umrah pada Travel Smarts Umrah Lampung," *Multazam : Jurnal Manajemen Haji dan Umrah*, vol. 3, no. 1, p. 1, Jun. 2023, doi: 10.32332/multazam.v3i1.5399.

[2] N. R. Aisy and Muzakki, "Pendampingan Manasik Haji Untuk Mengembangkan Nilai Agama Dan Moral Di RA Ar-Raudhah Desa Hampalit Corresponding Author," *Jurnal Pengabdian Masyarakat*, vol. 1, no. 4, pp. 199–204, Feb. 2024, doi: 10.59837/y47chb79.

[3] M. A. Yaqin and A. Tholib, "Implementasi Manasik Haji dengan Teknologi Vr (Virtual Reality) untuk Mencegah Penyebaran Virus Covid-19," *Dinamika Informatika*, vol. 14, no. 2, 2022, doi: 10.35315/informatika.v14i2.9185.

[4] A. Chulaivi, I. Z. Khazimi, M. Ikram, and A. Hafiz, "Efektivitas Manajemen Haji dan Umrah dalam Meningkatkan Kepuasan Jamaah di Indonesia: Sebuah Studi Literatur," *LANCAH: Jurnal Inovasi dan Tren*, vol. 2, no. 2b, Jul. 2024, doi: 10.35870/ljit.v2i2b.2957.

[5] R. Bhuiyan *et al.*, "Deep Dilated Convolutional Neural Network for Crowd Density Image Classification with Dataset Augmentation for Hajj Pilgrimage," *Sensors*, vol. 22, no. 14, Jul. 2022, doi: 10.3390/s22145102.

[6] W. Halboob, H. Altaheri, A. Derhab, and J. Almuhtadi, "Crowd Management Intelligence Framework: Umrah Use Case," *IEEE Access*, vol. 12, pp. 6752–6767, 2024, doi: 10.1109/ACCESS.2024.3350188.

[7] F. E. Permana and M. Hafil, "Lebih dari 18,5 Juta Muslim Tunaikan Haji dan Umroh di 2024," Republika. [Online]. Available: https://khazanah.republika.co.id/berita/sqh3xe430/lebih-dari-185-juta-muslim-tunaikan-haji-dan-umroh-di-2024.[22 Januari 2025]

[8] M. Dava Maulana *et al.*, "Algoritma XGboost Untuk Klasifikasi Kualitas Air Minum," Cimahi, 2023. doi: 10.36040/jati.v7i5.7308.

[9] T. T. H. Le, Y. E. Oktian, and H. Kim, "XGBoost for Imbalanced Multiclass Classification-Based Industrial Internet of Things Intrusion Detection Systems," *Sustainability (Switzerland)*, vol. 14, no. 14, Jul. 2022, doi: 10.3390/su14148707.

[10] A. Mehdary, A. Chehri, A. Jakimi, and R. Saadane, "Hyperparameter Optimization with Genetic Algorithms and XGBoost: A Step Forward in Smart Grid Fraud Detection," *Sensors*, vol. 24, no. 4, Feb. 2024, doi: 10.3390/s24041230.

[11] H. Joe and H. G. Kim, "Multi-label classification with XGBoost for metabolic pathway prediction," *BMC Bioinformatics*, vol. 25, no. 1, pp. 1–15, Dec. 2024, doi: 10.1186/s12859-024-05666-0.

[12] J. Abualdenien and A. Borrmann, "Ensemble-learning approach for the classification of Levels Of Geometry (LOG) of building elements," *Advanced Engineering Informatics*, vol. 51, pp. 1–15, Jan. 2022, doi: 10.1016/j.aei.2021.101497.

[13] W. Su *et al.*, "An XGBoost-Based Knowledge Tracing Model," *International Journal of Computational Intelligence Systems*, vol. 16, no. 1, Dec. 2023, doi: 10.1007/s44196-023-00192-y.

[14] E. Sahelvi, P. Cikita, R. Mela Sapitri, and L. Efrizoni, "Perbandingan Algoritma K-Nearest Neighbors dan Random Forest untuk Rekomendasi Gaya Hidup Sehat dalam Mencegah Penyakit Jantung," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 5, no. 5, pp. 830–840, Jul. 2025, doi: 10.57152/malcom.v5i3.1972.

[15] A. Tarisa Akbar, N. Yudistira, and A. Ridok, "Identifikasi Gagal Ginjal Kronis dengan Mengimplementasikan Metode Support Vector Machine beserta K-Nearest Neighbour (Svm-Knn)," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, vol. 10, no. 2, pp. 301–308, Apr. 2023, doi: 10.25126/jtiik.2023106059.

