

ANALISIS DAN PREDIKSI TREN FASHION BATIK INDONESIA MENGUNAKAN LSTM DAN K-MEANS BERBASIS DATA GOOGLE TRENDS

ANALYZING AND PREDICTING INDONESIAN BATIK FASHION TRENDS USING LSTM AND K-MEANS CLUSTERING BASED ON GOOGLE TRENDS DATA

Fryan Gerardi Parantak¹, Teguh Wahyono²

Teknik Informatika, Fakultas Teknologi Informasi, Universitas Kristen Satya Wacana^{1,2}

672022116@student.uksw.edu¹

ABSTRACT

The development of Indonesian batik fashion trends continues to evolve alongside the increasing use of digital technology and online information searches. This study aims to analyze and predict regional batik search trends using the Long Short-Term Memory (LSTM) method and classify the prediction results using K-Means Clustering based on Google Trends data from 2009 to 2026. The data were preprocessed through attribute selection, Min-Max normalization, time-series sequence generation, and train-test splitting before training the LSTM model. The evaluation results produced an MAE of 1.433845, an RMSE of 2.045695, and a MAPE of 29.658822%, indicating that the model was able to predict search trends effectively. The predicted search trends for July 2026 to June 2027 were subsequently grouped into three categories, namely high, medium, and low, using the K-Means algorithm. The findings demonstrate that the integration of LSTM and K-Means can provide both trend prediction and clustering of regional batik search patterns to support data-driven decision-making.

Keywords: Google Trends, LSTM, K-Means Clustering, Regional Batik, Trend Prediction.

ABSTRAK

Perkembangan tren fashion batik di Indonesia terus mengalami perubahan seiring meningkatnya pemanfaatan teknologi digital dan pencarian informasi melalui internet. Penelitian ini bertujuan menganalisis dan memprediksi tren pencarian batik daerah menggunakan metode Long Short-Term Memory (LSTM) serta mengelompokkan hasil prediksi menggunakan K-Means Clustering berdasarkan data Google Trends periode 2009–2026. Data diproses melalui seleksi atribut, normalisasi menggunakan Min-Max Scaler, pembentukan data deret waktu, dan pembagian data latih serta data uji sebelum dilakukan pelatihan model LSTM. Hasil evaluasi model menunjukkan nilai MAE sebesar 1,433845, RMSE sebesar 2,045695, dan MAPE sebesar 29,658822%, yang menunjukkan kemampuan model dalam memprediksi tren pencarian. Hasil prediksi selama Juli 2026 hingga Juni 2027 kemudian dikelompokkan menjadi tiga kategori, yaitu tinggi, sedang, dan rendah menggunakan algoritma K-Means. Penelitian ini menunjukkan bahwa integrasi LSTM dan K-Means mampu memberikan informasi prediksi sekaligus segmentasi tren pencarian batik daerah sebagai pendukung pengambilan keputusan berbasis data.

Kata Kunci: Google Trends, LSTM, K-Means Clustering, Batik Daerah, Prediksi Tren.

PENDAHULUAN

Perkembangan teknologi informasi telah membawa perubahan signifikan dalam berbagai aspek kehidupan, termasuk dalam pelestarian dan promosi budaya Indonesia. Salah satu warisan budaya yang turut mengalami perkembangan melalui pemanfaatan teknologi digital adalah batik. Sebagai identitas budaya bangsa, batik telah ditetapkan oleh *United Nations Educational, Scientific and Cultural Organization (UNESCO)* sebagai Warisan

Budaya Takbenda Dunia (*Intangible Cultural Heritage*) pada tanggal 2 Oktober 2009. Pengakuan tersebut semakin meningkatkan perhatian masyarakat terhadap batik serta mendorong berkembangnya industri batik di berbagai daerah di Indonesia (Hakim, 2018). Beragam jenis batik daerah, seperti Batik Solo, Batik Pekalongan, Batik Cirebon, Batik Madura, Batik Lasem, Batik Jambi, dan Batik Bali memiliki motif, filosofi, serta karakteristik yang berbeda sehingga

tingkat ketertarikan masyarakat terhadap masing-masing jenis batik juga terus mengalami perubahan.

Sebagai bagian dari industri fashion berbasis budaya, batik daerah memiliki karakteristik tren yang dinamis. Perubahan tren dapat terjadi dalam waktu singkat dan dipengaruhi oleh berbagai faktor, seperti musim, perkembangan gaya hidup, promosi digital, media sosial, hingga meningkatnya perhatian masyarakat terhadap isu budaya lokal (Shandy, 2024). Perubahan tersebut menyebabkan tingkat minat masyarakat terhadap berbagai jenis batik daerah terus berubah dari waktu ke waktu. Kondisi ini menjadi tantangan bagi pelaku UMKM, perajin batik, pemerintah, maupun pelaku industri kreatif dalam menyusun strategi produksi, promosi, serta pengembangan produk yang sesuai dengan kebutuhan pasar. Perubahan preferensi konsumen terhadap produk batik menuntut pelaku UMKM untuk terus berinovasi dan menyesuaikan strategi pemasaran agar mampu menjangkau pasar yang lebih luas. Selain itu, pemanfaatan media digital dan diversifikasi produk menjadi faktor penting dalam meningkatkan daya saing serta keberlanjutan industri batik di tengah persaingan pasar yang semakin dinamis (Hasanah & Selasih, 2026). Oleh karena itu, diperlukan suatu pendekatan yang mampu menganalisis perkembangan minat masyarakat secara lebih cepat, objektif, dan berbasis data.

Untuk memahami perubahan minat masyarakat tersebut, salah satu sumber data yang dapat dimanfaatkan adalah *Google Trends*. *Google Trends* merupakan layanan yang disediakan oleh *Google* untuk menampilkan tingkat popularitas suatu kata kunci berdasarkan aktivitas pencarian pengguna dalam periode dan wilayah tertentu (Trilestari, Shovmayanti, Kurniawan, & Prakosa, 2025). Data yang disajikan berupa indeks pencarian dengan rentang nilai 0–100 yang menunjukkan tingkat popularitas relatif suatu kata kunci pada periode tertentu. Dibandingkan survei konvensional, *Google Trends* mampu

memberikan gambaran perubahan minat masyarakat secara lebih cepat dan berkelanjutan sehingga banyak dimanfaatkan dalam penelitian mengenai perilaku konsumen, analisis pasar, hingga peramalan tren pada berbagai bidang. Dalam penelitian ini, data *Google Trends* yang digunakan mencakup periode tahun 2009–2026. Pemilihan periode tersebut didasarkan pada tahun pengakuan batik oleh UNESCO sebagai Warisan Budaya Takbenda Dunia pada tahun 2009, sehingga perkembangan minat masyarakat terhadap berbagai jenis batik daerah dapat dianalisis sejak awal pengakuan tersebut hingga periode terkini.

Salah satu pendekatan *machine learning* yang banyak digunakan dalam analisis data deret waktu (*time series*) adalah *Long Short-Term Memory* (LSTM). LSTM merupakan pengembangan dari *Recurrent Neural Network* (RNN) yang mampu mempelajari hubungan jangka panjang pada data sehingga efektif digunakan untuk memprediksi data yang memiliki pola nonlinier dan kompleks (Brownlee, 2020). Berbagai penelitian menunjukkan bahwa LSTM mampu menghasilkan tingkat akurasi yang tinggi dalam proses peramalan data deret waktu, termasuk data penjualan (Putri & Sadikin, 2021). Selain itu, penelitian (Maharani, Fitriani, Bagus, & Haq, 2026) juga menunjukkan bahwa LSTM efektif digunakan dalam memprediksi tren pencarian berbasis *Google Trends*. Selain melakukan prediksi, pengelompokan kata kunci berdasarkan karakteristik tren pencarian juga diperlukan untuk mengetahui kelompok batik daerah yang memiliki pola perkembangan serupa. Salah satu metode yang banyak digunakan untuk tujuan tersebut adalah *K-Means Clustering*, yaitu algoritma *unsupervised learning* yang mampu mengelompokkan data berdasarkan tingkat kemiripan karakteristik sehingga memudahkan proses analisis dan pengambilan keputusan (Analytics Vidhya, 2019). (Juliyanto et al. 2024) menunjukkan bahwa metode *K-Means* mampu

menghasilkan segmentasi data yang efektif sehingga dapat membantu dalam memahami karakteristik setiap kelompok dan mendukung penyusunan strategi yang lebih tepat. Oleh karena itu, pada penelitian ini metode *K-Means* digunakan untuk mengelompokkan kata kunci batik daerah berdasarkan pola tren pencarian pada Google Trends.

