

Artikel : [Akses terbuka/Open Access](#)

PEMANFAATAN GOOGLE TRENDS DALAM NOWCASTING JUMLAH KEDATANGAN PENUMPANG DI BANDARA KOMODO LABUAN BAJO

Sitasi : Bily. 2024, JSTAR 4(2), 69-82.

Kronologi naskah.

Submit : 4 November 2024

Revisi : 27 Desember 2024

Diterima : 27 Desember 2024



Penyedia Data Statistik Berkualitas untuk
Indonesia Maju

REFORMASI BIROKRASI



PEMANFAATAN GOOGLE TRENDS DALAM NOWCASTING JUMLAH KEDATANGAN PENUMPANG DI BANDARA KOMODO LABUAN BAJO

Silas Ino Bily¹

¹BPS Kabupaten Sumba Barat Daya, Indonesia

[‡]korespondensi *author*: silas.ino@bps.go.id

Abstract

This study investigates the application of Google Trends data as regressor variables for nowcasting passenger arrivals at Komodo Airport, the primary gateway to Labuan Bajo, a key tourism destination in Indonesia. Utilizing machine learning, the research aims to enhance the accuracy of passenger arrival prediction while deriving insights related to tourism interest in Labuan Bajo. The analysis reveals a significant correlation between trending search terms and actual passenger arrivals, indicating that shifts in online interest can effectively predict tourist behavior. Additionally, specific search queries reflecting potential visitors' preferences are identified, providing valuable insights for tourism stakeholders and professionals. The findings underscore the relevance of passenger arrival data as a critical indicator of tourism activity, assisting policymakers and businesses in making informed decisions to enhance the tourist experience in Labuan Bajo. By utilizing Google Trends data in the nowcasting process, this research contributes to the evolving discourse on data-driven strategies in tourism management and highlights the potential of leveraging Big Data to support sustainable tourism development in Labuan Bajo, aligning with government efforts to promote the region as a premier travel destination in Indonesia.

Keyword: *tourism, nowcasting, Big Data, Google Trends, Labuan Bajo*

1. Pendahuluan

Labuan Bajo sebagai salah satu destinasi pariwisata super prioritas yang telah ditetapkan Pemerintah Indonesia (Kemenparekraf, 2020) memiliki peran penting dalam pertumbuhan ekonomi kabupaten Manggarai Barat bahkan untuk provinsi Nusa Tenggara Timur (NTT). Berdasarkan data Badan Pusat Statistik (BPS), produk domestik regional bruto (PDRB) pada lapangan usaha penyediaan akomodasi dan makan minum menunjukkan pertumbuhan tertinggi dari tahun ke tahun di kabupaten Manggarai Barat. Pada tahun

2023, lapangan usaha ini bertumbuh sebesar 33,54% (BPS, 2024). Hal ini menunjukkan bahwa pariwisata merupakan sektor yang paling potensial di wilayah tersebut dan harus dikembangkan untuk pembangunan yang berkelanjutan, perluasan lapangan kerja, hingga peningkatan pertumbuhan ekonomi.

Bandara Komodo, yang baru saja ditetapkan sebagai bandara internasional melalui Keputusan Menteri Perhubungan Nomor 31/2024 menjadi gerbang utama bagi wisatawan yang berkunjung ke Labuan Bajo, khususnya

dengan dibukanya penerbangan internasional yang diprediksi akan semakin meningkatkan jumlah kedatangan penumpang di masa mendatang. Selain itu, menurut data BPS (2024) jumlah kedatangan penumpang di bandara Komodo pada bulan Agustus 2024 meningkat pesat melebihi jumlah penumpang yang tiba di bandara Eltari di Kupang yang sebelumnya merupakan bandara dengan jumlah penumpang terbanyak di provinsi NTT. Hal ini menempatkan posisi penting bandara Komodo dalam perkembangan pariwisata di NTT, khususnya di Labuan Bajo, kabupaten Manggarai Barat.

BPS tidak merilis data resmi mengenai kunjungan wisatawan setiap bulan di Labuan Bajo. Sebagai alternatif, BPS merilis data resmi terkait jumlah penumpang di bandara Komodo. Data ini sangat penting dan dapat digunakan sebagai *proxy* untuk analisis perkembangan pariwisata di Labuan Bajo. Namun, data penumpang pesawat secara resmi dirilis oleh BPS dengan keterlambatan atau *lag* waktu 2 (dua) bulan. Hal ini menjadi tantangan karena sangatlah penting bagi para pemangku kepentingan untuk mendapatkan data-data terkini guna pengambilan keputusan yang lebih baik di sektor pariwisata. Untuk menjawab tantangan ini, penggunaan metode *nowcasting* menjadi relevan.

Nowcasting merupakan metode untuk memprediksi data terkini atau data masa sekarang, atau sangat dekat dengan masa sekarang dengan menggunakan data historis atau sumber

data lain. Sumber data lain yang dapat dipertimbangkan dan sudah digunakan secara luas adalah *Big Data*. Salah satu sumber *Big Data* yang dapat dimanfaatkan adalah *Google Trends*, yang menyediakan indeks pencarian pada mesin pencarian Google yang relevan dengan minat perjalanan secara *real-time*.

Google adalah alamat web dan sekaligus mesin pencarian paling populer di dunia. Hingga Juli 2024, Google masih menguasai pangsa pasar terbesar dengan proporsi sebesar 91,02% (Statcounter, 2024). Kondisi yang sama juga berlaku di Indonesia. Distribusi pasar Google pada bulan Juli 2024 adalah 94,3% (Statcounter, 2024). Oleh karena itu, data pencarian Google menjadi relevan untuk mewakili data yang menggambarkan keseluruhan populasi.

