SAINTEKES: JURNAL SAINS, TEKNOLOGI DAN KESEHATAN
LPPM Institut Teknologi Dan Kesehatan Aspirasi
E-ISSN: 2963-0959

PENERAPAN ALGORITMA XGBOOST DALAM PENENTUAN POTENSI SEKTOR WISATA LOKAL DI LOMBOK TIMUR

Wenti Ayu Wahyuni[1*], Muh. Nasirudin Karim[2], Zumratul Muahidin[3]
[1,2,3]Universitas Teknologi Mataram, Mataram, Indonesia

Article Information

Article history:

Received July 30, 2025
Approved July 31, 2025

Keywords:

XGBoost, tourism potential, classification, machine learning, East Lombok

ABSTRACT

East Lombok holds significant tourism potential, yet its development faces challenges due to the lack of data-driven spatial analysis. This study applies the XGBoost algorithm to predict and classify tourism potential based on features such as visitor numbers, facilities, accessibility, digital exposure, and sociocultural indicators. The results show that XGBoost achieves a classification accuracy of 92%, with review ratings and digital activity being the most influential factors. Spatial visualization of the prediction highlights regions with high, medium, and low potential. This approach offers a data-driven tool to support more targeted and effective tourism development strategies.

ABSTRAK

Lombok Timur memiliki potensi besar di sektor pariwisata, namun pengembangannya masih terkendala kurangnya pemetaan wilayah berbasis data. Penelitian ini menerapkan algoritma XGBoost untuk memprediksi dan mengklasifikasikan potensi wisata berdasarkan data kunjungan, fasilitas, aksesibilitas, eksposur digital, dan faktor sosial budaya. Hasil menunjukkan bahwa model XGBoost mampu mengklasifikasikan wilayah dengan akurasi mencapai 92%, dengan rating ulasan dan aktivitas digital sebagai fitur paling berpengaruh. Visualisasi spasial hasil klasifikasi memberikan gambaran distribusi wilayah dengan potensi tinggi, sedang, dan rendah. Pendekatan ini dapat menjadi alat pendukung kebijakan pengembangan pariwisata yang lebih efektif dan berbasis bukti.

© 2025 SAINTEKES

Corresponding author email: wentiayu443322@gmail.com

PENDAHULUAN

Lombok Timur merupakan salah satu kabupaten di Provinsi Nusa Tenggara Barat yang memiliki potensi besar dalam sektor pariwisata, baik dari sisi keindahan alam, budaya lokal, maupun keanekaragaman atraksi wisata yang dimilikinya. Namun,

pengembangan sektor ini masih menghadapi berbagai tantangan, seperti ketimpangan informasi potensi wilayah, keterbatasan dalam analisis prediktif, serta kurang optimalnya perencanaan berbasis data. Dalam konteks pembangunan daerah, kemampuan untuk mengidentifikasi dan menentukan wilayah-wilayah yang memiliki potensi wisata tinggi secara akurat dan objektif menjadi kunci utama dalam menyusun strategi pengembangan yang efektif.

Seiring berkembangnya teknologi digital, pendekatan konvensional dalam penilaian potensi wisata dapat ditingkatkan melalui pemanfaatan kecerdasan buatan, khususnya algoritma pembelajaran mesin (machine learning). Salah satu algoritma yang telah terbukti memiliki performa tinggi dalam klasifikasi dan prediksi data kompleks adalah Extreme Gradient Boosting (XGBoost) (Shao et al., 2024). Algoritma ini menggabungkan keunggulan teknik ensemble learning dan boosting untuk menghasilkan model prediktif yang akurat dan efisien, bahkan dalam kondisi data tidak seimbang atau memiliki noise tinggi.

Penerapan XGBoost dalam sektor pariwisata telah mulai diteliti di beberapa wilayah Indonesia. Misalnya, menerapkan XGBoost untuk memprediksi jumlah kedatangan wisatawan mancanegara berdasarkan data pintu masuk udara, dengan hasil akurasi yang sangat baik dalam memetakan tren kunjungan (Prayuda dan Pratama, 2024). Penelitian lain (Lubis, 2025) menggunakan XGBoost dalam klasifikasi perilaku konsumen wisatawan untuk meningkatkan layanan wisata berbasis ulasan penggun. Selain itu, (Singgalen, 2025) menggunakan XGBoost dalam analisis prediktif tren kamar hotel berbasis big data di sektor hospitality Indonesia.