[16] Y. Firmansyah, E. Ernawati, and E. L. Prawiro, "Sistem Skoring untuk Memprediksi Kejadian Hipertensi pada Usia Produktif Di Kota Medan (Preliminary Study)," *Jurnal Muara Sains, Teknologi, Kedokteran dan Ilmu Kesehatan*, vol. 4, no. 1, p. 55, May 2020, doi: 10.24912/jmstkik.v4i1.6013.

[17] S. K. M. Hossain, S. A. Ema, and H. Sohn, "Rule-Based Classification Based on Ant Colony Optimization: A Comprehensive Review," 2022, *Hindawi Limited*. doi: 10.1155/2022/2232000.

[18] C. Herdian, A. Kamila, and I. G. Agung Musa Budidarma, "Studi Kasus Feature Engineering Untuk Data Teks: Perbandingan Label Encoding dan One-Hot Encoding Pada Metode Linear Regresi," *Technologia : Jurnal Ilmiah*, vol. 15, no. 1, p. 93, Jan. 2024, doi: 10.31602/tji.v15i1.13457.

[19] A. Ramadhani, N. Safaat Harahap, S. Agustian, I. Iskandar, and S. Sanjaya, "Perbandingan Performa Metode Klasifikasi Teks Multilabel Hadis Terjemahan Bukhari Menggunakan Support Vector Machine dan Long Short Term Memory," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 5, no. 3, pp. 896–907, Jul. 2025, doi: 10.57152/malcom.v5i3.2051.

[20] I. Ivanov, B. Toleva, and V. J. Hooper, "A fast and effective approach for classification medical data sets," *International Journal of Innovative Research and Scientific Studies*, vol. 6, no. 3, pp. 545–552, Apr. 2023, doi: 10.53894/ijirss.v6i3.1580.

[21] O. El Gannour, S. Hamida, Y. Lamalem, M. A. Mahjoubi, B. Cherradi, and A. Raihani, "Improving skin diseases prediction through data balancing via classes weighting and transfer learning," *Bulletin of Electrical Engineering and Informatics*, vol. 13, no. 1, pp. 628–637, Feb. 2024, doi: 10.11591/eei.v13i1.5999.

[22] A. D. Rachmatsyah, T. Sugihartono, and K. Irfan, "Perbandingan Teknik Optimasi Grid Search dan Randomized Search dalam Meningkatkan Akurasi Metode Klasifikasi SVM Pada Sentimen Ulasan Pengguna Aplikasi JKN Mobile," *SKANIKA: Sistem Komputer dan Teknik Informatika*, vol. 8, no. 1, pp. 13–22, Jan. 2025, doi: 10.36080/skanika.v8i1.3328.

[23] F. Baldo, J. Grando, Y. Yamada Correa, and D. Amorim Policarpo, "Adaptive Fast XGBoost for Multiclass Classification," *Journal of Information and Data Management*, vol. 14, no. 1, Oct. 2023, doi: 10.5753/jidm.2023.3150.

[24] S. Zhou, C. Chen, G. Han, and X. Hou, "Double additive margin softmax loss for face recognition," *Applied Sciences (Switzerland)*, vol. 10, no. 1, Jan. 2020, doi: 10.3390/app10010060.

[25] X. Li *et al.*, "Probabilistic solar irradiance forecasting based on XGBoost," *Energy Reports*, vol. 8, pp. 1087–1095, Aug. 2022, doi: 10.1016/j.egyr.2022.02.251.

[26] K. Shaheed, Q. Abbas, A. Hussain, and I. Qureshi, "Optimized Xception Learning Model and XgBoost Classifier for Detection of Multiclass Chest Disease from X-ray Images," *Diagnostics*, vol. 13, no. 15, pp. 1–25, Aug. 2023, doi: 10.3390/diagnostics13152583.

[27] P. Zhang, Y. Jia, and Y. Shang, "Research and application of XGBoost in imbalanced data," *Int J Distrib Sens Netw*, vol. 18, no. 6, Jun. 2022, doi: 10.1177/15501329221106935.

[28] M. Valavan and S. Rita, "Predictive-Analysis-based Machine Learning Model for Fraud Detection with Boosting Classifiers," *Computer Systems Science and Engineering*, vol. 45, no. 1, pp. 231–245, 2023, doi: 10.32604/csse.2023.026508.

[29] K. Riehl, M. Neunteufel, and M. Hemberg, "Hierarchical confusion matrix for classification performance evaluation," *J R Stat Soc Ser C Appl Stat*, vol. 72, no. 5, pp. 1394–1412, Nov. 2023, doi: 10.1093/jrsssc/qlad057.

[30] S. Widodo, H. Brawijaya, and S. Samudi, "Stratified K-fold cross validation optimization on machine learning for prediction," *Sinkron*, vol. 7, no. 4, pp. 2407–2414, Oct. 2022, doi: 10.33395/sinkron.v7i4.11792.