



ANALISIS PREDIKSI KATA KUNCI SITUS WEB MONSTERMAC DENGAN METODE *LONG SHORT-TERM MEMORY* (LSTM)

Fityan Hanif Assalmi¹, Wahyu Syaifullah Jauharis Saputra², Amri Muhaimin³

^{1,2,3}Sains Data, Universitas Pembangunan Nasional Veteran Jawa Timur

Surabaya, Jawa Timur, Indonesia 60294

20083010028@student.upnjatim.ac.id, wahyu.s.j.saputra.if@upnjatim.ac.id, amri.muhamin.stat@upnjatim.ac.id

Abstract

Amid increasingly fierce competition in the digital realm, many companies are striving to increase the number of visitors to their websites. One such competing company is MonsterMAC, a startup. This research aims to provide early warnings and analyze relevant keywords on the MonsterMAC website using the Long Short-Term Memory (LSTM) method. Visitor data from Google Analytics and keyword data from Google Trends for the period July 22, 2022, to July 15, 2023, have been collected and processed through several stages, such as preprocessing, model design, LSTM training, and testing, as well as visualization and interpretation of results. The modeling results show satisfactory performance, with MAE Train Real User = 0.0615, Vending Machine = 0.0218, IoT = 0.0284, Machine Learning = 0.0365, Digital Business = 0.0186, Business Intelligence = 0.0296. Furthermore, this research indicates that the number of visitors is predicted to increase but will also experience a sharp decline in the coming days. The use of the keyword "IoT" shows a significant increasing trend. Implementing the keyword "IoT" in SEO strategies has increased the number of visitors over the next seven days from 250 to 350. This research guides website owners in optimizing their content and SEO strategies to increase their visibility and competitiveness in a highly competitive digital environment. This research also emphasizes the importance of the LSTM method in keyword analysis and prediction to create more targeted SEO strategies.

Keywords: google analytics, google trends, keywords, long short-term memory (LSTM), Search Engine Optimization (SEO)

Abstrak

Dalam tengah persaingan bisnis yang semakin ketat di ranah digital, banyak perusahaan yang berlomba-lomba untuk meningkatkan jumlah pengunjung pada situs web mereka. Salah satu perusahaan yang bersaing adalah MonsterMAC, sebuah perusahaan *startup*. Penelitian ini bertujuan untuk memberikan peringatan dini dan menganalisis kata kunci yang relevan di situs web MonsterMAC menggunakan metode *Long Short-Term Memory* (LSTM). Data pengunjung dari *Google Analytics* dan data kata kunci dari *Google Trends* untuk rentang waktu 22 Juli 2022 hingga 15 Juli 2023 telah dikumpulkan dan diproses melalui beberapa tahapan, seperti *preprocessing*, perancangan model, pelatihan dan pengujian LSTM, serta visualisasi dan interpretasi hasil. Hasil pemodelan menunjukkan kinerja yang memuaskan, dengan MAE Train Real User = 0.0615, Vending Machine = 0.0218, IoT = 0.0284, Machine Learning = 0.0365, Bisnis Digital = 0.0186, Kecerdasan Bisnis = 0.0296. Selain itu, penelitian ini juga menunjukkan bahwa jumlah pengunjung diprediksi akan mengalami peningkatan, namun juga akan mengalami penurunan yang tajam dalam beberapa hari ke depan. Adapun penggunaan kata kunci "IoT" menunjukkan tren peningkatan yang signifikan. Implementasi kata kunci "IoT" dalam strategi SEO terbukti meningkatkan jumlah pengunjung selama tujuh hari berikutnya dari 250 hingga 350. Penelitian ini memberikan panduan bagi pemilik situs web untuk mengoptimalkan strategi konten dan SEO demi meningkatkan visibilitas serta daya saing mereka di dalam lingkungan digital yang sangat kompetitif. Penelitian ini juga menekankan pentingnya metode LSTM dalam analisis dan prediksi kata kunci guna menciptakan strategi SEO yang lebih tepat sasaran.

Kata kunci: google analytics, google trends, kata kunci, long short-term memory (LSTM), Search Engine Optimization (SEO)

1. PENDAHULUAN

Jurnal Transformasi digital telah mengubah cara perusahaan beroperasi dan berinteraksi dengan pelanggan. Perusahaan-perusahaan saat ini semakin mengandalkan teknologi untuk

meningkatkan efisiensi dan memperluas jangkauan mereka. Penelitian [1] bertujuan memahami sentimen pelanggan menggunakan pembelajaran mesin. Sebagai bagian dari transformasi ini, perusahaan *startup* teknologi seperti

MonsterMAC yang berfokus pada layanan *vending machine*, memainkan peran penting dalam menghadirkan inovasi baru dan memperkaya pengalaman konsumen [2]. Dalam penelitian [3] metode *Long Short-Term Memory* (LSTM) digunakan untuk menganalisis dan memprediksi tren COVID-19 berdasarkan data dari 22 Januari 2020 hingga 12 Februari 2022, menghasilkan hasil yang memuaskan. Oleh karena itu, peneliti saat ini berusaha menerapkan metode LSTM pada *dataset* yang lebih pendek, yaitu satu tahun, untuk kasus yang berbeda, yaitu pola kunjungan pengguna MonsterMAC.

Bagaimana penggunaan Metode *Long Short-Term Memory* (LSTM) dapat membantu MonsterMAC meningkatkan lalu lintas situs webnya, apa saja faktor-faktor yang mempengaruhi visibilitas situs web MonsterMAC, bagaimana cara mengatasinya dan bagaimana strategi SEO yang dapat diterapkan berdasarkan hasil analisis pola kunjungan pengguna menggunakan LSTM.