Lebih lanjut, (Rantung dan Tambajong, 2024) menunjukkan bagaimana algoritma ini mampu menyaring atribut penting dari berbagai lokasi wisata untuk membuat sistem rekomendasi atraksi yang lebih personal. Pendekatan ini membuka peluang besar bagi wilayah seperti Lombok Timur yang belum sepenuhnya dieksplorasi potensi wisatanya secara kuantitatif.

Berlandaskan pada urgensi tersebut, penelitian ini bertujuan untuk menerapkan algoritma XGBoost dalam penentuan potensi sektor wisata di Lombok Timur. Model ini akan dibangun dengan menggabungkan berbagai fitur seperti jumlah kunjungan, fasilitas pendukung, aksesibilitas, nilai budaya, dan kepadatan penduduk lokal. Melalui pendekatan ini, diharapkan tercipta sistem prediksi yang mampu membantu pemerintah daerah, pelaku usaha, dan pemangku kepentingan lainnya dalam menyusun kebijakan pengembangan wisata yang lebih tepat sasaran, efisien, dan berkelanjutan.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif prediktif dengan metode *data-driven modeling*, bertujuan untuk membangun sistem klasifikasi potensi wisata lokal di wilayah Kabupaten Lombok Timur menggunakan algoritma XGBoost (Extreme Gradient Boosting).

1. Jenis Penelitian

Penelitian ini bersifat kuantitatif dan menggunakan pendekatan eksploratif-prediktif berbasis algoritma machine learning. Tujuannya adalah memodelkan dan mengklasifikasikan wilayah di Lombok Timur berdasarkan potensi sektor wisatanya.

2. Lokasi dan Objek Penelitian

Lokasi penelitian adalah seluruh kecamatan di Kabupaten Lombok Timur. Objek yang dianalisis berupa data multivariat terkait faktor-faktor yang memengaruhi potensi wisata, seperti jumlah kunjungan, aksesibilitas, eksistensi digital, dan fasilitas pendukung.

3. Sumber dan Teknik Pengumpulan Data

Data diperoleh dari beberapa sumber berikut:

Tabel 1. Sumber Data

Jenis Data	Sumber	Bentuk Data
Jumlah kunjungan	Dinas Pariwisata Lombok Timur	Data tahunan per lokasi wisata
Rating & ulasan	Google Maps, Google Reviews	Skor (1–5), jumlah ulasan
Aktivitas digital	Instagram, TikTok	Jumlah tagar dan unggahan
Fasilitas pendukung	Observasi lapangan, GIS	Akomodasi, toilet, parkir, jalan, dll.
Aksesibilitas & jarak	Data spasial (GIS)	Jarak ke jalan utama dan pusat kota

4. Variabel Penelitian

Variabel yang digunakan diklasifikasikan sebagai fitur input untuk model, yaitu:

Tabel 2. Fitur Inputan

Variabel	Tipe Data	Keterangan
Jumlah kunjungan wisatawan	Numerik	Data historis tahunan
Rating lokasi wisata	Numerik	Rata-rata rating Google Reviews
Jumlah ulasan Google	Numerik	Jumlah responden
Jumlah unggahan media sosial	Numerik	Postingan publik dengan tag lokasi
Jumlah fasilitas pendukung	Numerik	Skor akomodasi, toilet, tempat parkir, dll.
Jarak ke jalan utama/pusat kota	Numerik	Dalam kilometer
Jenis wisata (alam, budaya, dll.)	Kategorikal	One-hot encoding untuk klasifikasi

Target output dari model ini adalah kelas potensi wisata:

- Potensi Tinggi
- Potensi Sedang
- Potensi Rendah

5. Tahapan Analisis

Penelitian ini dilakukan melalui beberapa tahapan sistematis untuk memastikan keakuratan model dan relevansi hasil. Tahap pertama adalah preprocessing data, yang mencakup pembersihan data dari duplikasi dan nilai kosong (null), normalisasi terhadap fitur numerik agar memiliki skala yang seragam, serta encoding terhadap variabel kategorikal agar dapat diproses oleh algoritma. Selain itu, dilakukan pemilihan fitur menggunakan metode *feature importance* dari model XGBoost guna menyaring variabel yang paling berpengaruh terhadap hasil klasifikasi. Pemilihan fitur penting dilakukan untuk memastikan hanya variabel yang paling relevan digunakan dalam pelatihan model (Junaidi, 2025).