Google Trends adalah aplikasi yang menyediakan indeks pencarian untuk berbagai topik, termasuk topik pencarian terkait perjalanan yang dilakukan oleh pelaku perjalanan. Data ini tersedia secara *real-time* berupa data bulanan, mingguan, dan harian untuk tahun 2004 hingga sekarang. Pola pencarian ini seringkali berkorelasi atau mencerminkan perilaku nyata di dunia, seperti kedatangan pengunjung ke suatu wilayah. Dengan karakteristik yang demikian, data ini menjadi data alternatif untuk menjembatani *lag* waktu ketersediaan berbagai indikator maupun statistik resmi (Hammer dkk, 2017). Selain itu penggunaan data *Google Trends* berpotensi memberikan akurasi

yang lebih baik ketika terjadi perubahan pola atau perilaku masyarakat seperti pada saat terjadi perubahan pada tren sosial, maupun keadaan lain yang dapat mempengaruhi kedatangan penumpang, dan menggunakan data historis saja tidak akurat lagi. *Google Trends* memiliki potensi yang dapat dimanfaatkan untuk berbagai analisis ekonomi dan sosial, seperti estimasi PDB, sentimen pelanggan, dan penjualan mobil dan motor (Hammer dkk, 2017). Selain itu, melalui penelitian Akbar dan Kurniawan (2020) ditemukan bahwa *Google Trends* juga bisa digunakan untuk melakukan *nowcasting* tingkat pengangguran terbuka di Provinsi Jawa Barat dengan akurasi yang baik.

Berdasarkan uraian-uraian di atas, penelitian ini bertujuan untuk mengidentifikasi kueri pencarian di *Google Trends* yang berkorelasi tinggi dengan kedatangan penumpang di bandara Komodo. Selain itu, variabel indeks pencarian tersebut akan diseleksi dan dimasukkan sebagai variabel prediktor ke dalam beberapa model *nowcasting* dalam upaya memperkirakan kedatangan penumpang di bandara Komodo, Labuan Bajo. Pada akhirnya, model dengan kesalahan paling kecil akan terpilih menjadi model terbaik.

2. Metodologi

Dalam penelitian ini dilakukan pemodelan untuk *nowcasting* jumlah kedatangan penumpang di bandara Komodo, Labuan Bajo dengan memanfaatkan data indeks pencarian

dari *Google Trends* sebagai variabel prediktor. Proses ini melalui beberapa tahap yaitu tahap pengumpulan data, tahap pra-pemrosesan data, dan tahap analisis.

Data yang digunakan yaitu data jumlah penumpang pesawat yang tiba di bandara Komodo, Labuan Bajo menurut bulan selama 2019 hingga 2024. Data tersebut menjadi variabel target peramalan (variabel Y). Selain itu, dilakukan juga pengumpulan data untuk variabel prediktor atau variabel X, yaitu data indeks pencarian *Google* yang dilakukan di dalam wilayah Indonesia yang berhubungan kuat dengan variabel Y.

Pengumpulan Data Indeks Pencarian

Tahap-tahap pengumpulan data indeks pencarian *Google* dijelaskan sebagai berikut:

1. Peneliti menentukan daftar kueri pencarian dasar (*seed queries*) dengan mengadopsi metode yang dilakukan oleh Li dkk, (2017). Kueri ditentukan berdasarkan 6 aspek perjalanan.
2. Peneliti mengunduh daftar kueri pencarian terkait (*related queries*). Kueri pencarian terkait terdapat pada fitur aplikasi *Google Trends* yang menunjukkan semua kueri pencarian yang terkait dengan topik yang kita berikan (kueri dasar). Ruang lingkup geografi data dibatasi untuk pencarian dalam wilayah Indonesia saja, karena penerbangan dari luar negeri baru dibuka pada tahun 2024 dan tidak relevan dengan periode data dalam penelitian ini.
3. Semua kueri pencarian terkait dikumpulkan dan dilakukan *filtering* untuk menghilangkan duplikat. Kemudian, kueri diseleksi lagi secara

manual untuk mengeliminasi kueri pencarian yang tidak memiliki makna yang relevan dengan Labuan Bajo.

4. Peneliti melakukan pengumpulan data indeks pencarian untuk setiap kueri pencarian terkait yang telah diunduh sebelumnya. Untuk mempermudah proses ini, mengingat jumlah topik kueri yang banyak peneliti menggunakan *web API* yang disediakan oleh Google dan *package* “gtrendsR” yang disediakan oleh Massicotte & Eddebuettel (2022). Proses ini dilakukan dengan aplikasi R Studio, sehingga mendukung otomatisasi proses tersebut. Data yang dikumpulkan adalah data dengan periode bulanan dari tahun 2019 hingga 2024 untuk menyesuaikan dengan periode variabel Y.
5. Peneliti melakukan seleksi variabel-variabel (*feature selection*) data Google yaitu, menyeleksi variabel Google yang berkorelasi tinggi dengan variabel Y. Nilai korelasi minimum yang ditetapkan yaitu minimal bernilai 0,70.
6. Setelah mendapatkan variabel-variabel indeks pencarian yang berkorelasi tinggi, peneliti kemudian melakukan seleksi variabel (*feature selection*) tahap kedua menggunakan metode *Least Absolute Shrinkage and Selection Operator* (LASSO) untuk mencegah *overfitting* dikarenakan terlalu banyaknya variabel prediktor.

