

Perbandingan Kewajaran dan Kepantasan Harga Penginapan dengan Analisis Statistik Geospasial dan Kecerdasan Buatan, Studi Kasus Pulau Bali sebagai Destinasi Wisata Utama

Handono Bayuadji, Shanti Pujilestari, Levyda, Ina Djamhur
Universitas Sahid Jakarta
e-mail: shanti_pujilestari@usahid.ac.id

Abstrak—Pulau Bali, sebagai destinasi pariwisata utama di Indonesia, menarik jutaan wisatawan setiap tahunnya. Penentuan harga penginapan yang sesuai dengan kualitas dan lokasi menjadi tantangan bagi pengusaha penginapan di Bali. Penentuan harga penginapan merupakan aspek krusial dalam industri perhotelan dan sektor pariwisata, yang harus mempertimbangkan lokasi geografis, fasilitas, kepuasan pelanggan, serta sarana dan prasarana di sekitarnya. Penelitian ini bertujuan menjelajahi dan membandingkan metode penentuan harga penginapan, melalui analisis statistik geospasial dan algoritma AI. Dengan pemahaman mendalam terhadap faktor-faktor yang memengaruhi harga penginapan, penelitian ini berupaya memberikan solusi yang akurat dan efisien dalam penentuan harga penginapan. Analisis statistik geospasial menunjukkan bahwa nilai rata-rata harga penginapan sejenis memiliki hubungan terbaik dengan seluruh algoritma AI yang digunakan. Algoritma AI yang paling sesuai dengan hasil analisis statistik geospasial adalah kelompok algoritma prediksi pada radius 3 km, dengan nilai korelasi sebesar 0.565171. Namun, perlu diperhatikan bahwa hasil penelitian ini mungkin kurang akurat akibat distorsi data yang berasal dari penginapan yang memasang harga tidak wajar. Fenomena ini dapat berdampak signifikan terhadap hasil analisis.

Kata Kunci— Analisis Statistik Geospasial; Kecerdasan Buatan; Harga Penginapan; Pariwisata; Pulau Bali.

I. PENDAHULUAN

Pulau Bali sebagai destinasi pariwisata utama di Indonesia, menarik jutaan wisatawan setiap tahunnya. Penentuan harga penginapan yang sesuai dengan kualitas dan lokasi merupakan tantangan tersendiri bagi pengusaha penginapan di Bali. Penentuan harga penginapan adalah salah satu aspek yang sangat penting dalam industri perhotelan dan sektor pariwisata.

Harga yang diberlakukan pada penginapan harus pantas dan wajar berdasarkan berbagai pertimbangan, termasuk lokasi geografis, fasilitas yang disediakan, kepuasan pelanggan, serta sarana dan prasarana di sekitarnya. Standarisasi dan penentuan kepantasan harga bertujuan untuk mengurangi perang harga yang sering terjadi antara penginapan yang berdampak pada kerugian atau penurunan mutu pelayanan. Pengusaha dan pengelola penginapan dituntut untuk mengikuti standarisasi harga baik yang tertulis maupun tidak tertulis sesuai dengan mekanisme yang terjadi di pasar [1].

Di dalam era digital saat ini, data menjadi relative lebih mudah untuk didapat. Peneliti memanfaatkan data yang tersebar di internet dengan cara melakukan *Web Scraping* untuk digunakan sebagai sumber data [2]. Selanjutnya data yang sudah terstruktur dapat digunakan untuk analisis statistik geospasial dan kecerdasan buatan (AI) dalam menentukan harga penginapan yang wajar dan pantas [3].