Penelitian sebelumnya yang dilakukan oleh [4] menunjukkan bahwa peringatan dini menggunakan LSTM dapat digunakan untuk mengambil keputusan yang akurat. Hasil penelitian tersebut sesuai dengan tujuan penelitian ini yaitu memberikan peringatan dini kepada MonsterMAC dan menganalisis serta memprediksi kata kunci yang paling relevan pada situs web MonsterMAC menggunakan Metode *Long Short-Term Memory* (LSTM). Selain itu, penelitian ini bertujuan untuk menunjukkan bahwa metode LSTM dapat digunakan untuk memprediksi data deret waktu dalam jangka waktu satu tahun dan memberikan rekomendasi strategi SEO yang lebih tepat sasaran berdasarkan hasil analisis.

Penelitian [5] bertujuan untuk memprediksi 7 hari ke depan, serupa dengan penelitian ini yang juga berfokus pada peramalan 7 hari ke depan, namun berbeda dalam hal data dan metode yang digunakan. Penelitian ini menitikberatkan pada analisis dan prediksi pola kunjungan pengguna serta kata kunci yang sering digunakan pada situs web MonsterMAC menggunakan Metode *Long Short-Term Memory* (LSTM). Pembatasan penelitian mencakup penggunaan data historis kunjungan pengguna dari periode 22 Juli 2022 hingga 22 Juli 2023 sebagai dasar analisis, tanpa mempertimbangkan faktor eksternal seperti perubahan industri atau tren pasar yang tidak langsung berhubungan dengan aktivitas *online* situs web. Selain itu, data kata kunci dari *Google Trends* untuk periode 22 Juli 2022 hingga 2023 yang sering digunakan oleh MonsterMAC meliputi *vending machine*, IoT, *machine learning*, bisnis digital, dan kecerdasan buatan.

2. METODE PENELITIAN

2.1 Metode pengumpulan data

Pengumpulan data adalah langkah awal dalam penelitian ini di mana data terkait kunjungan pengguna pada situs web MonsterMAC dikumpulkan. Sumber data utamanya berasal

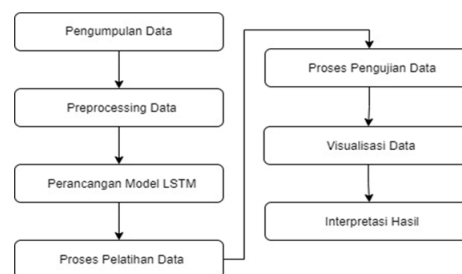
dari *google analytics* dan tujuannya adalah untuk memperoleh dataset yang mencakup informasi relevan seperti jumlah pengunjung, durasi kunjungan, dan pola perilaku pengguna lainnya. Dan untuk pengumpulan data kata kunci diperoleh dari *google trends* [6]. Berikut adalah tabel 1, yakni *Date* adalah Tanggal data tersebut didapatkan, di mana *User* adalah seseorang yang mengunjungi situs MonsterMAC beberapa kali selama periode tersebut, mereka dihitung sebagai satu pengguna, *New user* adalah pengunjung baru yang belum pernah mengunjungi situs MonsterMAC sebelumnya, *Real user* adalah jumlah dari *user* dan *new user*, *Keyword* adalah kata kunci yang digunakan pada tanggal tersebut, *Vending Machine*, IoT, *Machine Learning*, Kecerdasan Buatan Bisnis Digital adalah data kata kunci yang sering digunakan situs MonsterMAC dan akan digunakan sebagai batasan masalah untuk dianalisis

Tabel 1. Deskripsi data

No	Nama Variabel	Deskripsi variabel
1	<i>Date</i>	Tanggal data didapatkan
2	<i>User</i>	Pengunjung yang sudah pernah mengunjungi situs web MonsterMAC
3	<i>New User</i>	Pengunjung yang baru pertama kali ke situs web MonsterMAC
4	<i>Real User</i>	Jumlah dari <i>user</i> dan <i>new user</i>
5	<i>Keyword</i>	Kata kunci yang digunakan MonsterAC pada tanggal tersebut
6	<i>Vending Machine</i> , <i>Machine Learning</i> , IoT, Kecerdasan Buatan, Bisnis Digital.	Data kata kunci yang akan dianalisis

2.2 Tahapan penelitian

Secara garis besar metode penelitian adalah tahapan atau langkah kegiatan dalam penelitian sehingga data bisa diolah dan dianalisis. Tahap-tahap tersebut dapat dilihat pada gambar 1.



Gambar 1. Flowchart Penelitian

2.2.1. Pengumpulan Data

Pengumpulan data merupakan langkah awal dalam penelitian ini di mana pengumpulan data tersebut menggunakan dari dua sumber.

a) Data dari *Google Analytics* MonsterMAC

Data terkait kunjungan pengguna pada situs web MonsterMAC dikumpulkan. Sumber data berasal dari *Google Analytics* yang sudah disambungkan dengan situs web MonsterMAC. Tujuannya adalah untuk memperoleh *dataset* yang mencakup informasi relevan seperti jumlah pengunjung, durasi kunjungan, dan pola perilaku pengguna lainnya. Pengumpulan data yang komprehensif dan akurat sangat penting untuk memastikan analisis yang tepat dan hasil yang valid. Penelitian [7] memanfaatkan alat analisis web untuk manajemen pengetahuan yang efektif

b) Data kata kunci dari *Google Trends*

Data terkait pencarian kata kunci oleh seluruh pengguna *google*. Tujuannya untuk memperoleh *dataset* yang mencakup informasi berapa banyak pengguna yang mencari kata kunci pada hari tertentu. Penelitian [8] menggunakan *Google Trends* untuk menyajikan data mengenai seberapa sering kata kunci digunakan oleh pengguna dalam pencarian web di mesin pencari Google.