Beberapa penelitian terdahulu telah memanfaatkan *Google Trends* maupun metode *machine learning* dalam analisis tren. (Syarif et al. 2023) menggunakan metode SARIMA untuk memprediksi tren produk fashion berdasarkan data *Google Trends*. Penelitian tersebut menunjukkan bahwa metode SARIMA mampu menangkap pola musiman, namun masih memiliki keterbatasan dalam memodelkan hubungan nonlinier. Menurut (Kharis, Zili, Putri, & Robiansyah, 2023) *Google Trends* merupakan platform yang menyediakan informasi mengenai pola pencarian pengguna internet terhadap suatu topik berdasarkan data riwayat pencarian Google. Data tersebut dapat digunakan untuk menganalisis perkembangan minat masyarakat, perbandingan antarwilayah, maupun perubahan tren dari waktu ke waktu. (Maharani, Fitriani, Bagus, & Haq, 2026) membandingkan metode LSTM dan Prophet dalam memprediksi tren pencarian busana Lebaran berbasis *Google Trends* dan menunjukkan bahwa model LSTM menghasilkan tingkat akurasi prediksi yang lebih baik dibandingkan *Prophet* dalam memodelkan pola data pencarian pada *Google Trends*. Penelitian lain oleh (Tamami dan Arifin 2024) menunjukkan bahwa LSTM mampu menghasilkan prediksi dengan akurasi tinggi pada data penjualan, sedangkan (Ardi et al. 2023) dan (Alifa et al. 2024) membuktikan bahwa metode *K-Means* efektif digunakan untuk mengelompokkan pelanggan maupun produk berdasarkan karakteristik data. (Meriani & Rahmatulloh, 2024) juga membandingkan metode LSTM dan *Gated Recurrent Unit* (GRU) untuk memprediksi harga emas berbasis data time series. Hasil

penelitian menunjukkan bahwa model LSTM menghasilkan nilai MAE sebesar 0,0389, RMSE sebesar 0,0475, dan MAPE sebesar 5,2047%, sehingga mampu memberikan performa prediksi yang lebih baik dibandingkan model GRU. Untuk memahami perubahan minat masyarakat tersebut, salah satu sumber data yang dapat dimanfaatkan adalah *Google Trends*. *Google Trends* merupakan layanan yang disediakan oleh Google untuk menampilkan tingkat popularitas suatu kata kunci berdasarkan aktivitas pencarian pengguna dalam periode dan wilayah tertentu. Data yang disajikan berupa indeks pencarian dengan rentang nilai 0–100 yang menunjukkan tingkat popularitas relatif suatu kata kunci pada periode tertentu. Selain menampilkan popularitas pencarian, *Google Trends* juga dapat dimanfaatkan untuk menganalisis pola pencarian masyarakat, mengeksplorasi tren informasi, serta menjadi sumber data dalam analisis berbasis big data sehingga banyak digunakan pada berbagai penelitian di bidang perilaku masyarakat, pemasaran, maupun peramalan tren (Krismayani & Mafar, 2024). Dibandingkan survei konvensional, *Google Trends* mampu memberikan gambaran perubahan minat masyarakat secara lebih cepat dan berkelanjutan sehingga banyak dimanfaatkan dalam penelitian mengenai perilaku konsumen, analisis pasar, hingga peramalan tren pada berbagai bidang karakteristik data. Meskipun demikian, penelitian-penelitian tersebut masih menggunakan metode prediksi dan *clustering* secara terpisah atau berfokus pada objek selain batik daerah.

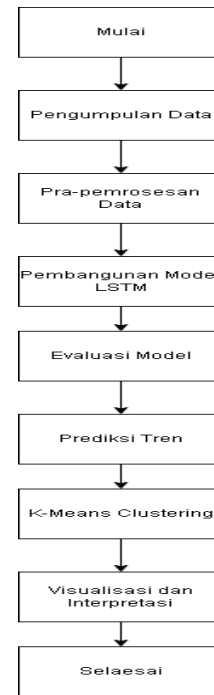
(Sarimole dan Hakim 2024) menerapkan metode *K-Means Clustering* untuk mengelompokkan data penjualan dalam menentukan prediksi stok barang. Hasil penelitian menunjukkan bahwa metode *K-Means* mampu membagi data menjadi tiga kelompok, yaitu kategori penjualan tinggi, sedang, dan rendah sehingga dapat mendukung proses

pengambilan keputusan terkait pengelolaan stok barang.

Berdasarkan uraian tersebut, masih terdapat celah penelitian (*research gap*), yaitu belum banyak penelitian yang mengintegrasikan metode *Long Short-Term Memory* (LSTM) dan *K-Means Clustering* untuk menganalisis, memprediksi, serta mengelompokkan tren pencarian batik daerah di Indonesia menggunakan data *Google Trends*. Oleh karena itu, penelitian ini berupaya menjawab permasalahan mengenai bagaimana membangun model LSTM yang mampu memprediksi tren pencarian batik daerah berdasarkan data *Google Trends* periode 2009–2026, bagaimana mengevaluasi tingkat akurasi model menggunakan metrik MAE, RMSE, dan MAPE, serta bagaimana mengelompokkan hasil prediksi menggunakan metode *K-Means Clustering* ke dalam kategori tren pencarian tinggi, sedang, dan rendah. Berdasarkan permasalahan tersebut, penelitian ini bertujuan menerapkan metode LSTM untuk memprediksi perkembangan tren pencarian berbagai jenis batik daerah serta menerapkan metode *K-Means Clustering* untuk mengelompokkan kata kunci batik daerah berdasarkan karakteristik pola tren pencariannya. Hasil penelitian diharapkan dapat memberikan informasi mengenai perkembangan minat masyarakat terhadap batik daerah serta menjadi referensi bagi pelaku UMKM, perajin batik, pemerintah, dan akademisi dalam mendukung pelestarian budaya, pengembangan produk, serta penyusunan strategi promosi berbasis data.

METODE

Pada penelitian ini akan dilakukan beberapa tahapan yaitu pengumpulan data, pra-pemrosesan data, pembangunan model LSTM, evaluasi model, prediksi tren, *K-Means Clustering*, visualisasi dan interpretasi. Tahapan ini di tunjukan pada Gambar 1.



Gambar 1. Tahapan Penelitian

Tahap pertama merupakan proses pengumpulan data yang akan digunakan sebagai masukan (input) pada model prediksi. Data diperoleh dari layanan *Google Trends*, yaitu layanan analisis tren pencarian yang disediakan oleh Google untuk menampilkan tingkat popularitas suatu kata kunci berdasarkan aktivitas pencarian pengguna pada wilayah dan periode tertentu. Nilai yang diberikan oleh *Google Trends* berada pada rentang 0–100, di mana nilai 100 menunjukkan tingkat popularitas tertinggi pada periode pengamatan, sedangkan nilai 0 menunjukkan tingkat pencarian yang sangat rendah atau tidak mencukupi untuk ditampilkan.

Data yang dikumpulkan terdiri atas delapan kata kunci batik daerah di Indonesia, yaitu Batik Papua, Batik Solo, Batik Jambi, Batik Pekalongan, Batik Yogyakarta, Batik Cirebon, Batik Bali, dan Batik Lasem. Seluruh data diambil dengan interval bulanan mulai Januari 2009 hingga Juni 2026 sehingga membentuk data deret *time* (*time series*) yang digunakan sebagai dasar dalam proses pelatihan dan prediksi model.

Data hasil pengambilan kemudian disimpan dalam format *Comma Separated Values* (CSV) dan dibaca menggunakan

pustaka Pandal. Selanjutnya atribut *Time* dikonversi ke dalam format *datetime* agar dapat dikenali sebagai data deret *time*. Berdasarkan proses pembacaan data diperoleh dataset yang terdiri atas 210 baris data dan 9 kolom, yaitu satu kolom *time* serta delapan kolom indeks pencarian batik daerah.

Tahap pra-pemrosesan diawali dengan pemisahan atribut *time* dari atribut indeks pencarian sehingga hanya delapan variabel batik daerah yang digunakan sebagai masukan model. Selanjutnya seluruh data dinormalisasi menggunakan metode Min-Max Scaling agar setiap variabel berada pada rentang nilai 0–1. Normalisasi dilakukan untuk mengurangi perbedaan skala antarvariabel sehingga proses pembelajaran model menjadi lebih stabil. Setelah itu data dibentuk menjadi data deret *time* menggunakan teknik *sliding window* dengan panjang urutan (*time step*) sebanyak 24 bulan, sehingga setiap data masukan terdiri atas dua puluh empat bulan sebelumnya untuk memprediksi indeks pencarian pada bulan berikutnya. Hasil pembentukan data menghasilkan 186 sampel dengan dimensi (186, 24, 8) sebagai data masukan dan 186 sampel dengan dimensi (186, 8) sebagai data target. Data tersebut kemudian dibagi menjadi 80% data pelatihan dan 20% data pengujian, sehingga diperoleh 148 data pelatihan dan 38 data pengujian.