Masalah dalam penelitian ini adalah menentukan harga penginapan yang wajar dan pantas di Pulau Bali, yang merupakan tujuan wisata utama di Indonesia. Proses penentuan harga penginapan yang tepat merupakan tantangan bagi pemilik dan pengelola penginapan di Bali, mengingat faktor-faktor yang kompleks seperti lokasi geografis dan administratif, fasilitas yang tersedia, kepuasan pelanggan, serta aspek-aspek lingkungan sekitar. Terdapat juga masalah perang harga antar penginapan yang dapat mengganggu stabilitas pasar dan kualitas layanan. Oleh karena itu, penelitian ini berupaya untuk menemukan solusi yang lebih akurat dan efisien dalam menentukan harga penginapan di Bali dengan menggabungkan dan membandingkan analisis statistik geospasial dan kecerdasan buatan (AI).

Tujuan utama penelitian ini adalah menjelajahi dan membandingkan metode penentuan harga penginapan, terutama analisis statistik geospasial dan algoritma AI. Dengan pemahaman yang lebih dalam tentang faktor-faktor yang memengaruhi harga penginapan, penelitian ini berharap dapat memberikan solusi yang lebih akurat dan efisien dalam menentukan harga penginapan, yang akan bermanfaat bagi pelaku usaha penginapan, pemerintah sebagai regulator, dan masyarakat umum.

Penelitian ini memberikan kontribusi dalam beberapa aspek penting. Pertama, penelitian ini bertujuan untuk menambah pemahaman tentang penggunaan analisis statistik geospasial dalam menentukan kepantasan dan kewajaran harga penginapan, dengan mempertimbangkan pengaruh lokasi serta harga penginapan sejenis di sekitarnya. Kedua, penelitian ini akan membandingkan efektivitas dan hubungan antara analisis statistik geospasial dan metode kecerdasan buatan (AI) dalam konteks

penentuan harga penginapan sejenis. Ketiga, hasil penelitian ini diharapkan dapat memberikan dasar yang kuat bagi pengambilan keputusan yang lebih akurat dan efisien terkait penetapan harga penginapan.

II. TINJAUAN PUSTAKA

A. Analisis Statistik Geospasial

Analisis statistik geospasial adalah pendekatan analisis data statistik yang mempertimbangkan faktor geografis dan keruangan dalam penentuan keputusan [4] [5]. Dalam konteks penentuan harga penginapan, analisis geospasial memungkinkan peneliti untuk mengintegrasikan data lokasi, wilayah administrasi, sarana dan prasarana disekitar penginapan, serta faktor geografis lainnya ke dalam proses pengambilan keputusan. Dengan menggabungkan informasi statistik geospasial, dapat membantu menentukan faktor-faktor yang mempengaruhi harga penginapan. Metode ini dapat digunakan untuk mengidentifikasi pola harga di berbagai lokasi dan menghitung harga yang pantas berdasarkan lokasi spesifik.

Dalam “Determination of hedonic hotel room prices with spatial effect in Antalya” dalam penelitian ini dijelaskan cara untuk membuat model harga hedonik yang mencakup efek spasial dengan mengidentifikasi karakteristik fasilitas akomodasi yang mempengaruhi tarif kamar di area Antalya (Turki). Dalam konteks ini, data dari 1444 hotel di Antalya diperoleh. Karena ada hubungan spasial di antara mereka, sebuah model regresi spasial diusulkan. Analisis data spasial eksplanatori diterapkan pada variabel-variabel dengan regresi bertahap. Model Harga Hedonik dibangun menggunakan Model Lag Spasial dan menunjukkan bahwa model ini dapat diterapkan pada penelitian serupa [6].

B. Prediksi Menggunakan AI

Prediksi harga penginapan adalah salah satu aspek penting dalam industri perhotelan dan perjalanan. Metode prediksi ini mencoba untuk memahami pola harga dan faktor-faktor yang mempengaruhi perubahan harga penginapan. Prediksi harga dapat dilakukan dengan menggunakan berbagai algoritma, sebagai contoh adalah *Regresi*, *Neural Network*, dan *Deep Learning* yang digunakan dalam penelitian ini. Kelebihan prediksi harga dapat menghasilkan nilai numerik yang tidak ada pada data pelatihan [7]

Penelitian sebelumnya telah menunjukkan bahwa dengan memahami pola harga sebelumnya dan mempertimbangkan faktor-faktor seperti musim liburan, peristiwa khusus, dan permintaan pasar, kita dapat memprediksi harga penginapan yang sesuai [8]. Ini dapat membantu pemilik penginapan untuk mengoptimalkan keuntungan mereka dan konsumen dalam membuat keputusan yang lebih baik.