2.2.2 Preprocessing

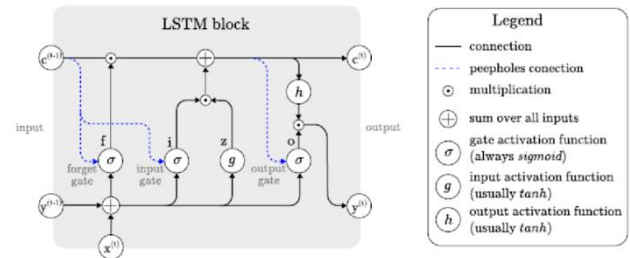
Preprocessing data melibatkan serangkaian langkah untuk membersihkan, memformat, dan mempersiapkan data sebelum digunakan untuk analisis [9]. Langkah-langkah ini mungkin mencakup penanganan nilai-nilai yang hilang, normalisasi data, atau transformasi data agar sesuai dengan persyaratan model LSTM. *Preprocessing* data yang baik akan membantu memastikan bahwa data yang digunakan untuk analisis adalah data yang berkualitas dan dapat diandalkan. [10]

2.2.3 Perancangan Model LSTM

Perancangan model LSTM melibatkan pembuatan struktur model *Long Short-Term Memory* (LSTM). Ini termasuk penentuan jumlah layer, konfigurasi *node*, dan pengaturan parameter lainnya. Model LSTM digunakan untuk memodelkan dan memahami pola dalam data deret waktu yang dihasilkan oleh kunjungan pengguna. Seperti penelitian [11] menggunakan model LSTM untuk meramalkan data seret waktu. Perancangan model yang tepat akan mempengaruhi kinerja dan akurasi model dalam memprediksi pola kunjungan pengguna.

Gambar 2 mengilustrasikan struktur pemrosesan data dalam LSTM, yang merupakan pengembangan dari algoritma RNN. Berbeda dengan arsitektur RNN standar yang terdiri dari jaringan modul berulang dengan fungsi layer sederhana, jaringan LSTM dideskripsikan sebagai beberapa modul berantai yang berulang, di mana setiap modul memiliki sub-modul yang berisi tiga fungsi gerbang

sigmoid. Fungsi *sigmoid* ini mengatur seberapa banyak informasi yang diteruskan. Terdapat tiga gerbang utama dalam sub-modul LSTM, yaitu *forget gate*, *input gate*, dan *output gate*. Penelitian [12] juga menjelaskan fungsi penting dari ketiga gerbang tersebut. Beberapa arsitektur LSTM yang dikembangkan, seperti yang terlihat pada gambar 2 biasanya memiliki tambahan satu gerbang lagi, yaitu *input modulation gate*. Secara umum, arsitektur LSTM dapat dibagi menjadi empat modul utama, yaitu *input gate* untuk membawa data masukan dari luar dan memprosesnya, *input modulation gate*, *forget gate* untuk menentukan data yang akan dihapus dan menentukan jeda waktu optimal untuk masukan berikutnya, dan *output gate* untuk memproses semua perhitungan masukan dan menghasilkan keluaran dari sel LSTM.



Gambar 2. Arsitektur LSTM

Dalam proses perhitungan LSTM, *input* data yang berbentuk *time-series* dimisalkan sebagai

$x = (x_1, x_2, \dots, x_n)$ dan *output* $y = (y_1, y_2, \dots, y_n)$ kemudian selanjutnya proses komputasi dilakukan dengan beberapa rumus sebagai berikut :

$$h_t = H (W_{hx} X_t + W_{hh} h_{t-1} + b_h) \quad (1)$$

$$P_t = W_{hy} Y_{t-1} + b_y \quad (2)$$

Di mana W adalah bobot dari matriks dan b adalah bias. Selanjutnya perhitungan *hidden state* dilakukan dengan rumus:

$$i_t = \sigma(W_{ix}x_t + W_{ih}h_{t-1} + W_{ic}c_{t-1} + b_i) \quad (3)$$

$$f_t = \sigma(W_{fx}x_t + W_{fh}h_{t-1} + W_{fc}c_{t-1} + b_f) \quad (4)$$

$$c_t = f_t * c_{t-1} + i_t * g(W_{cx}x_t + W_{ch}h_{t-1} + W_{cc}c_{t-1} + b_c) \quad (5)$$

$$o_t = \sigma(W_{ox}x_t + W_{oh}h_{t-1} + W_{oc}c_{t-1} + b_o) \quad (6)$$

$$h_t = o_t * h(c_t) \quad (7)$$

Di mana * merupakan hasil perkalian antara 2 matriks, sedangkan g dan h merupakan fungsi *sigmoid* dengan rentang nilai antara [-2,2] dan [-1,1]. σ merupakan fungsi *sigmoid* standar dengan rumus :

$$\sigma(x) = \frac{1}{1+e^x} \quad (9)$$

$$= \sum_{t=1}^n (y_t - p_t)^2 \quad (10)$$

2.2.4 Proses Pelatihan Data

Proses pelatihan data adalah langkah di mana model LSTM diberikan *dataset* untuk belajar pola dan hubungan dalam data. Dengan menggunakan metode pembelajaran mesin, model di latih dengan data historis sehingga dapat membuat prediksi yang akurat pada data yang belum pernah dilihat sebelumnya. Penelitian [13] menyarankan agar bentuk sebaran data asli tetap dipertahankan dan informasi yang ada dalam data tersebut tidak diubah. Proses pelatihan yang baik akan menghasilkan model yang lebih baik dalam memprediksi pola kunjungan pengguna di masa mendatang.