Setelah proses pra-pemrosesan selesai, dilakukan pembangunan model prediksi menggunakan algoritma *Long Short-Term Memory* (LSTM). Model dibangun menggunakan pustaka TensorFlow/Keras dengan arsitektur yang terdiri atas satu lapisan Input yang menerima data berukuran (24, 8), yaitu 24 bulan data historis dengan delapan variabel batik daerah. Selanjutnya data diproses oleh dua lapisan LSTM, masing-masing terdiri atas 128 unit neuron pada lapisan pertama dan 64 unit neuron pada lapisan kedua. Lapisan LSTM pertama menggunakan parameter `return_sequences=True` agar seluruh informasi urutan dapat diteruskan

ke lapisan berikutnya. Di antara kedua lapisan LSTM ditambahkan lapisan *Dropout* dengan nilai 0,1 untuk mengurangi kemungkinan terjadinya *overfitting* selama proses pelatihan model. Setelah itu digunakan lapisan Dense yang terdiri atas 32 neuron dengan fungsi aktivasi *Rectified Linear Unit* (ReLU) untuk mempelajari hubungan nonlinier antarvariabel, kemudian diakhiri dengan lapisan Dense sebanyak 8 neuron sebagai keluaran model yang merepresentasikan prediksi indeks pencarian untuk delapan jenis batik daerah secara bersamaan.

Model yang telah dibangun kemudian dikompilasi menggunakan optimizer Adam karena mampu menyesuaikan laju pembelajaran (*learning rate*) secara adaptif sehingga proses pelatihan menjadi lebih stabil. Fungsi kerugian (*loss function*) yang digunakan adalah *Mean Squared Error* (MSE) karena sesuai untuk permasalahan regresi dan prediksi data deret *time*, sedangkan metrik evaluasi yang digunakan selama proses pelatihan adalah *Mean Absolute Error* (MAE). Untuk meningkatkan kualitas pelatihan model, diterapkan teknik *Early Stopping* dengan nilai *patience* sebesar 15 *epoch* sehingga proses pelatihan akan dihentikan apabila nilai *validation loss* tidak mengalami perbaikan dalam 15 iterasi berturut-turut. Selain itu diterapkan *Reduce Learning Rate on Plateau* dengan faktor pengurangan sebesar 0,5 dan *patience* sebanyak 5 *epoch* untuk menurunkan nilai *learning rate* secara otomatis ketika proses pembelajaran mulai mengalami stagnasi.

Proses pelatihan model dilakukan menggunakan data pelatihan yang telah dipersiapkan sebelumnya dengan jumlah maksimum 200 *epoch* dan batch size sebesar 16. Selama proses pelatihan, model mempelajari pola perubahan indeks pencarian masing-masing jenis batik berdasarkan data historis selama 24 bulan sebelumnya. Pada setiap *epoch* dilakukan proses validasi menggunakan data pengujian sehingga perkembangan nilai *training loss* dan *validation loss* dapat

dipantau untuk memastikan model tidak mengalami *overfitting* maupun *underfitting*.

Setelah proses pelatihan selesai, model dievaluasi menggunakan data pengujian yang belum pernah digunakan selama proses pembelajaran. Nilai prediksi yang dihasilkan model kemudian dikembalikan ke skala aslinya melalui proses inverse transform menggunakan objek *MinMaxScaler* sehingga hasil prediksi dapat dibandingkan dengan data aktual. Kinerja model selanjutnya dievaluasi menggunakan tiga metrik pengukuran, yaitu *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE). Ketiga metrik tersebut digunakan untuk mengukur tingkat kesalahan prediksi model, di mana semakin kecil nilai yang dihasilkan menunjukkan bahwa model memiliki kemampuan prediksi yang semakin baik.

Setelah model menunjukkan performa yang memadai, dilakukan proses prediksi tren pencarian untuk 12 bulan berikutnya. Prediksi dilakukan secara iteratif (*recursive forecasting*) dengan menggunakan 24 bulan data historis terakhir sebagai masukan awal model. Pada setiap iterasi, hasil prediksi bulan sebelumnya dimasukkan kembali ke dalam urutan data sebagai input untuk memprediksi bulan berikutnya hingga diperoleh prediksi selama dua belas bulan berturut-turut. Seluruh hasil prediksi kemudian dikembalikan ke skala asli menggunakan proses inverse transform dan disusun ke dalam sebuah dataset baru yang berisi periode *time* beserta nilai prediksi masing-masing jenis batik daerah.

Tahap berikutnya adalah proses *K-Means Clustering* yang bertujuan mengelompokkan batik daerah berdasarkan kemiripan pola hasil prediksi selama dua belas bulan. Sebelum proses *clustering* dilakukan, data prediksi terlebih dahulu ditransformasikan sehingga setiap baris merepresentasikan satu jenis batik dengan

dua belas atribut nilai prediksi bulanan. Selanjutnya data dinormalisasi menggunakan *StandardScaler* agar seluruh atribut memiliki skala yang seragam sehingga tidak ada variabel yang mendominasi proses pengelompokan. Penentuan jumlah *cluster optimal* dilakukan menggunakan metode Elbow dengan menghitung nilai *Sum of Squared Error* (SSE) pada beberapa alternatif jumlah *cluster*. Hasil tersebut kemudian divalidasi menggunakan *Silhouette Score* untuk mengukur kualitas pemisahan antar *cluster*. Berdasarkan hasil evaluasi tersebut dipilih jumlah *cluster* yang paling representatif, kemudian dilakukan proses pengelompokan menggunakan algoritma *K-Means*. Setiap jenis batik selanjutnya memperoleh label *cluster* sesuai karakteristik pola tren pencariannya.

Tahap selanjutnya visualisasi dan interpretasi hasil. Hasil prediksi divisualisasikan dalam bentuk grafik deret *time* untuk menggambarkan perkembangan tren pencarian masing-masing jenis batik pada periode prediksi. Selanjutnya hasil pengelompokan divisualisasikan menggunakan *Principal Component Analysis* (PCA) sehingga distribusi setiap *cluster* dapat diamati dalam ruang dua dimensi. Berdasarkan rata-rata nilai indeks pencarian pada masing-masing *cluster*, setiap kelompok kemudian diberi kategori tinggi, sedang, dan rendah untuk memudahkan interpretasi hasil penelitian.

HASIL DAN PEMBAHASAN

Penelitian ini menerapkan metode *Long Short-Term Memory* (LSTM) untuk memprediksi tren pencarian delapan jenis batik daerah di Indonesia berdasarkan data Google Trends periode Januari 2009 hingga Juni 2026. Selanjutnya, hasil prediksi tersebut dikelompokkan menggunakan metode *K-Means Clustering* untuk mengetahui kategori tren pencarian setiap jenis batik. Tahapan hasil penelitian meliputi proses pengumpulan data, pra-pemrosesan data, pembangunan model LSTM, evaluasi model, prediksi tren,

hingga pengelompokan menggunakan *K-Means*.

Tahap awal penelitian dilakukan dengan mengumpulkan data Google Trends menggunakan delapan kata kunci batik daerah, yaitu Batik Papua, Batik Solo, Batik Jambi, Batik Pekalongan, Batik Yogyakarta, Batik Cirebon, Batik Bali, dan Batik Lasem. Data dikumpulkan dalam rentang waktu Januari 2009 hingga Juni 2026 dengan interval bulanan sehingga diperoleh 210 data observasi. Setiap data menunjukkan indeks pencarian Google

Trends dengan rentang nilai 0–100, yang merepresentasikan tingkat popularitas relatif suatu kata kunci pada bulan tertentu.

Hasil pembacaan dataset menunjukkan bahwa data terdiri atas 210 baris dan 9 kolom, yaitu satu kolom waktu (*Time*) dan delapan kolom indeks pencarian batik daerah. Kolom *time* kemudian dikonversi ke format *datetime* agar dapat diproses sebagai data deret *time* (*time series*). Cuplikan dataset yang digunakan dalam penelitian ditunjukkan pada Gambar 2.

Tabel 1. Lima baris awal dataset

Time	Batik Papua	Batik Solo	Batik Jambi	Batik Pekalongan	Batik Yogyakarta	Batik Cirebon	Batik Bali	Batik Lasem
2009-01-01	0	28	0	36	22	17	11	0
2009-02-01	0	40	0	69	25	20	15	0
2009-03-01	0	32	0	43	21	16	0	0
2009-04-01	0	23	0	45	15	11	0	0
2009-05-01	0	30	0	41	26	9	0	0

Berdasarkan Gambar 2, terlihat bahwa setiap baris merepresentasikan satu periode bulanan, sedangkan setiap kolom menunjukkan indeks pencarian masing-masing jenis batik. Nilai indeks yang lebih tinggi menunjukkan tingkat popularitas pencarian yang lebih besar pada periode tersebut.

Setelah data berhasil dikumpulkan, tahap berikutnya adalah melakukan pra-pemrosesan sebelum data digunakan pada model LSTM. Kolom *Time* dipisahkan dari data numerik sehingga hanya delapan variabel indeks pencarian yang digunakan sebagai fitur masukan model. Selanjutnya seluruh data dinormalisasi menggunakan metode *MinMaxScaler* sehingga nilai setiap variabel berada pada rentang 0 hingga 1. Proses normalisasi bertujuan untuk mengurangi perbedaan skala antarvariabel sehingga model LSTM dapat mempelajari pola historis data secara lebih efektif.