Metode Evaluasi

Data-data tersebut yang dihasilkan melalui tahap sebelumnya kemudian dilakukan standardisasi dan dibagi ke dalam 2 set yaitu *training set* dan *test set*.

Training set adalah data dari Januari

2019 hingga Desember 2023 dan akan digunakan dalam proses pelatihan model. Sedangkan *test set* adalah data dari Januari 2024 hingga Agustus 2024 dan akan digunakan untuk mengevaluasi akurasi hasil prediksi model. Metode ini disebut sebagai evaluasi *out-of-sample* di mana evaluasi menggunakan data yang belum pernah digunakan dalam pelatihan model.

Metrik yang akan digunakan dalam mengevaluasi model peramalan adalah *Mean Absolute Percentage Error* (MAPE). MAPE merupakan ukuran statistik yang menunjukkan seberapa besar kesalahan peramalan dibandingkan dengan nilai sebenarnya. Semakin kecil nilai MAPE, maka semakin akurat hasil peramalan. Nilai MAPE yang tidak melebihi 50% masih dapat digunakan, sedangkan nilai MAPE yang lebih dari 50% menunjukkan bahwa model peramalan tidak dapat digunakan lagi (Lewis, 1982).

Bahan dan Sumber Data

Dua variabel yang digunakan dalam penelitian berasal dari sumber yang berbeda. Pertama, data jumlah penumpang pesawat yang tiba di bandara Komodo Labuan Bajo menurut bulan diambil dari Berita Resmi Statistik BPS Provinsi NTT pada *website* resminya. Kedua, data indeks pencarian Google didapatkan dari aplikasi Google *Trends* menggunakan *web API*.

Data tentang jumlah penumpang pesawat dirilis setiap bulan dalam satuan orang. Semua penumpang yang tiba dihitung tanpa memisahkan menurut maksud kedatangan. Data ini termasuk semua kedatangan baik domestik maupun internasional.

Google *Trends* adalah aplikasi yang menyediakan data popularitas suatu kueri pencarian relatif terhadap suatu

periode waktu dan wilayah. Pengguna dapat mengatur spesifikasi data sesuai kebutuhan. Data *Google Trends* terdiri dari dua jenis yaitu (dikutip dari Google, 2024):

1. Data *realtime* adalah sampel acak penelusuran dalam 7 hari terakhir.
2. Data *non-realtime* adalah sampel acak dari penelusuran Google dari tahun 2004 hingga 36 jam sebelum pengambilan data di *Google Trends*.

Data yang diberikan adalah data indeks bukan jumlah total pencarian yang sebenarnya. Indeks ini berada pada rentang 0 hingga 100. Indeks ini dihitung dengan cara membandingkan jumlah pencarian yang mengandung kueri tertentu terhadap total semua pencarian pada periode dan wilayah yang sama.

Google Trends tidak menyediakan data untuk kasus pencarian yang dilakukan oleh sedikit orang sehingga pencarian dengan volume kecil akan diberikan nilai 0 (nol). Selain itu, penelusuran duplikat yang dilakukan oleh orang yang sama dalam jangka waktu singkat, dan kueri berisi karakter khusus lainnya akan dieliminasi.

Google Trends secara *default* menyediakan data mingguan jika rentang waktu yang diminta lebih dari tiga bulan dan kurang atau sama dengan lima tahun. Jika data yang diminta kurang atau sama dengan tiga bulan maka akan diberikan data harian. Sedangkan jika data yang diminta lebih dari lima tahun maka akan diberikan data bulanan. Oleh karena itu, pada kasus penelitian ini, data *Google Trends* yang didapatkan adalah data dalam periode bulanan.

Setelah melalui beberapa proses pengumpulan data seperti yang telah dijelaskan pada bagian metodologi data *Google Trends* siap dimasukkan ke dalam analisis *nowcasting*.

Metode Analisis Data

Dalam penelitian ini dilakukan *nowcasting* jumlah penumpang pesawat yang tiba di bandara Komodo Labuan Bajo setiap bulan dengan menggunakan variabel prediktor yaitu variabel-variabel indeks pencarian Google. Selain itu, dalam model juga ditambahkan variabel pembantu yaitu indeks waktu (*time index*). Menggunakan indeks waktu dalam model regresi membantu menangkap pola musiman, seperti dalam analisis data bulanan kedatangan penumpang. Draper & Smith (1998) menjelaskan bahwa variabel kategorikal untuk bulan memungkinkan model mengakomodasi efek musiman. Penelitian Cancurt & Subasi (2015) menunjukkan bahwa kombinasi indeks waktu dengan *Google Trends* dapat meningkatkan akurasi prediksi, karena indeks waktu menangkap pola musiman yang tidak dijelaskan sepenuhnya oleh data *Google Trends* sehingga membantu model menangkap fitur musiman pada data.

Model *nowcasting* tersebut dapat dirumuskan melalui persamaan berikut ini.

$$Y_n = f(X1_n, X2_n, \dots, Xi_n, MIndeks_n)$$

Keterangan:

Y_n = Jumlah penumpang pesawat yang tiba di bandara Komodo pada bulan sekarang atau bulan n .

Xi_n = Indeks pencarian Google untuk kueri pencarian ke- i pada bulan sekarang atau bulan n .

$MIndeks_n$ = Indeks waktu per bulan sekarang atau bulan n , bernilai 1 hingga 12, dan bertipe kategorial.