2.2.5 Proses Pengujian Data

Setelah model LSTM dilatih, proses pengujian data melibatkan penggunaan *dataset* terpisah yang tidak digunakan selama pelatihan untuk mengevaluasi kinerja model seperti yang disarankan oleh penelitian [14]. Hal ini membantu memastikan bahwa model dapat menggeneralisasi pola dengan baik pada data baru dan tidak hanya mempelajari data pelatihan. Pengujian data yang baik akan memberikan kepercayaan yang lebih besar terhadap model dalam memprediksi pola kunjungan pengguna secara akurat [15].

2.2.6 Visualisasi Data

Visualisasi data melibatkan pembuatan grafik, plot, atau visualisasi lainnya untuk menggambarkan hasil analisis dan prediksi. Dalam konteks ini, mungkin termasuk grafik waktu nyata kunjungan pengguna, perbandingan antara

prediksi model dan data aktual, atau visualisasi lainnya yang dapat membantu pemahaman dan interpretasi hasil. Visualisasi data yang baik akan memudahkan pemahaman terhadap hasil analisis dan memperkuat kesimpulan yang diambil dari data.

2.2.7 Interpretasi Hasil

Interpretasi hasil merupakan tahap penting dalam penelitian yang melibatkan proses analisis, penjelasan, dan pemberian makna pada data yang telah dikumpulkan selama penelitian. Tahap ini membantu peneliti untuk memahami implikasi temuan, mengaitkannya dengan pertanyaan penelitian, dan menyusun kesimpulan yang didasarkan pada data yang telah dianalisis. Interpretasi hasil yang baik akan memberikan wawasan yang berharga dan dapat dijadikan dasar untuk membuat keputusan atau tindakan selanjutnya.

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Tabel 2 berisi informasi mengenai kunjungan pengguna ke situs web MonsterMAC, yang didapatkan dari data *Google Analytics*. Data ini juga mencakup kata kunci yang sering digunakan dalam strategi SEO MonsterMAC, seperti '*vending machine*', 'IoT', '*machine learning*', 'bisnis digital', dan 'kecerdasan buatan'. Data kata kunci diperoleh dari *Google Trends*. Terdapat total 369 entri dengan 9 kolom, dan rentang waktu pengamatan data adalah dari 20 Juli 2022 hingga 15 Juli 2023.

Tabel 2 Data pengunjung dan data kata kunci yang sering digunakan MonsterMAC

Date	User	New User	Real User	Vending Machine	IoT	Machine Learning	Bisnis Digital	Kecerdasan Buatan
07/20/2022	104	96	200	40	69	75	46	35
07/21/2022	73	66	139	50	68	75	39	30
07/22/2022	63	61	124	40	70	75	35	25
...
07/14/2023	190	169	359	0	62	58	48	38
07/15/2023	164	150	314	0	62	54	49	39

Tabel 2 menggambarkan data harian untuk sejumlah variabel yang relevan dalam konteks bisnis atau aktivitas teknologi. Setiap baris dalam tabel mewakili satu hari, dimulai dari 20 Juli 2022 hingga 15 Juli 2023. Variabel "*User*" mencatat jumlah total pengguna yang menggunakan layanan atau produk pada hari itu, sedangkan "*New User*" mencatat jumlah pengguna baru yang menggunakan layanan atau produk pada hari tersebut. Variabel "*Real User*" mengindikasikan jumlah pengguna "*User*" dengan "*New User*" pada hari itu. Selain itu, tabel juga mencatat jumlah pencarian untuk kata kunci '*vending machine*', 'IoT', '*machine learning*', 'bisnis digital', dan 'kecerdasan buatan'

di mesin pencari *Google*. Dengan data ini, dapat dianalisis tren harian dalam penggunaan layanan, pertumbuhan pengguna baru, dan adopsi kata kunci yang sering dipakai pengguna *google*.

3.2 Preprocessing

Tahapan *preprocessing* melibatkan fungsi `str_to_datetime` untuk mengubah format string 'bulan/tanggal/tahun' menjadi objek `datetime`. Selain itu, dilakukan konversi tanggal '7/20/2022' menjadi objek `datetime`. Selanjutnya, data dalam kolom 'Date_int' dari DataFrame data diubah formatnya menjadi objek `datetime` menggunakan fungsi

str_to_datetime. Setelah itu, kolom 'Date_int' dijadikan sebagai indeks DataFrame, dan DataFrame data direset indeksnya menjadi indeks integer default. Tahapan ini mempersiapkan data untuk analisis dan pemodelan lebih lanjut. Pada gambar 3 adalah hasil *preprocessing* data yang sudah di lakukan.

```

0 2022-07-20
1 2022-07-21
2 2022-07-22
3 2022-07-23
4 2022-07-24
...
363 2023-07-18
364 2023-07-19
365 2023-07-20
366 2023-07-21
367 2023-07-22
Name: Date_int, Length: 368, dtype: datetime64[ns]
```