Tahap berikutnya adalah pembentukan data deret *time* menggunakan teknik *sliding window* dengan panjang urutan (*time step*) sebesar 24 bulan. Dengan pendekatan ini, model menggunakan data historis selama dua tahun sebelumnya untuk memprediksi indeks pencarian pada bulan berikutnya. Hasil transformasi menghasilkan 186 sampel data, dengan dimensi data masukan sebesar (186, 24, 8) dan data target sebesar (186, 8).

Selanjutnya data dibagi menjadi data pelatihan dan data pengujian menggunakan rasio 80:20. Berdasarkan proses tersebut diperoleh 148 sampel sebagai data pelatihan dan 38 sampel sebagai data pengujian. Pembagian ini bertujuan agar model dapat mempelajari pola historis pada data pelatihan, kemudian dievaluasi menggunakan data yang belum pernah dilihat sebelumnya.

Tabel 2. Hasil Pembentukan Data Time Series

No	Parameter	Nilai
----	-----------	-------

1	Jumlah data awal	210
2	Time step	24 bulan
3	Jumlah variabel	8
	Jumlah sampel	186
4	Bentuk data X	(186,24,8)
5	Bentuk data y	(186,8)

Tabel 1 menunjukkan hasil pembentukan data deret *time (time series)* menggunakan metode sliding window. Berdasarkan hasil pra-pemrosesan, data awal yang berjumlah 210 observasi dengan 8 variabel diubah menjadi 186 sampel menggunakan panjang urutan (*time step*) sebanyak 24 bulan. Nilai *time step* tersebut menunjukkan bahwa model memanfaatkan data historis selama dua tahun sebelumnya sebagai masukan untuk memprediksi nilai pada periode berikutnya. Hasil transformasi menghasilkan data masukan (X) berdimensi (186, 24, 8) yang berarti terdapat 186 sampel, masing-masing terdiri atas 24 periode waktu dengan 8 variabel batik. Sementara itu, data target (y) berdimensi (186, 8) yang menunjukkan nilai yang akan diprediksi untuk delapan jenis batik pada periode selanjutnya. Transformasi ini diperlukan agar data sesuai dengan format masukan yang dibutuhkan oleh model LSTM dalam mempelajari pola temporal dari data deret *time*.

Tabel 3. Pembagian Dataset

No	Dataset	Jumlah sampel
1	Data pelatihan	148
2	Data pengujian	38

Tabel 2 menunjukkan hasil pembagian dataset menjadi data pelatihan dan data pengujian dengan rasio 80:20. Berdasarkan pembagian tersebut diperoleh 148 sampel sebagai data pelatihan dan 38 sampel sebagai data pengujian. Data pelatihan digunakan untuk membangun model LSTM dan mempelajari pola historis tren pencarian batik, sedangkan data pengujian digunakan untuk mengevaluasi

kemampuan model dalam melakukan prediksi terhadap data yang belum pernah dipelajari sebelumnya. Pembagian data secara proporsional ini bertujuan untuk menghindari *overfitting* sekaligus memastikan bahwa performa model dapat dievaluasi secara objektif menggunakan data yang berbeda dari proses pelatihan.

Setelah proses pra-pemrosesan selesai, tahap berikutnya adalah membangun model *Long Short-Term Memory (LSTM)* menggunakan pustaka TensorFlow Keras dengan arsitektur Sequential. Arsitektur Sequential dipilih karena lapisan-lapisan penyusun model disusun secara berurutan, sehingga keluaran (output) dari suatu lapisan menjadi masukan (input) bagi lapisan berikutnya. Model ini digunakan untuk mempelajari pola historis data indeks pencarian Google Trends sehingga mampu melakukan prediksi nilai pada periode berikutnya.

Model diawali dengan Input Layer yang menerima data masukan berdimensi (24, 8). Dimensi tersebut menunjukkan bahwa model menerima data historis selama 24 periode dengan 8 variabel, yaitu Batik Papua, Batik Solo, Batik Jambi, Batik Pekalongan, Batik Yogyakarta, Batik Cirebon, Batik Bali, dan Batik Lasem.

Lapisan pertama menggunakan LSTM dengan 128 unit dan parameter `return_sequences=True`. Penggunaan parameter tersebut bertujuan agar seluruh keluaran dari setiap urutan data diteruskan ke lapisan LSTM berikutnya sehingga informasi historis dapat dipelajari secara lebih mendalam. Setelah lapisan LSTM pertama, ditambahkan *Dropout* sebesar 0,1 untuk mengurangi kemungkinan terjadinya *overfitting* selama proses pelatihan model.

Selanjutnya digunakan lapisan LSTM kedua dengan 64 unit. Berbeda dengan lapisan sebelumnya, lapisan ini hanya menghasilkan keluaran pada urutan terakhir sehingga informasi yang telah dipelajari dapat diteruskan menuju lapisan Dense. Setelah itu kembali ditambahkan *Dropout* sebesar 0,1 untuk meningkatkan

kemampuan generalisasi model terhadap data yang belum pernah dipelajari.

Tahap berikutnya adalah penggunaan Dense Layer dengan 32 neuron dan fungsi aktivasi *Rectified Linear Unit* (ReLU). Lapisan ini berfungsi mengolah representasi hasil ekstraksi fitur dari lapisan LSTM sebelum menghasilkan nilai prediksi. Pada lapisan terakhir digunakan Dense Output Layer yang terdiri atas 8 neuron, sesuai dengan jumlah variabel yang diprediksi secara bersamaan.

Setelah seluruh arsitektur model selesai dibangun, model dikompilasi menggunakan optimizer Adam dengan fungsi kerugian (loss function) *Mean Squared Error* (MSE) dan metrik evaluasi *Mean Absolute Error* (MAE). *Optimizer Adam* dipilih karena mampu melakukan pembaruan bobot secara adaptif sehingga proses pelatihan model menjadi lebih efisien, sedangkan MSE digunakan sebagai fungsi kerugian untuk mengukur selisih kuadrat antara nilai aktual dan nilai prediksi selama proses pelatihan.

Tabel 4 Arsitektur Model LSTM

Layer	Parameter	Nilai
lstm(LSTM)	(None, 24, 70,144 128)	
dropout(Dropout)	Data pengujian	0
lstm_1 (LSTM)	(None, 64)	49,408
dropout_1(Dropout)	(None, 64)	0
Dense (Dense)	(None, 32)	2,080
Dense_ (Dense)	(None,8)	264

Berdasarkan Gambar 3, model yang dibangun memiliki 121.896 parameter yang seluruhnya merupakan trainable parameters, sehingga seluruh parameter akan diperbarui selama proses pelatihan. Jumlah parameter tersebut menunjukkan bahwa model memiliki kapasitas yang cukup untuk mempelajari pola historis indeks pencarian delapan jenis batik daerah pada data Google Trends.

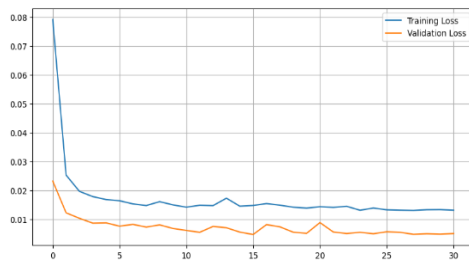
Setelah arsitektur model selesai dibangun, tahap berikutnya adalah melakukan pelatihan (training) menggunakan data pelatihan (X_{train} dan y_{train}) yang terdiri atas 148 sampel, sedangkan proses validasi dilakukan menggunakan 38 sampel data pengujian

(X_{test} dan y_{test}). Proses pelatihan dirancang dengan maksimum 200 *epoch* dan *batch size* sebesar 16. Selain itu, penelitian ini menerapkan dua callback, yaitu *EarlyStopping* dan *ReduceLROnPlateau*, untuk meningkatkan efisiensi proses pelatihan.

Callback EarlyStopping digunakan untuk menghentikan proses pelatihan secara otomatis apabila nilai *validation loss* tidak mengalami perbaikan selama 15 *epoch* berturut-turut. Parameter *restore_best_weights=True* diterapkan sehingga bobot model terbaik selama proses pelatihan akan digunakan kembali sebagai model akhir. Sementara itu, *callback ReduceLROnPlateau* digunakan untuk menurunkan nilai *learning rate* sebesar 50% apabila nilai *validation loss* tidak menunjukkan peningkatan selama 5 *epoch*. Penggunaan kedua callback tersebut bertujuan untuk mengurangi kemungkinan terjadinya *overfitting* serta meningkatkan proses konvergensi model.

Berdasarkan hasil pelatihan, proses training berhenti pada *epoch* ke-31 meskipun jumlah *epoch* maksimum ditetapkan sebanyak 200 *epoch*. Hal tersebut menunjukkan bahwa model telah mencapai kondisi optimal sebelum seluruh *epoch* selesai dijalankan sehingga proses pelatihan dihentikan secara otomatis oleh mekanisme *EarlyStopping*. Selama proses pelatihan juga terjadi penyesuaian nilai *learning rate* sebanyak beberapa kali melalui mekanisme *ReduceLR On Plateau*, yang menunjukkan bahwa model melakukan penyesuaian laju pembelajaran untuk memperoleh nilai loss yang lebih baik.