Dari data yang diambil di New York, ditemukan bahwa kualitas kamar dan lokasi sangat mempengaruhi harga kamar di industri hotel, tetapi faktor-faktor yang mempengaruhi harga kamar berbeda-beda di setiap jenis hotel. Dari penelitian ini hotel bisa mendapatkan manfaat dengan memahami harapan khusus dari pelanggan dalam segmen pasar tertentu, dan berusaha menyediakan fasilitas yang sesuai dengan harapan tersebut [9].

C. Klasifikasi Menggunakan AI

Klasifikasi menggunakan kecerdasan buatan (AI) adalah teknik yang digunakan untuk mengelompokkan data ke dalam kategori atau kelas yang berbeda. Dalam konteks penelitian ini, klasifikasi dapat digunakan untuk mengkategorikan penginapan ke dalam kelas harga tertentu berdasarkan berbagai atribut, seperti jenis akomodasi, tingkat bintang, fasilitas, sarana dan prasarana di sekitarnya wilayah administrasi serta data lainnya. Teknik-teknik klasifikasi AI, seperti *Naive Bayes*, *ID3*, dan *Decision Tree* yang digunakan dalam penelitian ini, dapat memproses data dan melakukan klasifikasi yang sesuai untuk setiap penginapan. Kekurangan dari algoritma ini adalah, klas yang dihasilkan hanya kelas yang ada pada label data pelatihan [10].

D. Kombinasi Analisis Statistik Geospasial dan AI

Kombinasi analisis statistik dan AI adalah pendekatan yang memanfaatkan kekuatan kedua metode tersebut untuk mencapai hasil yang lebih akurat dan relevan [11]. Dalam penelitian ini, kombinasi analisis statistik geospasial dan AI dapat memberikan wawasan yang lebih dalam dalam penentuan harga penginapan. Misalnya, analisis statistik geospasial dapat memberikan dasar pemahaman tentang pengaruh faktor-faktor geospasial, sementara AI dapat memproses data dengan lebih canggih dan mengidentifikasi pola-pola yang mungkin sulit dideteksi melalui metode statistik konvensional. Kombinasi ini dapat membantu dalam menentukan harga yang lebih tepat dan relevan berdasarkan konteks geospasial.

E. Kewajaran dan Kepantasan Harga Penginapan

Kewajaran harga penginapan mengacu pada harga penginapan dibandingkan penginapan sejenis lain di sekitarnya yang dianggap wajar dan adil oleh pelanggan dan penyedia penginapan. Penentuan harga yang pantas adalah keseimbangan antara keuntungan yang diinginkan oleh pemilik penginapan dan nilai yang diterima oleh pelanggan [12]. Memahami faktor-faktor yang mempengaruhi kepastian harga penginapan adalah tujuan penting dalam penelitian ini. Dengan mengkombinasikan analisis statistik geospasial, prediksi harga, klasifikasi menggunakan AI, dan analisis geospasial, penelitian ini bertujuan untuk membantu dalam menentukan harga yang pantas dengan lebih baik berdasarkan faktor-faktor geospasial dan preferensi pelanggan.