Gambar 3. Hasil *Preprocessing*

Selanjutnya, data dinormalisasi menggunakan metode *MinMaxScaler* dari pustaka *scikit-learn* sehingga nilai-nilai berada dalam rentang 0 hingga 1. Pertama, kelas *MinMaxScaler* diimpor dari modul *sklearn.preprocessing*. Kemudian, objek *MinMaxScaler* diinisialisasi dengan menetapkan parameter *feature_range* ke (0, 1) untuk menentukan rentang nilai setelah normalisasi. Proses normalisasi dilakukan dengan metode *fit_transform* yang diterapkan pada *dataset* *dataset_train*. Metode ini menghitung statistik yang diperlukan, seperti nilai minimum dan *maksimum*, dan kemudian menerapkan transformasi skala pada dataset sehingga semua fitur berada dalam rentang 0 hingga 1.

3.3 Perancangan Model LSTM

Penulis merancang dan mengimplementasikan model *Long Short-Term Memory* (LSTM) untuk menganalisis dan memprediksi jumlah pengguna aktual (*Real User*) pada situs web *MonsterMAC*, serta menganalisis kata kunci yang sering digunakan oleh *MonsterMAC*, seperti '*vending machine*', '*iot*', '*machine learning*', '*bisnis digital*', dan '*kecerdasan buatan*'. Proses pengembangan model dimulai dengan inisialisasi model LSTM sebagai model *sequential*, yang memungkinkan penambahan layer secara berurutan. Dua layer LSTM ditambahkan ke dalam model, masing-masing dengan 50 unit neuron. Layer pertama menggunakan *return_sequences=True* karena diikuti oleh layer LSTM lainnya, sedangkan layer kedua menggunakan *return_sequences=False* karena menjadi layer terakhir LSTM dalam rangkaian. Sebuah layer *output* dengan satu neuron ditambahkan untuk melakukan prediksi.

Model LSTM dikompilasi menggunakan *Adam optimizer* dengan *learning rate* 0.001. *Mean Squared Error* (MSE) digunakan sebagai fungsi loss, sementara *Mean Absolute Error* (MAE) dan *Mean Absolute Percentage Error* (MAPE) digunakan sebagai metrik evaluasi. Model dilatih menggunakan data pelatihan, dengan 3% dari data pelatihan

digunakan sebagai data validasi. *Batch size* sebesar 32 digunakan dan model dilatih selama 250 *epochs*.

Hasilnya, model LSTM berhasil dibangun, dikompilasi, dan dilatih untuk menganalisis dan memprediksi jumlah pengguna aktual serta kata kunci yang sering digunakan pada situs web *MonsterMAC*. Ringkasan arsitektur model menunjukkan jumlah parameter yang dapat di-*train*, memberikan gambaran tentang kompleksitas model.

3.4 Proses Pelatihan Data dan Pengujian Data

Dalam penelitian ini, data pelatihan telah dipersiapkan untuk digunakan dalam model *Long Short-Term Memory* (LSTM). Data tersebut awalnya diubah menjadi *array NumPy* untuk memungkinkan manipulasi dan pemrosesan lebih lanjut menggunakan fungsi-fungsi *NumPy*. Setelah itu, data tersebut diubah bentuk agar sesuai dengan format yang diperlukan oleh model LSTM.

Proses konversi dan *reshaping* data pelatihan menghasilkan bentuk *X_train* sebesar (348, 5, 1), menunjukkan terdapat 348 sampel dalam data pelatihan, masing-masing dengan 5 fitur atau langkah waktu, dan setiap langkah waktu hanya memiliki satu fitur. Bentuk dari *y_train* adalah (348, 1), yang berarti terdapat 348 sampel dalam data pelatihan, masing-masing dengan satu nilai target. Transformasi ini penting agar data pelatihan sesuai dengan format yang diperlukan oleh model LSTM, memungkinkan model untuk memahami urutan temporal data dan hubungan antar fitur dengan lebih efektif.

Selanjutnya, data pengujian diubah bentuk agar sesuai dengan format yang dibutuhkan oleh model LSTM. Hasilnya, bentuk dari *X_test* adalah (7, 1, 1), menunjukkan bahwa terdapat 7 sampel dalam data pengujian, setiap sampel hanya memiliki 1 fitur atau langkah waktu, dan setiap langkah waktu hanya memiliki satu fitur. Bentuk dari *y_test* adalah (7, 1), yang mengindikasikan terdapat 7 sampel dalam data pengujian, masing-masing dengan satu nilai target. Transformasi ini memastikan bahwa data pengujian memenuhi format yang dibutuhkan oleh model LSTM, memungkinkan model untuk memahami pola temporal dalam data dan memberikan prediksi yang akurat. Persiapan data pengujian dengan benar sangat penting dalam pengembangan model LSTM untuk prediksi deret waktu, karena dapat mempengaruhi kualitas hasil prediksi.

Model LSTM kemudian dikompilasi menggunakan *optimizer Adam* dan fungsi *loss MAE* (*Mean Absolute Error*). Pada tabel 3 merupakan hasil prediksi model yang didapatkan dari *MAE train*, *RMSE train*.