Perkembangan nilai training loss dan validation loss selama proses pelatihan ditunjukkan pada Gambar 4.



Gambar 2. Grafik Training Loss dan Validation Loss LSTM

Berdasarkan Gambar 4, nilai training loss maupun *validation loss* mengalami penurunan yang cukup signifikan pada awal proses pelatihan. Setelah beberapa *epoch*, kedua kurva menunjukkan kecenderungan yang semakin stabil hingga mencapai nilai yang relatif konstan. Selain itu, jarak antara kurva training loss dan *validation loss* tidak menunjukkan perbedaan yang terlalu besar. Kondisi tersebut mengindikasikan bahwa model mampu mempelajari pola data pelatihan dengan baik serta memiliki kemampuan generalisasi yang cukup baik terhadap data pengujian. Dengan demikian, model LSTM yang dibangun dinilai telah memiliki performa yang memadai untuk digunakan pada proses prediksi tren pencarian batik daerah.

Setelah proses pelatihan selesai dilakukan, tahap berikutnya adalah mengevaluasi performa model LSTM menggunakan data pengujian (X_{test}). Sebelum proses evaluasi dilakukan, hasil prediksi model dan data aktual dikembalikan terlebih dahulu ke skala aslinya menggunakan fungsi *inverse_transform()* dari *MinMaxScaler*. Proses ini bertujuan agar nilai prediksi yang dihasilkan dapat dibandingkan secara langsung dengan nilai aktual indeks pencarian Google Trends.

Evaluasi model dilakukan menggunakan tiga metrik pengukuran, yaitu *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE). MAE digunakan untuk mengukur rata-rata selisih absolut antara nilai aktual dan nilai prediksi, RMSE digunakan untuk mengukur besarnya kesalahan prediksi dengan memberikan penalti yang lebih

besar terhadap kesalahan yang tinggi, sedangkan MAPE digunakan untuk mengetahui tingkat kesalahan prediksi dalam bentuk persentase sehingga lebih mudah diinterpretasikan.

Tabel 5 Hasil Evaluasi Model LSTM

No	Metrik	Nilai
1	MAE	1,434
2	RMSE	2,046
3	MAPE	29,66 %

Berdasarkan Tabel 3, model LSTM menghasilkan nilai MAE sebesar 1,434, yang menunjukkan bahwa rata-rata selisih antara nilai prediksi dan nilai aktual sekitar 1,4 poin indeks pencarian Google Trends. Nilai RMSE sebesar 2,046 menunjukkan bahwa rata-rata kesalahan prediksi masih berada pada rentang yang relatif kecil meskipun memberikan penalti yang lebih besar terhadap prediksi yang memiliki kesalahan tinggi. Sementara itu, nilai MAPE sebesar 29,66% menunjukkan bahwa tingkat kesalahan prediksi model berada di bawah 30%, sehingga model masih mampu memberikan hasil prediksi yang cukup baik terhadap data yang digunakan dalam penelitian.

Meskipun nilai MAPE belum berada pada kategori sangat tinggi, hasil tersebut menunjukkan bahwa model LSTM telah mampu mempelajari pola historis pencarian batik daerah dan menghasilkan prediksi yang cukup mendekati nilai aktual. Perbedaan antara nilai aktual dan prediksi dipengaruhi oleh fluktuasi data Google Trends yang cenderung berubah secara dinamis akibat berbagai faktor, seperti perubahan minat masyarakat, musim, promosi digital, maupun peristiwa tertentu yang tidak dapat dipelajari sepenuhnya oleh model.

Selain dilakukan secara keseluruhan, evaluasi juga dilakukan terhadap masing-masing jenis batik untuk mengetahui performa model pada setiap variabel. Hasil evaluasi tersebut disajikan pada Tabel 4.

Tabel 6. Evaluasi model keseluruhan

N o	Jenis Batik	MAE	RMSE	MAPE (%)
1	Batik Papua	0.8788 11	1.0946 59	23.349 798
2	Batik Solo	1.6466 17	2.2263 20	11.976 824
3	Batik Jambi	0.5703 19	0.6857 42	36.103 313
4	Batik Pekalongan	4.2950 28	4.4770 68	63.496 119
5	Batik Yogyakarta	1.7551 41	1.9256 27	37.752 384
6	Batik Cirebon	0.8151 95	1.0623 66	20.981 964
7	Batik Bali	1.0145 24	1.2845 75	13.599 424
8	Batik Lasem	0.4951 27	0.5683 33	30.010 749

Berdasarkan Tabel 4, performa model LSTM menunjukkan hasil yang berbeda pada masing-masing jenis batik. Batik Solo menghasilkan nilai MAE sebesar 1.646617, RMSE sebesar 2.226320, dan MAPE sebesar 11.976824%, yang merupakan nilai MAPE terendah dibandingkan jenis batik lainnya. Hasil tersebut menunjukkan bahwa model mampu memprediksi tren pencarian Batik Solo dengan tingkat kesalahan yang relatif rendah. Selain itu, Batik Bali dan Batik Cirebon juga menunjukkan performa prediksi yang cukup baik dengan nilai MAPE masing-masing sebesar 13.599424% dan 20.981964%.

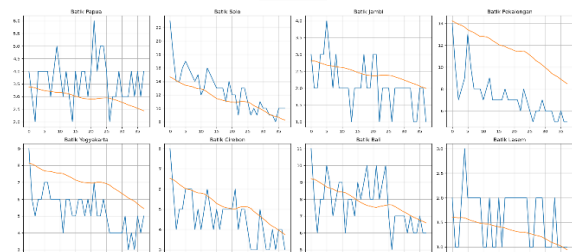
Sebaliknya, Batik Pekalongan memperoleh nilai MAE sebesar 4.295028, RMSE sebesar 4.477068, dan MAPE sebesar 63.496119%, yang merupakan nilai tertinggi pada seluruh variabel. Nilai tersebut menunjukkan bahwa model mengalami kesulitan dalam memprediksi pola pencarian Batik Pekalongan dibandingkan jenis batik lainnya. Tingginya nilai kesalahan mengindikasikan bahwa pola pencarian Batik Pekalongan memiliki fluktuasi yang lebih besar

sehingga lebih sulit dipelajari oleh model LSTM.

Sementara itu, Batik Papua, Batik Jambi, Batik Yogyakarta, dan Batik Lasem menghasilkan nilai MAPE berturut-turut sebesar 23.349798%, 36.103313%, 37.752384%, dan 30.010749%. Hasil tersebut menunjukkan bahwa model masih mampu mengikuti pola perubahan indeks pencarian pada keempat jenis batik tersebut, meskipun tingkat kesalahan prediksinya lebih tinggi dibandingkan Batik Solo, Batik Bali, dan Batik Cirebon.

Secara keseluruhan, hasil evaluasi menunjukkan bahwa kemampuan model LSTM dalam melakukan prediksi dipengaruhi oleh karakteristik pola pencarian masing-masing jenis batik pada Google Trends. Jenis batik yang memiliki pola pencarian lebih stabil cenderung menghasilkan nilai kesalahan yang lebih rendah, sedangkan jenis batik dengan pola pencarian yang lebih fluktuatif menghasilkan tingkat kesalahan yang lebih tinggi.

Setelah dilakukan evaluasi menggunakan metrik MAE, RMSE, dan MAPE, tahap berikutnya adalah membandingkan hasil prediksi model dengan data aktual pada setiap jenis batik. Perbandingan tersebut bertujuan untuk memberikan gambaran visual mengenai kemampuan model LSTM dalam mengikuti pola perubahan indeks pencarian Google Trends selama periode pengujian. Hasil perbandingan antara nilai aktual dan nilai prediksi untuk masing-masing jenis batik ditunjukkan pada Gambar 5.



Gambar 3. Perbandingan Nilai Aktual dan Prediksi pada Delapan Jenis Batik

Berdasarkan Gambar 5, secara umum kurva hasil prediksi mampu mengikuti pola

perubahan kurva data aktual pada sebagian besar jenis batik. Hal ini menunjukkan bahwa model LSTM telah berhasil mempelajari hubungan temporal dari data historis sehingga mampu menghasilkan prediksi yang mendekati nilai aktual. Meskipun demikian, pada beberapa periode masih terdapat selisih antara nilai aktual dan nilai prediksi, terutama ketika terjadi perubahan indeks pencarian yang cukup tajam.

Pada Batik Solo, kurva prediksi memiliki pola yang paling mendekati data aktual. Hasil tersebut sejalan dengan nilai MAPE sebesar 11.976824% yang diperoleh pada tahap evaluasi. Kondisi ini menunjukkan bahwa pola pencarian Batik Solo relatif lebih konsisten sehingga lebih mudah dipelajari oleh model LSTM.

Hal yang serupa juga terlihat pada Batik Bali dan Batik Cirebon, di mana kurva prediksi mampu mengikuti perubahan data aktual dengan penyimpangan yang relatif kecil. Kesesuaian tersebut mendukung hasil evaluasi sebelumnya yang menunjukkan nilai MAPE masing-masing sebesar 13.599424% dan 20.981964%.