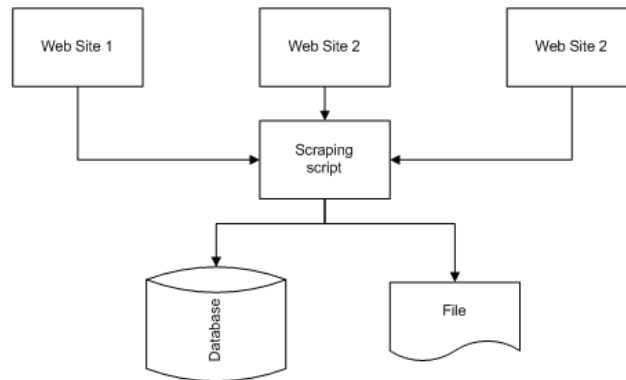
III. METODE PENELITIAN

A. Pengumpulan Data

Pengumpulan data merupakan langkah pertama dalam penelitian ini. Pengumpulan data dalam penelitian ini akan dilakukan menggunakan beberapa teknik yang sesuai dengan tujuan. Berikut adalah teknik pengumpulan data yang akan digunakan dalam pengumpulan data:

1) Web Scraping untuk Pengumpulan Data Harga Penginapan

Teknik web scraping akan digunakan untuk mengumpulkan data harga penginapan. Dalam penelitian ini, akan dilakukan scraping data harga penginapan dari situs booking hotel yang relevan di Provinsi Bali. Proses web scraping akan melibatkan penggunaan perangkat lunak Python, seperti BeautifulSoup dan Scrapy. Selain itu juga digunakan software yang dirancang khusus untuk mengambil data secara otomatis dari halaman web. Data hasil scraping akan disimpan dalam basis data MySQL dan file Excel untuk keperluan pengolahan dan analisis selanjutnya.



Gambar 1.

Web Scraping untuk Pengumpulan Data Harga Penginapan

2) Ekstraksi Data Koordinat

Selain data harga penginapan, data koordinat atau lokasi geografis penginapan dan fasilitas di sekitarnya juga akan dikumpulkan. Data koordinat ini akan digunakan untuk analisis geospasial dan pemetaan harga penginapan di Provinsi Bali. Proses ekstraksi data koordinat akan melibatkan penggunaan beberapa software sederhana yang dibuat khusus sesuai dengan format data dari situs Booking, Agoda Hotel dan Google Maps. Data koordinat yang diperoleh akan disimpan dalam basis data MySQL dan file Excel. Penggunaan MySQL sebagai basis data akan memungkinkan penyimpanan dan pengelolaan data yang efisien, sehingga mempermudah analisis statistic geospasial dan analisis dengan menggunakan beberapa macam algoritma AI.

B. Variabel Penelitian

Dalam penelitian ini, dilakukan operasionalisasi variabel-variabel penelitian untuk memungkinkan pengukuran dan analisis data dengan beberapa algoritma AI. Terdapat beberapa variable yang mempengaruhi harga penginapan seperti jenis, lokasi, sarana dan prasaran pendukung, ulasan dan peringkat, Ulasan dan peringkat dalam web booking penginapan juga dapat digunakan sebagai strategi pemasaran yang efektif untuk kesuksesan pemasaran penginapan [13].

1) Variabel Independen (Bebas)

Variabel bebas terdiri dari:

- JENIS_PENGINAPAN:** Variabel ini menggambarkan jenis penginapan seperti hotel, villa, guest house, atau homestay. Type data variable ini adalah nominal.
- RATING_REVIEW_SKOR:** Variabel ini mengindikasikan skor atau rating yang diberikan oleh pengguna atau pelanggan terhadap penginapan. Skor ini berupa angka dari 1 hingga 5 dengan type data numerik yang dapat dikonversi menjadi nominal
- RATING_REVIEW_NAMA:** Variabel ini menggambarkan kategori atau kualitas rating yang diberikan oleh pengguna atau pelanggan, seperti "Sangat Baik", "Baik", "Cukup", atau "Buruk". Type data variable ini adalah nominal
- RATING_REVIEW_COUNT:** Variabel ini mengindikasikan jumlah reviewer yang memberikan rating atau ulasan terhadap penginapan. Skala untuk variable ini adalah 1 sampai dengan 10, Type data variable ini adalah numerik
- KABUPATEN:** Variabel ini menggambarkan kabupaten di Provinsi Bali di mana penginapan berada, seperti Gianyar, Denpasar, Buleleng, dan sebagainya. Type data variable ini adalah nominal.
- LOKASI_WILAYAH:** Variabel ini menggambarkan lokasi geografis penginapan, seperti "Pantai Kuta", "Ubud", "Nusa Dua" dan lain sebagainya. Type data variable ini adalah nominal.
- HOTEL_PESAING_COUNT:** Variabel ini mengindikasikan jumlah hotel pesaing yang terletak dalam radius 1 km sampai 10 km dari penginapan. Variabel ini dapat dioperasionalkan dengan menghitung jumlah hotel pesaing yang terdaftar dalam daerah radius 1 km sampai 10 km dari setiap penginapan. Type data variable ini adalah numerik
- MART_COUNT:** Variabel ini mengindikasikan jumlah toko atau minimarket yang terletak dalam radius 5 km dari penginapan. Variabel ini dapat dioperasionalkan dengan menghitung jumlah toko atau minimarket yang terdaftar dalam daerah radius 1 km sampai 10 km dari setiap penginapan. Type data variable ini adalah numerik.
- SPBU_COUNT:** Variabel ini mengindikasikan jumlah pompa bensin atau SPBU (Stasiun Pengisian Bahan Bakar Umum) yang terletak dalam radius 1 km sampai 10 km dari penginapan. Type data variable ini adalah numerik.

- j. MALL_COUNT: Variabel ini mengindikasikan jumlah pusat perbelanjaan atau mall yang terletak dalam radius 1 km sampai 10 km dari penginapan. Type data variable ini adalah numerik.
- k. WISATA_COUNT: Variabel ini mengindikasikan jumlah obyek wisata yang terletak dalam radius 1 km sampai 10 km dari penginapan. Type data variable ini adalah numerik.
- l. PANTAI_JARAK_METER: Variabel ini mengindikasikan jarak penginapan terhadap pantai terdekat dalam satuan meter. Type data variable ini adalah numerik.

2) Variabel Dependen (Terikat)

Variabel HARGA_HOTEL (RATES_FROM) merupakan variabel dependen dalam penelitian ini. Variabel ini menggambarkan harga penginapan pada kondisi sesungguhnya. Variabel ini merupakan label yang digunakan dalam proses analisis data. Type data variable ini adalah numerik, namun untuk dapat menerapkan algoritma *klasifikasi AI*, dapat dilakukan konversi menjadi nominal. Untuk mengurangi jumlah kelas dalam proses klasifikasi, maka digunakan jenis mata uang dollar dan pembulatan.

C. Populasi dan Sampel

Populasi yang menjadi fokus penelitian ini mencakup semua penginapan di Provinsi Bali. Namun, mengingat keterbatasan waktu dan sumber daya yang ada, penelitian ini hanya akan memanfaatkan data dari situs pemesanan penginapan seperti Agoda, Booking, dan Google Maps yang berhasil di scraping, serta menggunakan data sekunder lain yang tersedia. Data peta yang digunakan untuk analisis spasial melibatkan peta administrasi, peta jaringan jalan, dan peta pendukung lainnya yang dapat diakses melalui Open Street Map (OSM) dan Open GIS Consortium (OGC).

Hasil penelitian sejenis dengan menggunakan situs *booking.com* sebagai sumber data menunjukkan bahwa jumlah penginapan di Pulau Bali selama periode pembatasan karena Covid-19 adalah 5.389 properti. Daerah dengan kepadatan tinggi di Bali ada di Bagian Selatan yaitu pada Kabupaten Badung. Dengan mempertimbangkan fasilitas pendukung pariwisata di Kabupaten Badung, kabupaten tersebut memiliki tingkat persaingan tertinggi [2].