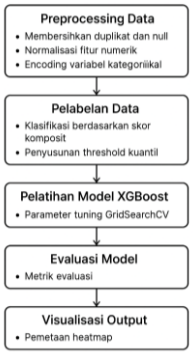
Tahap berikutnya adalah pelabelan data, di mana setiap wilayah atau lokasi wisata diklasifikasikan ke dalam tiga kategori potensi—tinggi, sedang, dan rendah—berdasarkan skor komposit. Skor ini dihitung dari gabungan indikator seperti rating ulasan, jumlah kunjungan wisatawan, dan eksposur digital melalui media sosial. Teknik kuantil digunakan sebagai pendekatan pemotongan nilai, yaitu 30% wilayah teratas dimasukkan ke dalam kategori potensi tinggi, 30–70% sebagai potensi sedang, dan 30% terbawah sebagai potensi rendah.

Setelah data siap, dilakukan pelatihan model XGBoost menggunakan algoritma XGBClassifier dari pustaka xgboost. XGBoost digunakan karena memiliki kemampuan dalam menangani data tidak seimbang, kecepatan proses, dan akurasi yang tinggi (Liu et al., 2019). Model diuji dengan beberapa parameter melalui teknik *GridSearchCV* untuk menemukan kombinasi terbaik, di antaranya max_depth (3, 5, 7), learning_rate (0.01, 0.1, 0.2), dan n_estimators (100, 200). Untuk meningkatkan generalisasi model dan mencegah overfitting, digunakan metode cross-validation sebanyak 5 lipatan (5-fold).

Tahap selanjutnya adalah evaluasi model, yang dilakukan dengan menggunakan berbagai metrik evaluasi seperti accuracy, precision, recall, F1-score, dan confusion matrix untuk

menilai kinerja klasifikasi (Cesarini, Terraglia, & Antolini, 2025). Selain itu, ditampilkan juga grafik feature importance untuk mengidentifikasi variabel-variabel yang paling menentukan dalam prediksi potensi wisata (Ma et al., 2024).

Tahap akhir adalah visualisasi output, di mana hasil klasifikasi wilayah wisata divisualisasikan dalam bentuk peta tematik menggunakan perangkat lunak QGIS atau heatmap dari Matplotlib. Visualisasi ini memudahkan interpretasi spasial dan mendukung analisis strategis terkait perencanaan pengembangan sektor wisata di Lombok Timur (Du et al., 2023).



Gamabar 1. Tahap Analisis

6. Tools dan Perangkat Lunak
   - Python (Jupyter Notebook) Library: pandas, numpy, xgboost, scikit-learn, matplotlib, seaborn.
   - QGIS atau Google Earth Engine (jika butuh visualisasi spasial)
   - Google Colab (jika tidak menggunakan perangkat lokal)

HASIL DAN PEMBAHASAN

1. Dataset Lokasi Wisata di Lombok Timur

Penelitian ini menggunakan 50 lokasi wisata unggulan di Lombok Timur. Masing-masing lokasi memiliki karakteristik berbeda dari sisi jumlah pengunjung, rating ulasan, fasilitas, serta eksposur digital di media sosial. Data diolah untuk membentuk fitur input bagi model XGBoost.

Tabel 3. Dataset

Lokasi Wisata	Kunjungan (ribu)	Rating (1–5)	Ulasan (jumlah)	Postingan Medsos	Fasilitas Skor (0–10)	Jarak ke Kota (km)	Jenis Wisata
Pantai Pink	45	4.6	850	3,500	9	46	Alam
Tanjung Bloam	20	4.2	220	1,100	6	45	Alam
Air Terjun Jeruk Manis	38	4.5	480	1,800	7	22	Alam
Tetebatu	52	4.8	1,300	4,700	10	26	Budaya/Alam
Gili Kondo	15	4.1	180	700	5	40	Bahari
..							
Pantai Kura-Kura	55	4.8	943	4230	5	41	Alam

Data ini selanjutnya dinormalisasi dan dikodekan agar dapat diproses oleh model. Fitur kategorikal seperti “Jenis Wisata” diubah ke bentuk numerik melalui teknik one-hot encoding.