Metode Peramalan

Adapun metode peramalan yang digunakan adalah metode *Machine Learning* dengan algoritma *Multi Layer Perceptron* (MLP). Beberapa kombinasi parameter akan dilakukan untuk membentuk beberapa model dan kemudian akan dipilih model terbaik.

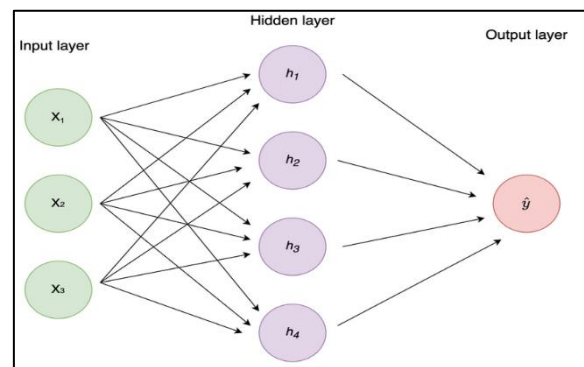
Pertama, data terlebih dahulu dibagi menjadi dua bagian yaitu data untuk *training* (*training set*) dan data untuk tes (*test set*). Model akan dilatih dengan data pada *training set*, lalu akan diuji akurasi dengan *test set*. Model dengan kesalahan minimal akan dipilih menjadi model terbaik. Metode evaluasi akurasi yang digunakan adalah *Mean Absolute Percentage Error* (MAPE).

Multi-Layer Perceptron (MLP)

Multilayer Perceptrons (MLPs) adalah salah satu jenis *Artificial Neural Networks* (ANN) yang terdiri dari beberapa lapisan neuron. Setiap neuron dalam jaringan ini terhubung dengan setiap neuron di lapisan berikutnya, membentuk jaringan yang sepenuhnya terhubung. Struktur MLP biasanya mencakup lapisan input (*input layer*), satu atau lebih lapisan tersembunyi (*hidden layer*), dan lapisan output (*output layer*). Bagan representasi dari formula untuk MLP dengan satu lapisan tersembunyi ditampilkan pada gambar 1.

Neuron-neuron di *hidden layers* menggunakan fungsi aktivasi nonlinier seperti sigmoid, tanh (*hyperbolic tangent*), atau ReLU (*Rectified Linear Unit*). Nonlinieritas ini memungkinkan MLP untuk belajar pola dan hubungan

kompleks dalam data. Setiap neuron juga memiliki bias yang membantu menyesuaikan output neuron dengan lebih fleksibel. Proses pembelajaran melibatkan penyesuaian bobot dan bias neuron melalui metode yang disebut *backpropagation*, yang bertujuan untuk meminimalkan kesalahan dalam output jaringan.



Gambar 1. Skema Algoritma *Multilayer Perceptron*

Dalam sebuah *Multilayer Perceptron* (MLP), data input pertama-tama diproses oleh *hidden layers*, yang mengubah data menjadi representasi antara yang menggambarkan pola dan fitur penting yang diekstraksi dari input. Data ini kemudian diteruskan ke *output layer*. *Output layer* kemudian memproses lebih lanjut data yang telah diubah ini untuk menghasilkan hasil akhir, seperti prediksi atau klasifikasi.

Sifat nonlinier dari MLP memungkinkannya untuk memodelkan hubungan kompleks dalam data, yang mengarah pada peningkatan akurasi dalam model prediktif. Fleksibilitas algoritma ini memungkinkannya untuk melakukan berbagai tugas seperti klasifikasi, regresi, dan lain-lain. Oleh karena itu, sangat cocok digunakan

untuk menghasilkan prediksi yang akurat.

Dalam pemodelan dengan *machine learning* tidak dapat dipastikan sedari awal parameter-parameter yang sedemikian rupa memberikan akurasi terbaik pada model, melainkan perlu dilakukan pencarian parameter yang memberikan akurasi terbaik melalui proses mencoba semua kombinasi konfigurasi yang mungkin. Proses percobaan untuk mendapatkan parameter terbaik ini disebut dengan *tuning*. Pada *tuning* praktisi perlu menentukan konfigurasi-konfigurasi yang akan digunakan untuk mendapatkan parameter terbaik. Konfigurasi-konfigurasi ini disebut dengan *hyperparameter*.

Ada beberapa cara yang dapat digunakan untuk melakukan proses tersebut, salah satu di antaranya yaitu metode *grid search*. *Grid search* adalah salah satu metode paling sederhana untuk melakukan *tuning* model terbaik. *Grid search* menggunakan kombinasi nilai-nilai *hyperparameter* yang ada pada *grid* nilai yang telah ditentukan di awal. Pada akhirnya, berdasarkan berbagai kombinasi tersebut dapat ditemukan model yang sedemikian rupa memberikan akurasi terbaik. Walaupun dalam beberapa penelitian seperti Liashchynskiy & Liashchynskiy, (2019), dan Chowdhury dkk, (2022) telah ditemukan bahwa metode ini memakan waktu yang lebih lama, peneliti akan menggunakan metode ini karena sederhana dan lebih memastikan untuk mendapatkan parameter terbaik karena

mencoba semua kemungkinan kombinasi (Liashchynskiy & Liashchynskiy, 2019). Selain itu, jumlah *hyperparameter* dan dimensi *dataset* penelitian ini tidak begitu besar sehingga tidak ada kendala proses yang lama.