D. Teknik Analisis Data

Pada tahap analisis data, akan digunakan beberapa teknik untuk memproses dan menganalisis data yang telah dikumpulkan. Berikut adalah teknik analisis data yang akan digunakan dalam penelitian ini.

1) Analisis Geospasial untuk Mencari Jumlah Fasilitas di Sekitar Hotel

Pada tahap ini, dilakukan analisis geospasial untuk mencari jumlah fasilitas di sekitar hotel. Analisis ini bertujuan untuk mendapatkan informasi tentang keberadaan dan distribusi fasilitas-fasilitas penting yang ada di sekitar hotel, seperti restoran, toko, taman, dan lain sebagainya. Hal ini dapat memberikan pemahaman yang lebih baik mengenai lingkungan sekitar hotel dan dapat menjadi faktor penting dalam menentukan harga penginapan.

Analisis geospasial dilakukan dengan memanfaatkan data koordinat yang telah diekstraksi dari situs booking hotel maupun data fasilitas yang diambil dari Google Maps. Data koordinat hotel digunakan sebagai titik acuan, sedangkan data fasilitas digunakan untuk mencari fasilitas-fasilitas yang berada dalam radius tertentu dari lokasi penginapan.

Proses analisis geospasial dapat melibatkan perangkat lunak GIS (*Geographic Information System*) seperti ArcGIS, QGIS, MapInfo atau perangkat lunak sejenis [2]. Dalam perangkat lunak tersebut, data koordinat hotel dan data fasilitas dapat diimpor dan dianalisis. Metode *buffering* atau radius searching dapat digunakan untuk mencari fasilitas-fasilitas yang berada dalam jarak tertentu dari hotel.

Hasil analisis geospasial ini akan memberikan informasi tentang jumlah fasilitas di sekitar hotel dalam kategori jarak yang ditentukan, yaitu 1 km, 2 km, 3 km, 5 km dan 10 km. Hal ini dapat menjadi pertimbangan penting dalam penentuan harga penginapan, karena fasilitas-fasilitas di sekitar hotel dapat mempengaruhi kenyamanan dan daya tarik bagi calon tamu. Hasil dari analisis geospasial ini selanjutnya akan digunakan untuk melakukan analisis statistik geospasial dan analisis dengan menggunakan algoritma AI.

Dengan melakukan analisis geospasial yang komprehensif, penelitian ini dapat memberikan pemahaman yang lebih baik tentang kondisi lingkungan sekitar hotel dan pengaruhnya terhadap harga penginapan. Hal ini dapat menjadi informasi berharga dalam pengembangan sistem prediksi harga penginapan yang lebih akurat dan berbasis data spasial.

2) Analisis Statistik Geospasial

Analisis statistik terhadap data geospasial dapat digunakan dengan mencari penginapan dengan jenis dan kelas yang sama, kemudian mengambil nilai rata-rata dan modulusnya. Dalam konteks ini, langkah-langkah yang dapat diikuti adalah sebagai berikut:

- a. Identifikasi penginapan dengan jenis dan kelas yang sama dalam radius 1 km, 2 km, 3 km, 5 km dan 10 km berdasarkan analisis geospasial yang telah dilakukan sebelumnya.
- b. Identifikasi *nilai rata-rata harga* untuk penginapan dengan jenis dan kelas yang sama. Ini dilakukan dengan menjumlahkan semua harga penginapan yang ditemukan dan membaginya dengan jumlah penginapan yang relevan.
- c. Identifikasi modus harga, yaitu harga yang paling sering muncul di antara penginapan dengan jenis dan kelas yang sama. Ini dapat memberikan informasi tentang harga yang umum atau dominan dalam kategori tersebut.