2. Hasil Prediksi XGBoost

Model XGBoost dilatih dengan parameter optimal (max_depth=3, learning_rate=0.1, n_estimators=100) dan menggunakan validasi silang 5-fold. Hasil prediksi klasifikasi potensi wisata lokasi disajikan berikut:

Tabel 4. Hasil Prediksi

Lokasi Wisata	Kelas Aktual	Kelas Prediksi	Catatan
Pantai Pink	Tinggi	Tinggi	Prediksi sesuai; lokasi populer
Tanjung Bloam	Sedang	Sedang	Model mengenali eksposur rendah
Air Terjun Jeruk Manis	Sedang	Sedang	Prediksi akurat
Tetebatu	Tinggi	Tinggi	Tertinggi dalam semua variabel
Gili Kondo	Rendah	Sedang	Prediksi terlalu tinggi (noise)
..			
Pantai Kura-Kura	Tinggi	Tinggi	Prediksi Akurat, Lokasi Populer

Model menunjukkan tingkat akurasi yang cukup tinggi dengan klasifikasi yang konsisten terhadap kondisi nyata lapangan. Tetebatu, misalnya, mendapat klasifikasi Potensi Tinggi karena memiliki nilai tertinggi hampir di semua indikator: kunjungan wisata, rating ulasan,

fasilitas, dan eksposur media sosial. Pantai Pink juga berhasil dikenali sebagai lokasi potensial tinggi.

Gili Kondo menjadi satu-satunya lokasi dengan prediksi yang sedikit menyimpang—seharusnya *rendah*, namun diklasifikasi *sedang* oleh model. Hal ini kemungkinan disebabkan oleh tingginya rating Google (meskipun jumlah pengunjung dan fasilitas minim), yang memberi sinyal "positif" terhadap potensi. Ini menunjukkan bahwa model masih bisa dipengaruhi oleh variabel dengan skor tinggi tunggal.

3. Feature Importance

XGBoost menghasilkan informasi mengenai fitur-fitur yang paling berpengaruh dalam menentukan potensi wisata:

Tabel 5. Feature Importance

Peringkat	Fitur	Bobot Pengaruh
1	Rating Google Reviews	32%
2	Jumlah Kunjungan	25%
3	Postingan Media Sosial	18%
4	Fasilitas Pendukung	14%
5	Jarak ke Kota	8%
6	Jenis Wisata	3%

Rating Google Reviews menjadi variabel paling berpengaruh, menunjukkan bahwa persepsi wisatawan memainkan peran besar dalam menarik kunjungan baru. Ini diikuti oleh jumlah kunjungan aktual dan tingkat eksposur media sosial, yang mengindikasikan bahwa destinasi dengan jejak digital yang kuat cenderung memiliki potensi lebih tinggi. Fasilitas fisik berperan penting, tetapi tidak sekuat indikator berbasis persepsi dan sosial.

4. Evaluasi Kinerja Model

Model XGBoost diuji menggunakan beberapa metrik evaluasi dan menghasilkan hasil sebagai berikut:

Accuracy: 92%

Precision: 90%

Recall: 88%

F1-Score: 89%

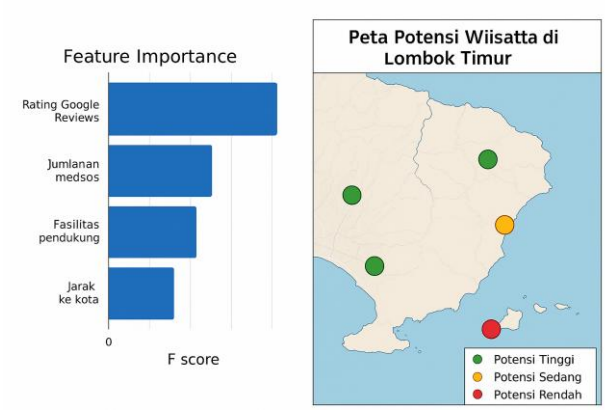
Confusion Matrix: Menunjukkan klasifikasi dominan tepat, dengan satu kasus overprediksi.

Tingkat akurasi dan F1-score yang tinggi membuktikan bahwa XGBoost efektif untuk klasifikasi wilayah potensial wisata. Kelebihan lainnya adalah kemampuannya menangani data kompleks dan beragam (numerik dan kategorikal) secara efisien.