Tabel 3. Hasil Prediksi Model

No	Objek Penelitian	MAE Train	RMSE Train
1	<i>Real User</i>	0.0615	0.00748
2	<i>Vending Machine</i>	0.0218	0.04488

No	Objek Penelitian	MAE Train	RMSE Train
3	IoT	0.0284	0.03862
4	Machine Learning	0.0365	0.05401
5	Bisnis Digital	0.0186	0.03040
6	Kecerdasan Bisnis	0.0296	0.04205

Berdasarkan analisis performa model menggunakan *Mean Absolute Error* (MAE) dan *Root Mean Square Error* (RMSE) pada data pelatihan dengan berbagai jenis data, ditemukan variasi tingkat akurasi yang signifikan di antara *dataset* yang berbeda. Model yang diterapkan pada data Bisnis Digital menunjukkan performa terbaik dengan nilai MAE sebesar 0.0186 dan RMSE sebesar 0.03040, mengindikasikan tingkat kesalahan prediksi yang sangat rendah. Sebaliknya, model yang diterapkan pada data *Real User* memiliki nilai MAE tertinggi sebesar 0.0615, meskipun nilai RMSE-nya sangat rendah pada angka 0.00748, menunjukkan kesalahan absolut rata-rata yang lebih besar tetapi jarang menyimpang jauh dari nilai sebenarnya.

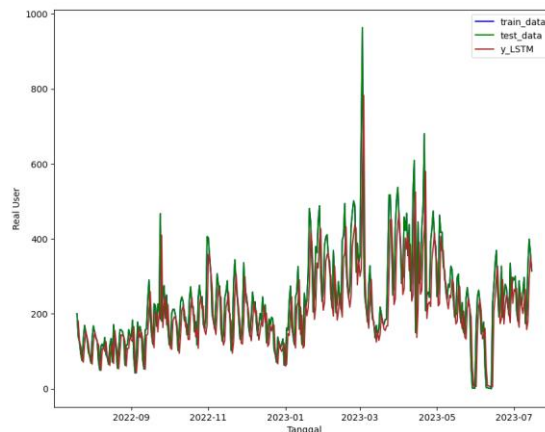
Model yang diterapkan pada data *Vending Machine* dan *Machine Learning* memiliki nilai RMSE yang lebih tinggi dibandingkan MAE, yang mengindikasikan adanya beberapa kesalahan prediksi yang signifikan. Model yang diterapkan pada data IoT dan Kecerdasan Bisnis menunjukkan performa yang cukup baik dengan nilai MAE dan RMSE yang relatif rendah, menandakan akurasi prediksi yang baik dan konsisten. Secara keseluruhan, model yang diterapkan pada data Bisnis Digital menunjukkan performa terbaik, sementara model yang diterapkan pada data *Real User* menunjukkan kesalahan absolut rata-rata terbesar, meskipun konsistensi kesalahan pada model ini memerlukan perhatian lebih lanjut.

Penelitian ini menunjukkan bahwa meskipun model yang digunakan sama, karakteristik kesalahan yang berbeda tergantung pada jenis data yang digunakan. Pemilihan *dataset* yang tepat sangat tergantung pada konteks dan kebutuhan spesifik dari aplikasi yang dimaksud. Analisis lebih lanjut diperlukan untuk memahami distribusi kesalahan dan faktor-faktor yang mempengaruhi performa model pada berbagai jenis data.

3.5 Visualisasi Data

Setelah menyelesaikan semua tahapan analisis menggunakan model *Long Short-Term Memory* (LSTM), langkah berikutnya adalah melakukan visualisasi data guna mempermudah pemahaman. Visualisasi tersebut berguna untuk meningkatkan strategi SEO (*Search Engine Optimization*) dengan memberikan pemahaman yang lebih baik mengenai tren penggunaan kata kunci dan perilaku pengunjung. Dengan pemahaman yang lebih baik ini, strategi SEO dapat dirancang lebih efektif sesuai dengan

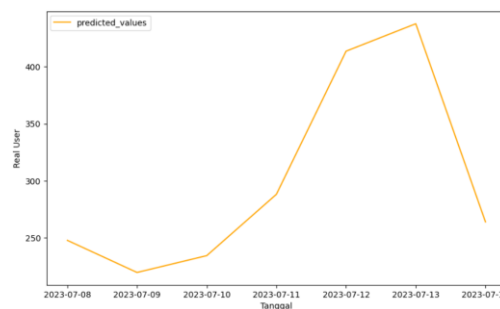
preferensi pengguna, sehingga meningkatkan visibilitas dan peringkat situs web MonsterMAC di mesin pencari. Gambar 4 menampilkan visualisasi data prediksi untuk memverifikasi kesesuaian prediksi dengan data asli.



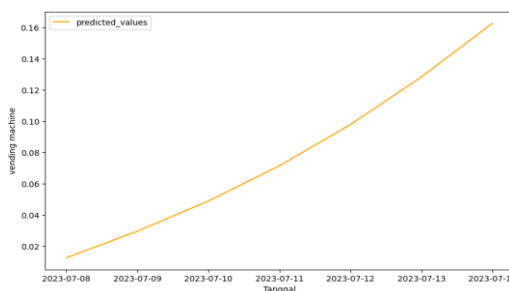
Gambar 4 Visualisasi Prediksi Data

Dari grafik tersebut, dapat disimpulkan bahwa pemodelan LSTM sudah mendekati data asli.