Analisis ini dapat memberikan pemahaman tentang kisaran harga untuk penginapan dengan jenis dan kelas yang sama dalam radius tertentu. Selain itu, perlu diperhatikan bahwa analisis ini hanya melibatkan penginapan dengan jenis dan kelas yang sama, dan tidak mempertimbangkan faktor-faktor lain yang dapat mempengaruhi harga seperti lokasi, fasilitas, sarana prasarana, review pelanggan dan kebijakan harga masing-masing penginapan. Jadi, menggunakan analisis spasial dan

statistik dengan mencari nilai rata-rata dan modus harga pada penginapan dengan jenis dan kelas yang sama dapat memberikan informasi umum tentang tren harga dalam kategori tersebut di suatu area.

E. Metode AI Klasifikasi dan Prediksi

Metode AI digunakan untuk mengelompokkan penginapan ke dalam kategori harga yang sesuai dan memprediksi harga penginapan. Metode AI mencakup algoritma klasifikasi dan prediksi, yang telah diuji dalam penelitian sebelumnya. Rincian metode AI sebagai berikut:

- 1) **Klasifikasi dengan Naive Bayes, ID3, Decision Tree, dan Random Forest:** Metode klasifikasi ini digunakan untuk mengkategorikan penginapan ke dalam kategori harga yang telah ditentukan. Harga dianggap sebagai label nominal, dan hasil prediksi akan mengambil salah satu nilai label harga pada data pelatihan [14].
- 2) **Prediksi dengan Neural Network, Deep Learning, dan Regresi:** Metode prediksi ini memperlakukan harga penginapan sebagai bilangan numerik. Hasil prediksi dapat menghasilkan harga yang lebih spesifik dan tidak terbatas pada harga yang ada dalam data pelatihan [15].

Metodologi analisis data yang digunakan mengintegrasikan analisis statistik geospasial dan metode AI untuk memberikan pemahaman yang lebih baik dalam menentukan harga yang pantas untuk penginapan sejenis dalam konteks geospasial. Penelitian ini diharapkan memberikan wawasan yang lebih mendalam dalam penentuan harga penginapan berdasarkan aspek geospasial dan prediksi harga.

F. Korelasi Hasil Perhitungan Harga

Analisis korelasi adalah alat statistik yang digunakan untuk mengevaluasi hubungan antara dua variabel atau lebih dalam dataset. Ini membantu dalam mengukur sejauh mana variabel-variabel tersebut berhubungan satu sama lain. Oleh karena itu, analisis korelasi dapat digunakan untuk mencari hubungan dan pola data [13] [16]. Tahap ini bertujuan untuk menganalisis korelasi atau hubungan antara hasil perhitungan harga penginapan dari analisis geospasial dengan hasil dari model AI, yang mencakup metode klasifikasi dan prediksi harga.

Tahap ini sangat penting dalam mengevaluasi konsistensi dan akurasi metode AI dalam konteks penentuan harga penginapan berdasarkan analisis geospasial. Hasil korelasi akan memberikan informasi yang lebih mendalam tentang bagaimana dua pendekatan ini berinteraksi dan sejauh mana mereka dapat digunakan secara bersamaan dalam penelitian penentuan kewajaran harga penginapan di Pulau Bali.

IV. HASIL DAN PEMBAHASAN

A. Hasil Pengambilan Data

Data yang berhasil dikumpulkan pada penelitian ini berjumlah 9.966 penginapan, berikut adalah rekapitulasi data berdasarkan pengelompokan beberapa variable penginapan di Pulau Bali.

1) Jumlah Penginapan Berdasarkan Level Bintang

Dalam pengelompokan ini, data penginapan akan dikelompokkan berdasarkan tingkat bintang yang dimiliki oleh penginapan tersebut, seperti 1 bintang, 2 bintang, 3 bintang, 4 bintang, atau 5 bintang. Setiap penginapan akan diklasifikasikan ke dalam salah satu kategori ini berdasarkan tingkat kualitas atau fasilitas yang ditawarkan. Jumlah penginapan akan dihitung dalam setiap kategori bintang untuk mendapatkan gambaran tentang distribusi penginapan berdasarkan level bintang di Pulau Bali.