Pada kasus penelitian ini, yaitu menggunakan MLP, proses pemodelan akan dilakukan menggunakan aplikasi R Studio dengan *package* “nnet” (Ripley & Venables, 2023). Sedangkan, untuk melakukan *tuning* dengan metode *grid search* akan menggunakan *package* “caret” (Kuhn, 2023).

Pada proses ini dilakukan *tuning* untuk beberapa *hyperparameter* penting dengan metode *grid search* yaitu, *size*, *decay*, dan jumlah iterasi. Sedangkan *hyperparameter* lainnya akan menggunakan nilai *default*. Penentuan *hyperparameter* yang akan digunakan adalah sebagai berikut:

1. *Size* yaitu jumlah neuron pada *hidden layer* akan digunakan sebanyak 1, 2, 3, 5, 7.
2. *decay* adalah parameter regulasi yang mencegah *overfitting*, nilai yang akan digunakan yaitu 0,1; 0,01; dan 0,001.
3. Untuk jumlah iterasi maksimal peneliti menggunakan nilai maksimal 100.

Peneliti menggunakan nilai *hyperparameter* pada batasan nilai yang kecil karena lebih sesuai dengan *dataset* yang kecil dan tidak berdimensi besar, dan kurang rentan terhadap *overfitting* (Brigato & Locchi, 2020).

3. Hasil dan Pembahasan

Kueri Pencarian Dasar

Peneliti menyusun daftar kueri-kueri pencarian dasar berdasarkan 6 aspek perjalanan menurut Li dkk, (2017). Peneliti juga menambahkan aspek baru yaitu tujuan wisata di sekitarnya. Kueri-kueri pencarian dasar yang dibentuk dapat dilihat pada tabel 1 di bawah ini. Kueri-kueri ini akan digunakan sebagai input untuk mendapatkan daftar kueri pencarian terkait di Aplikasi Google Trends.

Tabel 1. Daftar kueri pencarian dasar menurut 6 (enam) aspek perjalanan

| Aspek | Kueri Pencarian Dasar |
|-----------------------------------|--|
| <i>Dining</i> (makanan) | Labuan Bajo restoran, Labuan Bajo <i>Food</i> , Labuan Bajo makanan, Labuan Bajo <i>cafe</i> |
| <i>Lodging</i> (penginapan) | Hotel Labuan Bajo, <i>Villa</i> labuan Bajo, Labuan Bajo <i>resort</i> |
| <i>Shopping</i> (perbelanjaan) | <i>Mall</i> Labuan Bajo, belanja Labuan Bajo |
| <i>Traffic</i> (lalu lintas) | Bandara Komodo, <i>Komodo Airport</i> , penerbangan Labuan Bajo, penerbangan komodo |
| <i>Tour</i> (tur) | Trip Labuan Bajo, trip pulau Padar, trip komodo, <i>travel</i> Labuan Bajo, pantai Labuan Bajo, Labuan Bajo <i>beach</i> , <i>Pink</i> <i>beach</i> |
| <i>Recreation</i> (rekreasi) | <i>Things to do in</i> Labuan Bajo, wisata Labuan Bajo, atraksi Labuan Bajo, atraksi komodo |
| Tujuan wisata di sekitarnya | Manggarai, Flores, Wae Rebo |

Kueri Pencarian Terkait

Setelah dilakukan input kueri-kueri dasar (*seed queries*) di Aplikasi Google Trends didapatkan daftar kueri pencarian terkait yang unik dan memiliki makna yang relevan dengan Labuan Bajo. Jumlah kueri pencarian terkait yang didapatkan sebanyak 199 kueri pencarian yang unik dan memiliki makna yang relevan dengan Labuan Bajo.

Variabel Indeks Pencarian yang Terseleksi

Setelah melalui proses *feature selection* didapatkan hasil-hasil yang dijelaskan sebagai berikut.

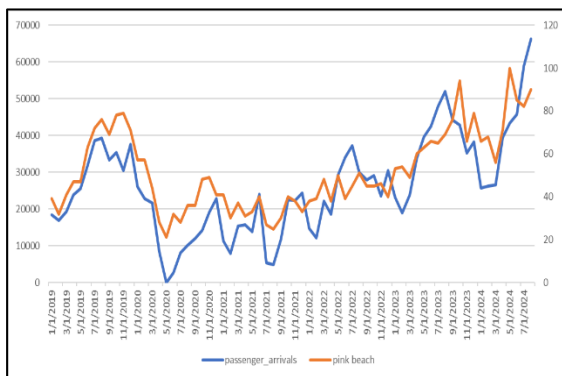
Dapat dilihat pada tabel 2, semua kueri pencarian yang berkorelasi tinggi dengan data jumlah penumpang pesawat yang tiba di bandara Komodo Labuan Bajo. Jumlah kueri pencarian yang berkorelasi adalah sebanyak 20 kueri pencarian.

Dari tabel ini juga bisa diperoleh informasi mengenai topik-topik pencarian yang populer yang digunakan oleh pelaku perjalanan yang datang ke Labuan Bajo. Informasi ini dapat digunakan untuk analisis lebih lanjut tentang tren pariwisata di Labuan Bajo.