Sebaliknya, pada Batik Pekalongan terlihat adanya perbedaan yang lebih besar antara kurva aktual dan kurva prediksi, terutama pada periode ketika indeks pencarian mengalami perubahan yang cukup signifikan. Kondisi ini sejalan dengan hasil evaluasi yang menunjukkan nilai MAPE sebesar 63.496119%, sehingga model mengalami kesulitan dalam memprediksi pola pencarian Batik Pekalongan secara akurat.

Sementara itu, Batik Papua, Batik Jambi, Batik Yogyakarta, dan Batik Lasem menunjukkan bahwa model masih mampu mengikuti arah perubahan tren pencarian, meskipun pada beberapa periode terdapat selisih antara nilai aktual dan hasil prediksi. Perbedaan tersebut menunjukkan bahwa model LSTM telah berhasil menangkap pola umum data, namun masih memiliki keterbatasan dalam mengikuti fluktuasi yang terjadi secara tiba-tiba.

Secara keseluruhan, hasil visualisasi memperkuat hasil evaluasi kuantitatif yang telah diperoleh sebelumnya. Kesesuaian pola antara kurva aktual dan kurva prediksi menunjukkan bahwa model LSTM mampu mempelajari karakteristik data Google Trends dan menghasilkan prediksi yang representatif terhadap tren pencarian batik daerah selama periode pengujian.

Setelah model LSTM berhasil dilatih dan dievaluasi, tahap berikutnya adalah melakukan prediksi tren pencarian batik daerah untuk 12 periode berikutnya. Proses prediksi dilakukan menggunakan 24 data terakhir sebagai masukan (input) model. Selanjutnya, hasil prediksi pada setiap periode digunakan kembali sebagai masukan untuk memprediksi periode berikutnya hingga diperoleh prediksi selama 12 bulan, yaitu mulai Juli 2026 hingga Juni 2027.

Hasil prediksi yang diperoleh kemudian dikembalikan ke skala aslinya menggunakan fungsi `inverse_transform()` sehingga nilai yang dihasilkan kembali berada pada skala indeks pencarian Google Trends. Hasil prediksi tren pencarian delapan jenis batik daerah ditunjukkan pada Gambar 6.

Tabel 7. Hasil Prediksi Tren Pencarian Batik Daerah Periode Juli 2026 – Juni 2027

Ti m e	Bati k Pap ua	Bati k Solo	Bati k Jam bi	Batik Pekal ongan	Batik Yogya karta	Bati k Cire bon	Bati k Bali	Bati k Las em
20								
26-07-01	2.38 3514	8.14 5912	1.95 9975	8.2132 51	5.2626 74	3.62 0066	6.52 5505	0.89 7706
20								
26-08-01	2.32 9719	7.96 4906	1.92 8690	7.9721 02	5.1182 55	3.48 9014	6.43 3638	0.85 3906
20								
26-09-01	2.29 9609	7.87 7363	1.90 8375	7.9018 38	5.0945 33	3.41 4301	6.34 3115	0.83 8510
20								
26-10-01	2.26 4756	7.76 4409	1.88 7017	7.7808 96	5.0538 79	3.32 7699	6.24 6045	0.81 7086
20								

26-11-01	2.24	7.73	1.87	7.7727	5.0879	3.28	6.15	0.81
20	5865	9988	2337	56	08	3715	3366	4691
26-12-01	2.23	7.74	1.85	7.7830	5.1351	3.25	6.06	0.81
20	1486	3485	9817	95	65	4165	9212	7134
26-01-20	2.20	7.72	1.84	7.7253	5.1444	3.20	5.98	0.80
01-01-20	9937	0462	4723	66	05	8548	5384	9864
26-02-01	2.20	7.79	1.83	7.8105	5.2328	3.21	5.91	0.82
02-01-20	8209	3016	7998	25	29	5709	9314	6126
26-03-01	2.21	7.93	1.84	8.0962	5.3540	3.28	5.87	0.86
03-01-20	8307	8334	0240	30	93	7471	1880	0627
26-04-01	2.21	8.01	1.83	8.2707	5.4051	3.32	5.83	0.87
04-01-20	6400	3294	8407	49	12	8142	2709	8045
26-05-01	2.22	8.14	1.84	8.5274	5.4861	3.39	5.80	0.90
05-01-20	2658	3649	1141	56	56	6712	2145	4970
26-06-01	2.22	8.30	1.84	8.7362	5.5459	3.44	5.79	0.92
06-01-20	4884	6944	3930	14	31	5982	0480	2187

Gambar 6. Hasil Prediksi Tren Pencarian Batik Daerah Periode Juli 2026 – Juni 2027

Berdasarkan Gambar 6, setiap jenis batik menunjukkan karakteristik tren pencarian yang berbeda selama periode prediksi. Secara umum, nilai prediksi berada pada kisaran yang relatif stabil tanpa mengalami perubahan yang sangat signifikan. Hal ini menunjukkan bahwa model LSTM memprediksi pola pencarian pada periode mendatang masih mengikuti kecenderungan historis yang telah dipelajari dari data Google Trends.

Pada Batik Solo, nilai prediksi berada pada kisaran 7,72 hingga 8,31, dengan kecenderungan meningkat secara bertahap menjelang akhir periode prediksi. Pola tersebut menunjukkan bahwa Batik Solo diperkirakan tetap menjadi salah satu jenis batik dengan tingkat pencarian yang relatif tinggi dibandingkan jenis batik lainnya.

Hasil serupa juga terlihat pada Batik Pekalongan yang memiliki nilai prediksi berkisar antara 7,72 hingga 8,74. Meskipun pada tahap evaluasi model Batik Pekalongan menghasilkan tingkat kesalahan yang relatif tinggi, hasil prediksi menunjukkan kecenderungan peningkatan

indeks pencarian hingga akhir periode pengamatan.

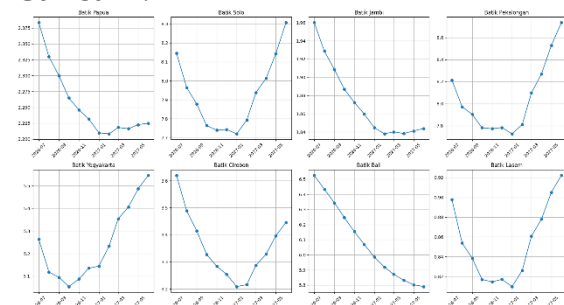
Sementara itu, Batik Yogyakarta diprediksi memiliki nilai indeks pencarian yang berada pada kisaran 5,05 hingga 5,55, sedangkan Batik Bali berada pada kisaran 5,79 hingga 6,53. Kedua jenis batik tersebut menunjukkan pola perubahan yang relatif stabil tanpa fluktuasi yang besar.

Untuk Batik Papua, Batik Jambi, Batik Cirebon, dan Batik Lasem, nilai prediksi berada pada kisaran yang lebih rendah dibandingkan jenis batik lainnya. Meskipun demikian, keempat jenis batik tersebut tetap menunjukkan pola perubahan yang konsisten selama periode prediksi sehingga dapat digunakan sebagai dasar dalam proses pengelompokan menggunakan metode *K-Means*

Clustering.

Secara keseluruhan, hasil prediksi menunjukkan adanya perbedaan tingkat indeks pencarian antarjenis batik. Perbedaan tersebut mengindikasikan bahwa minat masyarakat terhadap masing-masing jenis batik tidak berada pada tingkat yang sama. Oleh karena itu, diperlukan proses pengelompokan untuk mengidentifikasi kelompok batik yang memiliki karakteristik tren pencarian yang serupa.

Untuk mempermudah interpretasi hasil prediksi, nilai prediksi selama periode Juli 2026 hingga Juni 2027 divisualisasikan dalam bentuk grafik garis untuk setiap jenis batik. Visualisasi ini bertujuan untuk menunjukkan kecenderungan perubahan indeks pencarian yang dihasilkan oleh model LSTM pada masing-masing jenis batik. Hasil visualisasi ditunjukkan pada Gambar 7.



Gambar 4. Visualisasi Prediksi Tren Pencarian Batik Daerah Periode Juli 2026–Juni 2027

Berdasarkan Gambar 7, terlihat bahwa setiap jenis batik memiliki pola perkembangan tren pencarian yang berbeda. Batik Solo dan Batik Pekalongan menunjukkan nilai prediksi yang relatif tinggi selama periode pengamatan. Meskipun mengalami sedikit penurunan pada beberapa bulan awal, kedua jenis batik tersebut kembali menunjukkan tren peningkatan hingga akhir periode prediksi. Kondisi ini mengindikasikan bahwa minat masyarakat terhadap kedua jenis batik diperkirakan tetap tinggi pada tahun berikutnya.

Batik Yogyakarta dan Batik Bali menunjukkan pola yang relatif stabil dengan perubahan yang tidak terlalu tajam. Batik Yogyakarta mengalami sedikit peningkatan secara bertahap pada akhir periode prediksi, sedangkan Batik Bali menunjukkan kecenderungan menurun secara perlahan meskipun tetap berada pada tingkat pencarian yang cukup tinggi dibandingkan beberapa jenis batik lainnya.