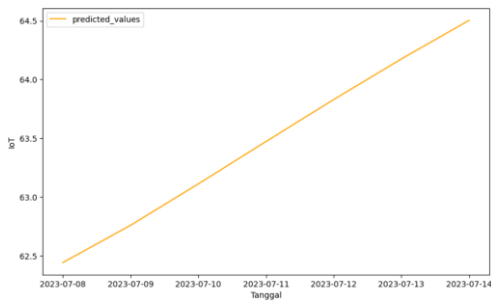
Gambar 5 menampilkan prediksi data kunjungan dan penggunaan kata kunci selama periode tujuh hari, yakni dari tanggal 8 Juli 2023 hingga 14 Juli 2023. Hasil prediksi menunjukkan bahwa jumlah kunjungan awalnya diestimasi akan meningkat, tetapi kemudian mengalami penurunan drastis dari 500 menjadi 275 pada tanggal 13 Juli hingga 14 Juli 2023. Sementara itu, penggunaan kata kunci "IoT" mengalami lonjakan signifikan dari 62 menjadi 64 selama periode yang sama. Hal ini mengindikasikan bahwa topik atau produk yang berkaitan dengan *Internet of Things* (IoT) kemungkinan akan menjadi sorotan utama atau tren dalam pencarian *online* minggu depan.



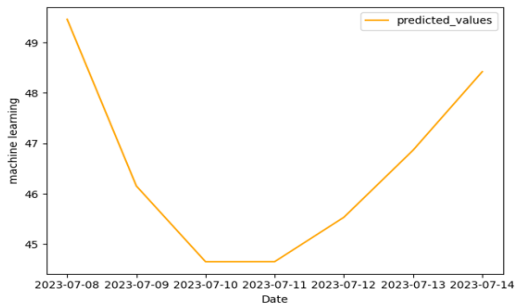
Gambar 5. (a) Visualisasi Prediksi Data Kata Kunci "Real User"



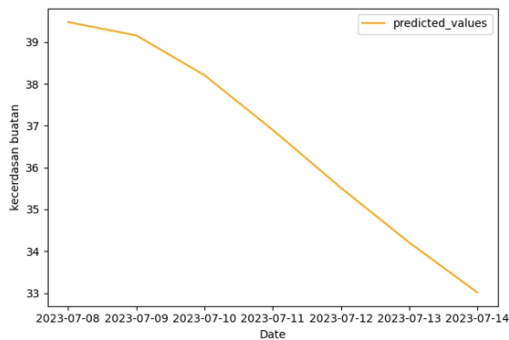
Gambar 5. (b) Visualisasi Prediksi Data Kata Kunci "vending machine"



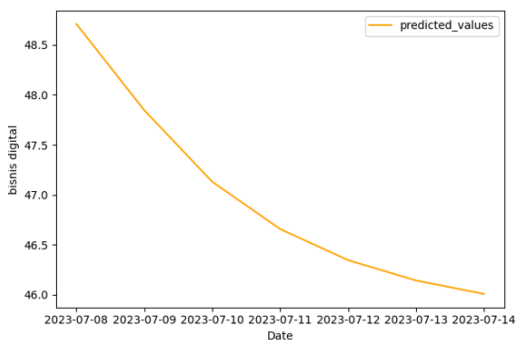
Gambar 5. (c) Visualisasi Prediksi Data Kata Kunci “IoT”



Gambar 5. (d) Visualisasi Prediksi Data Kata Kunci “machine learning”



Gambar 5. (e) Visualisasi Prediksi Data Kata Kunci “kecerdasan buatan”

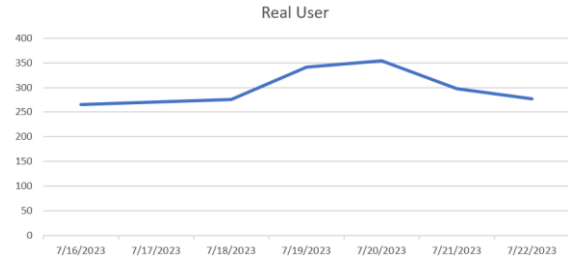


Gambar 5. (f) Visualisasi Prediksi Data Kata Kunci “bisnis digital”

3.6 Pembuktian Model

Untuk menguji analisis kata kunci dengan model LSTM, peneliti memasukkan kata kunci "IoT" selama periode 7 hari sebelumnya, lalu memvisualisasikannya setelah 7 hari. Tujuannya adalah untuk mengetahui apakah prediksi peningkatan kata kunci "IoT" dalam 7 hari ke depan akan terbukti tepat. Penggunaan model LSTM dalam analisis ini memungkinkan pengamatan pola dan tren yang mungkin sulit dideteksi dengan metode analisis konvensional.

Dengan mempertimbangkan data historis, peneliti berharap dapat mengenali potensi kenaikan kata kunci "IoT" dan menggunakan informasi tersebut untuk pengambilan keputusan yang lebih baik dalam strategi pemasaran atau penelitian. Gambar 6 adalah visualisasi data Real User dari hasil pemasukan kata kunci IoT selama 7 hari di situs web MonsterMAC



Gambar 6. Visualisasi Real User

4. KESIMPULAN

Hasil penelitian menggunakan model *Long Short-Term Memory* (LSTM) menunjukkan bahwa jumlah pengunjung situs web MonsterMAC mengalami kenaikan namun kemudian turun drastis dari 500 menjadi 275 pada tanggal 13 hingga 14 Juli 2023. Sebaliknya, penggunaan kata kunci "IoT" mengalami lonjakan dari 62 menjadi 64 selama periode yang sama. Model LSTM berhasil digunakan untuk menganalisis data selama satu tahun dan menunjukkan variasi performa tergantung pada jenis data. Evaluasi model dengan *Mean Absolute Error* (MAE) dan *Root Mean Square Error* (RMSE) menunjukkan bahwa model pada data

Bisnis Digital memiliki performa terbaik dengan MAE sebesar 0.0186 dan RMSE sebesar 0.03040, sementara model pada data *Real User* memiliki nilai MAE tertinggi sebesar 0.0615, meskipun RMSE-nya rendah pada 0.00748. Kesimpulannya, model LSTM memberikan peringatan dini kepada situs web MonsterMAC untuk segera mengganti kata kunci, yang diharapkan dapat membantu dalam mengatasi penurunan pengunjung. Diharapkan hasil analisis ini dapat membantu MonsterMAC dalam meningkatkan visibilitas dan daya saing situs web mereka.