Tabel 2. Daftar kueri pencarian yang diseleksi dengan nilai korelasi

| No | Kueri pencarian | Nilai Korelasi |
|----|---------------------|----------------|
| 1 | pink beach | 0.856 |
| 2 | lombok to komodo | 0.818 |
| 3 | bali to labuan bajo | 0.817 |
| 4 | lombok pink beach | 0.808 |

| No | Kueri pencarian | Nilai Korelasi |
|----|-----------------------|----------------|
| 5 | labuan bajo beach | 0.796 |
| 6 | labuan bajo lombok | 0.774 |
| 7 | sudamala resort | 0.771 |
| 8 | komodo island tour | 0.785 |
| 9 | bali to komodo | 0.759 |
| 10 | komodo island | 0.753 |
| 11 | labuan bajo lombok | 0.751 |
| 12 | labuan bajo airport | 0.751 |
| 13 | surabaya labuan bajo | 0.744 |
| 14 | bali labuan bajo | 0.743 |
| 15 | labuan bajo komodo | 0.719 |
| 16 | manggarai | 0.710 |
| 17 | komodo tour | 0.708 |
| 18 | komodo airport | 0.707 |
| 19 | labuan bajo | 0.702 |
| 20 | lombok to labuan bajo | 0.701 |



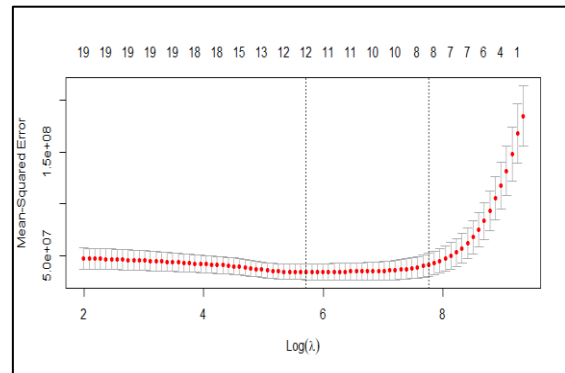
Gambar 2. Perbandingan data jumlah penumpang pesawat di Bandara Komodo dengan Indeks pencarian “pink beach”

Dapat dilihat pada gambar 2 di atas, terdapat pergerakan yang sama antara kedua variabel pada waktu yang sama, sehingga hal ini mendukung penggunaan variabel indeks pencarian sebagai prediktor.

Dikarenakan jumlah kueri yang berkorelasi cukup banyak maka peneliti melakukan seleksi variabel yang lebih signifikan dengan metode LASSO.

Pada metode LASSO sangat penting untuk menentukan nilai λ yang akan mengontrol jumlah variabel yang

direduksi. Oleh karena itu, peneliti menentukan nilai λ optimal dengan melakukan *k-fold cross validation* dengan memanfaatkan fungsi *cv.glmnet()* pada Aplikasi R.



Gambar 3. Nilai MSE untuk berbagai nilai λ

Pada gambar 3 di atas ini dapat dilihat nilai λ dengan kesalahan minimal adalah dengan 12 variabel, sedangkan nilai λ dengan kesalahan pada 1 standar eror adalah dengan 8 variabel. Namun, untuk mencegah *overfitting* peneliti memilih 8 variabel saja yaitu variabel dengan nilai koefisien paling signifikan.

Daftar kueri pencarian yang signifikan menurut LASSO dapat dilihat pada tabel 3 berikut ini. Variabel-variabel dengan koefisien tertinggi adalah “surabaya labuan bajo”, diikuti oleh “pink beach” dan “lombok pink beach”.

Tabel 3. Daftar 8 kueri pencarian hasil proses LASSO dan nilai koefisiennya

| No | Kueri Pencarian | Nilai Koefisien |
|----|-------------------|-----------------|
| 1 | pink beach | 120,802 |
| 2 | lombok to komodo | 64,415 |
| 3 | lombok pink beach | 89,783 |
| 4 | labuan bajo beach | 68,973 |
| 5 | sudamala resort | 78,478 |

| No | Kueri Pencarian | Nilai Koefisien |
|----|----------------------|-----------------|
| 6 | komodo island tour | 40,220 |
| 7 | surabaya labuan bajo | 171,419 |
| 8 | labuan bajo | 51,028 |

Model Nowcasting

Model *nowcasting* dibentuk dengan melakukan percobaan menggunakan berbagai kombinasi parameter melalui proses *tuning* parameter. Beberapa parameter yang dikondisikan yaitu jumlah *hidden layer* (*size*), parameter regulasi (*decay*), dan jenis fungsi linier atau non-linier. Berikut ringkasan beberapa model terbaik yang didapatkan melalui *training* dengan *training set*.

Tabel 4. Model terbaik berdasarkan beberapa kombinasi parameter

| Model | Kombinasi parameter terbaik |
|------------|-----------------------------|
| linier | Size = 7, decay = 0.1 |
| Non-linier | Size = 5, decay = 0.01 |

Evaluasi Hasil Prediksi

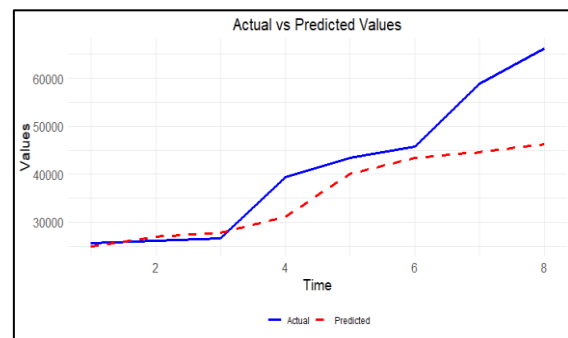
Untuk mengevaluasi keakuratan model dalam memprediksi, peneliti melakukan evaluasi hasil peramalan menggunakan *test set* yaitu data *out-of-sample* yang tidak digunakan dalam pelatihan model. Hal ini bertujuan untuk melihat kemampuan model memprediksi ketika menggunakan data yang benar-benar baru.