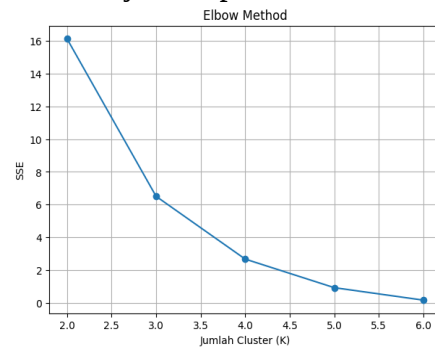
Sementara itu, Batik Papua, Batik Jambi, Batik Cirebon, dan Batik Lasem memiliki nilai prediksi yang lebih rendah dibandingkan jenis batik lainnya. Keempat jenis batik tersebut memperlihatkan perubahan yang relatif kecil selama periode prediksi, sehingga menunjukkan pola pencarian yang cenderung stabil tanpa adanya lonjakan indeks pencarian yang signifikan.

Secara keseluruhan, hasil visualisasi menunjukkan bahwa model LSTM mampu menghasilkan prediksi yang konsisten terhadap pola historis data Google Trends. Perbedaan tingkat indeks pencarian antarjenis batik pada hasil prediksi ini selanjutnya menjadi dasar dalam proses pengelompokan menggunakan metode *K-Means Clustering* untuk mengidentifikasi kelompok batik dengan karakteristik tren pencarian yang serupa.

Sebelum algoritma *K-Means* diterapkan, dilakukan penentuan jumlah

cluster menggunakan Elbow Method. Pada penelitian ini, pengujian dilakukan dengan mencoba jumlah *cluster* sebanyak 2 hingga 6 ($k = 2-6$). Untuk setiap nilai k , dihitung nilai *Sum of Squared Error* (SSE) yang menunjukkan total jarak setiap data terhadap pusat *clusternya* (*centroid*).

Hasil pengujian menggunakan Elbow Method ditunjukkan pada Gambar 8.



Gambar 5. Hasil Elbow Method untuk Penentuan Jumlah Cluster

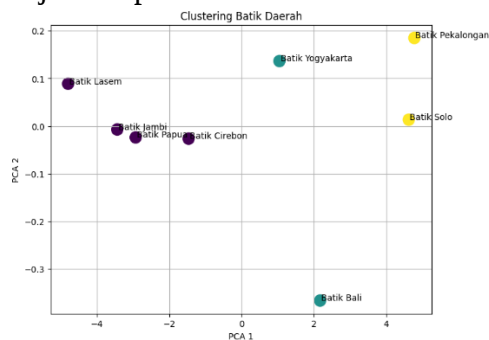
Berdasarkan Gambar 6, nilai SSE mengalami penurunan yang cukup besar hingga $k = 3$, kemudian penurunannya mulai melandai pada jumlah *cluster* berikutnya. Pola tersebut menunjukkan adanya titik siku (*elbow point*) pada $k = 3$, sehingga jumlah *cluster* yang digunakan dalam penelitian ini ditetapkan sebanyak tiga *cluster*. Pemilihan tiga *cluster* dianggap mampu memberikan keseimbangan antara tingkat homogenitas data dalam setiap kelompok dan jumlah kelompok yang terbentuk.

Setelah jumlah *cluster* ditentukan, algoritma *K-Means* diterapkan untuk mengelompokkan delapan jenis batik berdasarkan hasil prediksi tren pencarian selama dua belas bulan.

Hasil proses *K-Means* menunjukkan bahwa delapan jenis batik berhasil dikelompokkan ke dalam tiga *cluster* yang memiliki karakteristik tren pencarian yang berbeda. Untuk mempermudah interpretasi hasil pengelompokan, data divisualisasikan menggunakan metode Principal Component Analysis (PCA). Metode PCA digunakan untuk mereduksi dimensi data menjadi dua komponen utama sehingga posisi masing-masing jenis batik dapat divisualisasikan dalam bidang dua dimensi.

tanpa menghilangkan karakteristik utama data.

Visualisasi hasil *clustering* ditunjukkan pada Gambar 9.



Gambar 6. Visualisasi Hasil K-Means Clustering Menggunakan PCA

Berdasarkan Gambar 9, terlihat bahwa delapan jenis batik berhasil dipisahkan ke dalam tiga kelompok sesuai dengan hasil pengelompokan menggunakan algoritma *K-Means*. Setiap titik pada grafik merepresentasikan satu jenis batik, sedangkan warna yang berbeda menunjukkan *cluster* yang dihasilkan. Posisi antar titik menunjukkan tingkat kemiripan pola tren pencarian hasil prediksi selama dua belas bulan. Semakin dekat posisi dua titik, semakin tinggi tingkat kemiripan pola prediksi yang dimiliki kedua jenis batik tersebut. Sebaliknya, titik yang berjauhan menunjukkan adanya perbedaan karakteristik tren pencarian.

Hasil visualisasi menunjukkan bahwa Batik Solo dan Batik Pekalongan berada pada kelompok yang sama dan terpisah dari kelompok lainnya. Hal ini menunjukkan bahwa kedua jenis batik tersebut memiliki pola tren pencarian yang relatif mirip dan memiliki tingkat indeks pencarian yang lebih tinggi dibandingkan jenis batik lainnya. Sementara itu, Batik Yogyakarta dan Batik Bali membentuk kelompok tersendiri yang menunjukkan karakteristik tren pencarian pada tingkat menengah. Adapun Batik Papua, Batik Jambi, Batik Cirebon, dan Batik Lasem berada pada kelompok lain dengan karakteristik tren pencarian yang relatif lebih rendah.

Visualisasi menggunakan PCA memberikan gambaran yang lebih mudah dipahami mengenai hasil pengelompokan

yang dilakukan oleh algoritma *K-Means*. Meskipun PCA tidak digunakan dalam proses pembentukan *cluster*, metode ini membantu menunjukkan bahwa objek yang berada pada *cluster* yang sama memiliki karakteristik yang serupa, sedangkan objek yang berada pada *cluster* yang berbeda memiliki karakteristik yang berbeda. Dengan demikian, hasil visualisasi mendukung bahwa proses *clustering* telah berhasil mengelompokkan jenis batik berdasarkan kemiripan pola tren pencarian hasil prediksi model LSTM. Selanjutnya, untuk memberikan makna terhadap masing-masing *cluster*, dilakukan perhitungan rata-rata nilai prediksi pada setiap *cluster*. Hasil perhitungan tersebut digunakan sebagai dasar dalam pemberian kategori Tinggi, Sedang, dan Rendah.

Tabel 8. Rata-rata Nilai Prediksi Setiap Cluster

No	Cluster	Rata - rata
1	0	2.083966
2	1	5.662239
3	2	7.989260

Tabel 5 menunjukkan rata-rata nilai prediksi dari setiap *cluster* yang dihasilkan oleh algoritma *K-Means*. Nilai rata-rata tersebut diperoleh dari rata-rata seluruh hasil prediksi pada masing-masing *cluster* selama periode Juli 2026 hingga Juni 2027. Perhitungan rata-rata dilakukan untuk memberikan interpretasi terhadap karakteristik setiap *cluster* sehingga dapat diberi label tingkat tren pencarian.

Berdasarkan Tabel 5, *Cluster 2* memiliki rata-rata nilai prediksi tertinggi sebesar 7.989260, sehingga diinterpretasikan sebagai kelompok dengan tren pencarian tinggi. Selanjutnya, *Cluster 1* memiliki rata-rata sebesar 5.662239, yang menunjukkan tingkat tren pencarian sedang. Adapun *Cluster 0* memiliki rata-rata terendah, yaitu 2.083966, sehingga dikategorikan sebagai kelompok dengan tren pencarian rendah. Hasil tersebut menunjukkan adanya perbedaan karakteristik yang cukup jelas di antara ketiga *cluster* yang terbentuk.

Tabel 9. Hasil kategori Tinggi–Sedang–Rendah

Batik	Cluster	Kategori
Batik Papua	0	Rendah
Batik Solo	2	Tinggi
Batik Jambi	0	Rendah
Batik Pekalongan	2	Tinggi
Batik Yogyakarta	1	Sedang
Batik Cirebon	0	Rendah
Batik Bali	1	Sedang
Batik Lasem	0	Rendah

Berdasarkan hasil pengelompokan pada Tabel 6, diperoleh tiga kategori tren pencarian, yaitu tinggi, sedang, dan rendah. Pengelompokan tersebut didasarkan pada hasil algoritma *K-Means* yang telah diinterpretasikan menggunakan rata-rata nilai prediksi setiap *cluster*.

Kategori tinggi terdiri atas Batik Solo dan Batik Pekalongan. Kedua jenis batik tersebut berada pada *Cluster* 2 yang memiliki rata-rata nilai prediksi tertinggi, yaitu 7.989260. Hasil ini menunjukkan bahwa Batik Solo dan Batik Pekalongan diprediksi memiliki tingkat pencarian yang paling tinggi selama periode Juli 2026 hingga Juni 2027. Tingginya indeks pencarian tersebut mengindikasikan bahwa kedua jenis batik masih memiliki daya tarik yang kuat di kalangan masyarakat dan berpotensi menjadi fokus dalam pengembangan produk maupun strategi pemasaran berbasis digital.