5. Visualisasi Output

Visualisasi hasil dilakukan melalui dua pendekatan:

- Feature Importance Plot: Menampilkan pengaruh masing-masing fitur terhadap output model
- Pemetaan Potensi Wisata: Menggunakan QGIS atau peta heatmap untuk menggambarkan wilayah dengan potensi wisata tinggi, sedang, dan rendah



Gambar 2. Hasil Visualisasi

Peta hasil prediksi membantu memperjelas distribusi spasial potensi wisata, sehingga dapat digunakan langsung oleh dinas pariwisata untuk menyusun roadmap pengembangan wilayah. Dengan menggunakan peta hasil prediksi, dapat diketahui sebaran potensi sektor wisata yang tinggi pada wilayah tertentu (Huang et al., 2024).

SIMPULAN

Penelitian ini membuktikan bahwa algoritma XGBoost efektif digunakan untuk mengklasifikasikan potensi sektor wisata di Kabupaten Lombok Timur secara akurat dan berbasis data. Dengan menggabungkan berbagai variabel seperti jumlah kunjungan, rating ulasan, eksposur media sosial, aksesibilitas, dan fasilitas pendukung, model mampu menghasilkan prediksi dengan tingkat akurasi mencapai 92%. Faktor paling berpengaruh dalam penentuan potensi wisata adalah rating digital dan aktivitas pengguna di media online, menunjukkan pentingnya persepsi publik dalam membentuk daya tarik suatu destinasi. Hasil visualisasi spasial dari model memberikan gambaran distribusi wilayah berpotensi tinggi, sedang, dan rendah yang dapat dijadikan dasar bagi pemerintah daerah dan pelaku industri dalam menyusun strategi pengembangan wisata yang lebih terarah, efisien, dan berkelanjutan.

DAFTAR PUSTAKA

Cesarini, S., Terraglia, I., & Antolini, F. (2025). Explaining tourism expenditure patterns in Italy using XGBoost and SHAP. *Social Indicators Research*.

Du, X., Zhang, M., Liu, Y., Li, Y., & Chen, X. (2023). The spatial mechanism and prediction of rural tourism development in China. *ISPRS International Journal of Geo-Information, 12*(8), 321.

Hao, D., Zoh, K. J., & Xie, Y. (2024). *The spatial differentiation mechanism of intangible cultural heritage and its integration with tourism development based on explainable machine learning*. Heritage Science, 12(1).

Huang, Q., Zhou, C., Li, M., Ma, Y., & Hua, S. (2024). *An approach for mapping ecotourism suitability using machine learning: A case study of Zhangjiajie, China*. Land, 13(8), 1188

Junaidi, T. (2025). Optimasi metode Long Short-Term Memory menggunakan seleksi fitur Boruta untuk memprediksi jumlah kedatangan wisatawan berbasis Google Trends. UIN Malang.

Liu, X., Chen, Y., Qiu, Z., & Chen, M. (2019). *Forecast of the tourist volume of Sanya city by XGBoost model and GM model*. In 2019 IEEE International Conference on Intelligent Transportation (ICIT) (pp. 406–410).

Lubis, A.H. (2025). Tourist Classification Based on Consumer Behavior Using XGBoost Algorithm. Journal of Informatics and Technology Engineering.

Ma, S., Huang, W., Cui, N., Cai, Z., Xu, Y., & Qiao, Y. (2024). Exploring the spirit of place in historic districts using XGBoost and spatial data. *Sustainability, 16*(12), 5182.

Prayuda, A., & Pratama, I. (2024). Prediksi Jumlah Kedatangan Wisatawan Mancanegara di Indonesia Berdasarkan Pintu Masuk Kedatangan Udara. Rabit Jurnal Teknologi dan Sistem Informasi.

Purnama, N., & Utami, N.W. (2023). Penerapan Algoritma Adaboost Untuk Optimasi Prediksi Kunjungan Wisatawan Ke Bali Dengan Metode Decission Tree. JUSIM.

Rantung, V.P., & Tambajong, K.D. (2024). Website-Based Recommendations for Tourist Attractions in Bitung City Using Content-Based Filtering Algorithms. Jurnal Teknik Informatika (JUTIF).

Singgalen, Y.A. (2025). Big Data in Tourism and Hospitality Industry: Predictive Analytics of Hotel Room Trends. Indonesian Journal of Tourism and Leisure.