Kedua model yang didapatkan pada proses sebelumnya digunakan untuk melakukan prediksi dan dievaluasi berdasarkan nilai MAPE. Hasil peramalan *out-of-sample* kedua model

dapat dilihat pada tabel dan grafik berikut.

Tabel 5. Nilai MAPE hasil peramalan *out-of-sample* menggunakan *test set*

| Model | MAPE <i>out-of-sample</i> (%) |
|------------|-------------------------------|
| Linier | 12,05 |
| Non linier | 18,09 |



Gambar 4. Hasil peramalan menggunakan model dengan MAPE terbaik dibandingkan dengan data aktual, Januari-Agustus 2024

Pada tabel 5 di atas dapat dilihat bahwa model terbaik untuk kebutuhan peramalan ini adalah model linier. Nilai MAPE dari model tersebut adalah sebesar 12,05% dan nilai ini dikategorikan sebagai nilai yang akurat (Lewis, 1982).

Pada gambar 4 dapat dilihat hasil peramalan (garis merah putus-putus) dibandingkan dengan jumlah aktualnya (garis biru). Terlihat bahwa model dapat memprediksi tren atau arah pergerakan data. Namun, pada bulan Agustus 2024 model masih kurang akurat memprediksi data. Jumlah penumpang pada bulan Agustus 2024 adalah yang paling banyak. Hal ini menjadi hal yang perlu dipertimbangkan untuk pengembangan metode yang lebih akurat lagi, seperti

menambahkan prediksi untuk penumpang internasional.

4. Simpulan dan Saran

Berdasarkan hasil-hasil dan pembahasan di atas, dapat ditarik beberapa kesimpulan dan saran untuk penelitian lanjutan, pembuat kebijakan dan praktisi di bidang pariwisata.

Beberapa kesimpulan yang dapat ditarik yaitu:

1. Untuk mengatasi keterlambatan rilis data resmi dapat dilakukan *nowcasting* untuk memprediksi data terkini dengan memanfaatkan sumber data lain yang tersedia dengan waktu yang lebih cepat.
2. Sumber data *Google Trends* yang menyediakan data indeks pencarian Google dapat dimanfaatkan untuk melakukan *nowcasting* data jumlah penumpang pesawat bulanan yang tiba di bandara Komodo Labuan Bajo.
3. Melalui proses pengumpulan data dan *feature selection* didapatkan 8 kueri pencarian yang berkorelasi tinggi dan menjadi prediktor yang signifikan untuk peramalan.
4. Model peramalan dengan algoritma MLP dengan menggunakan data *-out-of-sample* memberikan nilai MAPE sebesar 12,05%, yang menandakan model tersebut akurat untuk digunakan.

Beberapa saran yang dapat diberikan untuk penelitian selanjutnya yaitu:

1. Mengingat pola perilaku sosial yang sangat dinamis, sangat perlu untuk melakukan pembaruan model yang ada dan menyesuaikan dengan keadaan terkini, seperti penyesuaian variabel dan penambahan jumlah data historis untuk pelatihan model. Semakin banyak data yang digunakan untuk pelatihan, maka model akan semakin mengenali pola dalam data dan meningkatkan akurasi.
2. Mengingat data jumlah penumpang pesawat juga sudah termasuk penumpang internasional yang rutanya baru saja dibuka, maka perlu dibentuk model untuk meramalkan jumlah penumpang internasional, kemudian dijumlahkan dengan angka prediksi penumpang domestik.
3. Perlu dilakukan pemodelan dengan membandingkan berbagai metode yang lain yang dapat menambah akurasi peramalan.
4. Perlu dipertimbangkan untuk menggunakan kombinasi data lain dari sumber yang berbeda, seperti sumber *Big Data* lainnya yang juga relevan terutama untuk prediktor yang berkorelasi negatif yang menentukan berkurangnya jumlah penumpang pesawat yang tiba. Sumber data yang dapat dimanfaatkan yaitu data berita bencana alam atau pandemi. Selain itu, berikut beberapa saran

yang dapat diberikan kepada pembuat kebijakan maupun praktisi di bidang pariwisata:

1. Istilah-istilah pencarian populer yang berkorelasi dengan kedatangan penumpang di Labuan Bajo, seperti “*pink beach*”, “*sudamala resort*”, “*komodo island*”, dan lain-lain telah menggambarkan lokasi yang paling diminati oleh pencari informasi wisata, sehingga pemerintah atau praktisi pariwisata dapat memanfaatkan informasi ini untuk menyusun strategi untuk mempromosikan atraksi-atraksi tersebut dan membuat kebijakan yang meningkatkan fasilitas dan pengalaman wisatawan untuk menjadi lebih baik. Di sisi lain, dapat disusun juga strategi untuk mempromosikan atraksi pariwisata lainnya yang masih kurang populer.
2. Terdapat juga istilah-istilah pencarian terkait rute perjalanan paling populer yang berkorelasi tinggi dengan kedatangan penumpang seperti, “*bali to labuan bajo*”, “*lombok to labuan bajo*”, “*surabaya to labuan bajo*”. Informasi ini dapat dimanfaatkan untuk menyusun strategi pengaturan layanan di bandara dengan mengoptimalkan layanan pada penerbangan dari daerah asal penumpang yang populer. Selain itu, dapat disusun strategi layanan ketika permintaan meningkat.