Kategori sedang terdiri atas Batik Yogyakarta dan Batik Bali yang berada pada *Cluster* 1 dengan rata-rata nilai prediksi sebesar 5.662239. Kedua jenis batik tersebut diprediksi memiliki tingkat pencarian yang relatif stabil pada tingkat menengah. Meskipun tidak setinggi kategori tinggi, hasil ini menunjukkan bahwa Batik Yogyakarta dan Batik Bali masih memiliki potensi untuk terus dikembangkan melalui inovasi produk, promosi digital, maupun penguatan identitas budaya daerah.

Sementara itu, kategori rendah terdiri atas Batik Papua, Batik Jambi, Batik

Cirebon, dan Batik Lasem yang berada pada *Cluster* 0 dengan rata-rata nilai prediksi sebesar 2.083966. Nilai tersebut menunjukkan bahwa keempat jenis batik diprediksi memiliki tingkat pencarian yang relatif lebih rendah dibandingkan kelompok lainnya. Hasil ini tidak menunjukkan bahwa kualitas atau nilai budaya batik tersebut lebih rendah, tetapi mengindikasikan bahwa tingkat perhatian masyarakat berdasarkan data pencarian Google Trends masih relatif kecil. Oleh karena itu, diperlukan upaya peningkatan promosi, digitalisasi pemasaran, serta pelestarian budaya untuk meningkatkan eksposur dan minat masyarakat terhadap batik-batik tersebut.

Secara keseluruhan, hasil *clustering* menunjukkan bahwa metode *K-Means* mampu mengelompokkan jenis batik berdasarkan kemiripan pola tren pencarian hasil prediksi model LSTM. Integrasi kedua metode tersebut memberikan informasi yang tidak hanya menggambarkan prediksi tingkat pencarian pada masa mendatang, tetapi juga menghasilkan segmentasi batik berdasarkan karakteristik tren pencariannya. Informasi ini dapat dimanfaatkan oleh pelaku UMKM, pengrajin batik, pemerintah, maupun akademisi sebagai dasar dalam menyusun strategi promosi, pengembangan produk, dan pelestarian batik daerah secara lebih terarah.

SIMPULAN

Penelitian ini berhasil menerapkan metode *Long Short-Term Memory* (LSTM) untuk memprediksi tren pencarian delapan jenis batik daerah di Indonesia berdasarkan data Google Trends periode 2009–2026. Model LSTM menunjukkan kemampuan prediksi yang cukup baik dengan nilai evaluasi keseluruhan sebesar MAE 1,433845, RMSE 2,045695, dan MAPE 29,658822%. Hasil tersebut menunjukkan bahwa model mampu mengikuti pola historis data dan menghasilkan prediksi tren

pencarian untuk periode Juli 2026 hingga Juni 2027.

Hasil prediksi selanjutnya digunakan sebagai masukan pada metode K-Means Clustering untuk mengelompokkan jenis batik berdasarkan kemiripan pola tren pencariannya. Berdasarkan hasil clustering, diperoleh tiga kategori tren pencarian, yaitu tinggi, sedang, dan rendah. Kategori tinggi terdiri atas Batik Solo dan Batik Pekalongan, kategori sedang terdiri atas Batik Yogyakarta dan Batik Bali, sedangkan kategori rendah terdiri atas Batik Papua, Batik Jambi, Batik Cirebon, dan Batik Lasem. Pengelompokan tersebut memberikan gambaran mengenai tingkat perhatian masyarakat terhadap masing-masing jenis batik berdasarkan hasil prediksi tren pencarian.

Secara keseluruhan, integrasi metode LSTM dan *K-Means Clustering* mampu memberikan informasi mengenai prediksi perkembangan tren pencarian sekaligus mengelompokkan jenis batik berdasarkan karakteristik tren pencariannya. Hasil penelitian ini diharapkan dapat menjadi referensi bagi pelaku UMKM, pengrajin batik, pemerintah, dan akademisi dalam menyusun strategi promosi, pengembangan produk, serta pelestarian batik daerah berbasis data digital.

DAFTAR PUSTAKA

- Alifa, T. T., Astuti, R., & Basysyar, F. M. (2024). Implementasi data mining menggunakan algoritma K-Means clustering dalam analisis penjualan produk. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(1), 602–607.
- f R. B., Nastiti, F. E., & Sumarlinda, S. (2023). Algoritma K-Means clustering untuk segmentasi pelanggan (Studi kasus: Fashion Viral Solo). *Infotech Journal*, 9(1), 124–131.
- Avinash, A., Widjaja, A., & Karnalim, O. (2024). Analisis perbandingan algoritma machine learning untuk forecasting persediaan produk barang pokok. *Jurnal Teknik Informatika dan Sistem Informasi*, 10(2), 361–378.
- Juliyanto, R. M., Patria, F., Pamungkas, H. D., & Nugraha, F. S. (2024, November 23). Penerapan algoritma clustering untuk segmentasi pelanggan berdasarkan perilaku pembelian. *Prosiding Seminar Nasional AMIKOM Surakarta (SEMNAS)* 2024, 117–127. AMIKOM Surakarta.
- Kharis, S. A., Zili, A. H., Putri, A., & Robiansyah, A. (2023, Oktober). Analisis Tren Minat Masyarakat Indonesia terhadap Artificial Intelligence dalam Menyongsong Society 5.0: Studi Menggunakan Google Trends. *G-Tech : Jurnal Teknologi Terapan*.
- Krismayani, I., & Mafar, F. (2024). Tren Pencarian Informasi Masyarakat Indonesia. *Jurnal Kajian Ilmu Perpustakaan, Informasi dan Kearsipan*.
- Hakim, L. M. (2018, Juni 17). Batik Sebagai Warisan Budaya Bangsa dan Nation Brand Indonesia. *Nation State: Journal of International Studies*.
- Hasanah, N., & Selasih, D. (2026, Juni). STRATEGI INOVASI PRODUK DAN PEMASARAN BATIK CIWARINGIN UNTUK MASA DEPAN UMKM DAN KOPERASI: DENGAN STUDI KASUS MENGGUNAKAN KUALITATAIF DESKRIPTIF. *Jurnal Ekonomidan Perbankan Syariah*.
- Meriani, A. P., & Rahmatulloh, A. (2024). PERBANDINGAN GATED RECURRENT UNIT (GRU) DAN ALGORITMA LONG SHORTTERM MEMORY (LSTM) LINEAR REFRESSIONDALAM PREDIKSI HARGA EMAS MENGGUNAKAN MODEL TIME SERIES. *JITET (Jurnal Informatika dan Teknik Elektro Terapan)*.
- Maharani, M. F., Fitriani, L. D., Bagus, Y., & Haq,

- D. Z. (2026). PERBANDINGAN LSTM DAN PROPHET DALAM PREDIKSI TRENDS. *Journal of Data Science Theory and Application*.
- Putri, E. S., & Sadikin, M. (2021). Prediksi penjualan produk untuk mengestimasi kebutuhan bahan baku menggunakan perbandingan algoritma LSTM dan ARIMA. *Format Jurnal Ilmiah Teknik Informatika*, 10(2), 162–172.
- Syarif, M. H., Maghfur, N. M., & Dermawan, B. A. (2023). Forecasting fashion product trends using the SARIMA method on Google Trends data. *Jurnal Ilmiah Wahana Pendidikan*, 9(8), 350–359.
- Sarimole, F. M., & Hakim, L. (2024). Klasifikasi Barang Menggunakan Metode Clustering K-Means Dalam. *Jurnal Sains dan Teknologi*, 5 No. 3. doi: <https://doi.org/10.55338/saintek.v5i1.2709>
- Tamami, G., & Arifin, M. (2024). Penggunaan LSTM dalam membangun prediksi penjualan untuk aplikasi Laptop Lens. *Jurnal Fasilkom*, 14(2), 301–308.
- Trilestari, L., Shovmayanti, N. A., Kurniawan, D., & Prakosa, F. A. (2025). PEMANFAATAN GOOGL TRENDS UNTUK MENGANALISIS POLA PENGGUNAAN MEDIA SOSIA DI INDONESIA. *Jurnal Bincang Komunikasi*.
- m/lstm-for-time-series-prediction-in-pytorch/
- Shandy, E. (2024, September 16). Pengaruh media sosial terhadap trend fashion anak muda Indonesia. Kanal24. <https://kanal24.co.id/pengaruh-media-sosial-terhadap-trend-fashion-anak-muda-indonesia/>

Website

- Analytics Vidhya. (2019, August). Comprehensive guide: K-Means clustering. Analytics Vidhya. <https://www.analyticsvidhya.com/blog/2019/08/comprehensive-guide-k-means-clustering/>
- Brownlee, J. (2020). LSTM for time series prediction in PyTorch. *Machine Learning Mastery*. <https://machinelearningmastery.com>