Ucapan Terima Kasih

Penulis F.H.A mengucapkan terima kasih kepada dosen – dosen Sains Data Universitas Pembangunan Nasional veteran Jawa Timur yang telah memberikan masukan untuk penulisan jurnal ini.

DAFTAR PUSTAKA.

- [1] Prabakaran.N, A.Anbarasi, N.Deepa, and Pandiaraja.P, “Enabling an On-demand Access to Community Sentiments using LSTM RNNs Web Service Architecture,” *ScienceDirect*, vol. 230, p. 584, 2023.
- [2] S. T. F. Sirojul Munir, Anifatul Aufah, Kuati Septiani, “ANALISIS DAN RANCANG BANGUN PROTOTYPE WEBMARKET PLACE

- UMKM JUARA,” *J. Teknol. Terpadu*, vol. 6, 2020, [Online]. Available: <https://journal.nurulfikri.ac.id/index.php/JTT/article/view/265/173>
- [3] S. Hansun, I. Charles, and T. Gherman, “The role of the mass vaccination programme in combating the COVID-19 pandemic: An LSTM-based analysis of COVID-19 confirmed cases,” *Heliyon*, vol. 9, no. 14397, p. 3, 2023.
- [4] Z. Lin *et al.*, “Early warning method for power supply service quality based on three-way decision theory and LSTM neural network,” *ScienceDirect*, vol. 8, p. 538, 2022, [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2352484722004917>
- [5] A. T. Damaliana, A. Muhaimin, and P. A. Riyantoko, “Peramalan Lonjakan Kasus Harian Covid-19 Di Indonesia Dengan Model Arima,” *SENADA*, vol. 3, 2023, [Online]. Available: <https://prosiding-senada.upnjatim.ac.id/index.php/senada/article/view/112>
- [6] F. T. Anggraeny, D. F. Rosida, W. S. J. Saputra, and H. Prasetyo, “Kolaborasi Pemasaran Digital Menggunakan Media Sosial dan Marketplace untuk Meningkatkan Produktivitas UMKM,” *J. Nas. Pengabd. Masy.*, vol. 2, 2021, [Online]. Available: <https://journal.jis-institute.org/index.php/jnpm/article/view/486>
- [7] T. Naprawski, “The Impact of Web Analytics Tools on Knowledge Management,” *ScienceDirect*, vol. 225, p. 3407, 2023, [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S187705092301493X>
- [8] A. Rovetta, “An integrated infoveillance approach using google trends and Talkwalker: Listening to web concerns about COVID-19 vaccines in Italy,” *ScienceDirect*, vol. 4, no. 100272, p. 2, 2023, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2772442523001399?ref=pdf_download&fr=RR-2&rr=866fdcfb4bdf40e0
- [9] S. García and J. M. B. and F. H. Sergio Ramírez-Gallego, Julián Luengo, “Big data preprocessing: methods and prospects,” *Springer*, p. 3, 2016, doi: 10.1186/s41044-016-0014-0.
- [10] R. Maulana, A. Voutama, and T. Ridwan, “Analisis Sentimen Ulasan Aplikasi MyPertamina pada Google Play Store menggunakan Algoritma NBC,” *J. Teknol. Terpadu*, vol. 9, p. 44, 2023, [Online]. Available: <https://journal.nurulfikri.ac.id/index.php/JTT/article/view/609/306>
- [11] K. E. ArunKumar, D. V. Kalaga, C. M. S. Kumar, M. Kawaji, and T. M. Brenza, “Forecasting of COVID-19 using deep layer Recurrent Neural Networks (RNNs) with Gated Recurrent Units (GRUs) and Long Short-Term Memory (LSTM) cells,” *ScienceDirect*, vol. 146, p. 3, 2021, [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0960077921002149>
- [12] S. V. Belavadia, S. Rajagopala, R. Ra, and R. Mohan, “Air Quality Forecasting using LSTM RNN and Wireless Sensor Networks,” *ScienceDirect*, vol. 170, 2020.
- [13] K. E. ArunKumar, D. V. Kalaga, C. M. S. Kumar, Masahiro, Kawaji, and T. M. Brenza, “Comparative analysis of Gated Recurrent Units (GRU), long Short-Term memory (LSTM) cells, autoregressive Integrated moving average (ARIMA), seasonal autoregressive Integrated moving average (SARIMA) for forecasting COVID-19 trends,” *ScienceDirect*, vol. 61, 2022.
- [14] A. T. A *et al.*, “Reconstruction of missing streamflow series in human-regulated catchments using a data integration LSTM model,” *ScienceDirect*, vol. 52, 2024.
- [15] D. Fitriati, “PERBANDINGAN KINERJA CNN LeNet5 DANEXTREME LEARNING MACHINE PADA PENGENALAN CITRA TULISAN TANGAN ANGKA,” *J. Teknol. Terpadu*, vol. 2, 2016, [Online]. Available: <https://journal.nurulfikri.ac.id/index.php/JTT/article/view/45/33>