3. Mengingat data *Google Trends* juga tersedia dalam periode harian atau mingguan, dan data disajikan dengan indeks bukan jumlah kumulatif, maka praktisi dapat memanfaatkan data awal pada minggu pertama setiap bulan untuk memprediksi jumlah kedatangan penumpang pada bulan tersebut dan diperbarui setiap minggunya tanpa harus menunggu pada akhir bulan. Dengan ini, didapatkan informasi awal untuk menyusun strategi lebih awal dan lebih dinamis.
4. Secara umum, mengingat perkembangan teknologi informasi yang pesat, sangat mungkin untuk mengembangkan strategi kebijakan dan pelayanan yang lebih dinamis dengan memanfaatkan analisis *Big Data* dan *Machine Learning* yang berpotensi memberikan hasil yang lebih cepat, akurat, dan efisien daripada hanya menggunakan statistik resmi dan metode analisis klasik.

Daftar Pustaka

- Akbar, I. A., & Kurniawan, R. (2020). Pemodelan Nowcasting Tingkat Pengangguran Terbuka Menggunakan Data *Google Trends* dengan Metode Antlion Optimization-Support Vector Regression: Studi Kasus di Provinsi Jawa Barat pada 2005-2019. *Prosiding Seminar Nasional Official Statistics, 2020(1)*, 105–114.

- <https://doi.org/10.34123/semnasoffstat.v2020i1.504>
- Badan Pusat Statistik. (n.d.). Laju Pertumbuhan PDRB Menurut Lapangan Usaha (Persen). Diakses dari <https://manggaraibaratkab.bps.go.id/id/statistics-table/2/MzgjMg==/laju-pertumbuhan-pdrb-menurut-lapangan-usaha--persen-.html>
- Badan Pusat Statistik. (2024, 1 Oktober). Jumlah penerbangan angkutan udara pada Agustus 2024 adalah 3.557 penerbangan. Diakses pada 25 Oktober 2024 dari <https://ntt.bps.go.id/id/pressrelease/2024/10/01/1378/jumlah-penerbangan-angkutan-udara-pada-agustus-2024-adalah-3-557-penerbangan.html>
- Brigato, L., & Locchi, L. (2020). A Close Look at Deep Learning with Small Data. In *2020 25th International Conference on Pattern Recognition (ICPR)* (pp. 2490-2497). IEEE. <https://doi.org/10.1109/ICPR48806.2021.9412492>
- Cankurt, S., & Subasi, Abdulhamit. (2015). Developing tourism demand forecasting models using machine learning techniques with trend, seasonal, and cyclic components. 33.
- Chowdhury, A. A., Das, A., Hoque, K. K. S., & Karmaker, D. (2021). A Comparative Study of Hyperparameter Optimization Techniques for Deep Learning. In *2020 25th International Conference on Pattern Recognition (ICPR)* (pp. 6495-6501). IEEE. <https://doi.org/10.1109/ICPR48806.2021.9412492>
- Draper, N. R., & Smith, H. (1998). *Applied Regression Analysis* (3rd ed., Vol. 326, Wiley Series in Probability and Statistics). John Wiley & Sons. ISBN 0471170828, 9780471170822.
- Google. (2024). *Google Trends*. Diakses dari <https://support.google.com/trends/?hl=en>
- Hammer, C., Kostroch, D. C., & Quiros-Romero, G. (2017). Big data: Potential, Challenges, and Statistical Implications. *IMF Staff Discussion Notes, No. 2017/006*. Diakses dari <https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2017/09/13/Big-Data-Potential-Challenges-and-Statistical-Implications-45106>
- Kementerian Pariwisata dan Ekonomi Kreatif. (2019). *5 destinasi super prioritas*. Diakses dari <https://info5dsp.kemenparekraf.go.id/>
- Kementerian Perhubungan. (2024). Keputusan Menteri Nomor 31/2024 (KM 31/2024) tentang Penetapan Bandar Udara Internasional. Jakarta: Kementerian Perhubungan. Diakses dari <https://jdih.kemenhub.go.id/peraturan>
- Kuhn, M. (2023). caret: Classification and Regression Training (Version 6.0-94). Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/caret/index.html>
- Lewis, C. D. (1982). *Industrial and Business Forecasting Methods: A Practical Guide To Exponential Smoothing And Curve Fitting*.

Butterworth Scientific.

Li, X., Pan, B., & Huang, X. (2017). Forecasting tourism demand with composite search index. *Tourism Management*, 57, 66-79. <https://doi.org/10.1016/j.tourman.2016.05.022>

Liashchynskyi, P.B., & Liashchynskyi, P. (2019). Grid Search, Random Search, Genetic Algorithm: A Big Comparison for NAS. *ArXiv*, abs/1912.06059.

Massicotte, P., & Eddelbuettel, D. (2022). Package 'gtrendsR': Perform and Display Google Trends Queries (Version 1.5.1). Dikutip dari <https://github.com/PMassicotte/gtrendsR/issues>

Ripley, B. D., & Venables, W. N. (2023). nnet: Feed-Forward Neural Networks and Multinomial Log-Linear Models (Version 7.3-19). Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/nnet/index.html>

Statcounter Global Stats. (2024). Search Engine Market Share Worldwide. Diakses dari <https://gs.statcounter.com/search-engine-market-share#monthly-202406-202407>

Statcounter Global Stats. (2024). Search Engine Market Share Indonesia. Diakses dari <https://gs.statcounter.com/search-engine-market-share/all/indonesia>