Tabel 1.
Jumlah Penginapan Berdasarkan Level Bintang

No	Star_Rating	Jumlah
1	0	514
2	1	308
3	1.5	6
4	2	724
5	2.5	24
6	3	3,083
7	3.5	50
8	4	2,553
9	4.5	92
10	5	2,612

2) Jumlah Penginapan Berdasarkan Lokasi Kabupaten

Dalam pengelompokan ini, data penginapan akan dikelompokkan berdasarkan lokasi geografis mereka di kabupaten di Pulau Bali. Setiap penginapan akan diklasifikasikan berdasarkan kabupaten tempat mereka berada, seperti Badung, Denpasar, Gianyar, Karangasem, Tabanan, dan sebagainya. Jumlah penginapan akan dihitung dalam setiap kabupaten untuk melihat sebaran penginapan di seluruh Pulau Bali dan memahami distribusi geografis mereka.

Tabel 2.
 Jumlah Penginapan Berdasarkan Lokasi Kabupaten

No	Kabupaten	Jumlah
1	BADUNG	5,457
2	BANGLI	57
3	BULELENG	393
4	GIANYAR	2,533
5	JEMBRANA	32
6	KARANGASEM	217
7	KLUNGKUNG	467
8	KOTA DENPASAR	612
9	TABANAN	198

3) Jumlah Penginapan Berdasarkan Tipe Penginapan

Dalam pengelompokan ini, data penginapan akan dikelompokkan berdasarkan jenis atau tipe penginapan, seperti hotel, villa, resort, guest house, atau hostel. Setiap penginapan akan diklasifikasikan ke dalam salah satu kategori ini berdasarkan jenis layanan dan fasilitas yang disediakan. Jumlah penginapan akan dihitung dalam setiap kategori tipe penginapan untuk melihat variasi jenis penginapan yang tersedia di Pulau Bali.

Tabel 3.
 Jumlah Penginapan Berdasarkan Tipe Penginapan

No	Jenis Penginapan	Jumlah
1	Apartment	406
2	Bungalow	257
3	Cottage	32
4	Guesthouse	197
5	Homestay	233
6	Hostel	50
7	Hotel	786
8	House	561
9	Inn	2
10	Lodge	2
11	Motel	1
12	Resort	570
13	Tent	6
14	Villa	6,863

4) Distribusi harga berdasarkan jenis penginapan

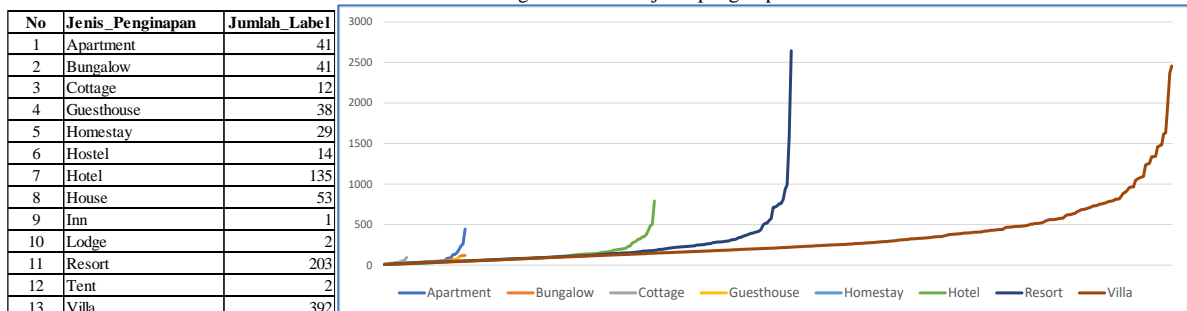
Distribusi harga berdasarkan jenis penginapan adalah aspek penting dalam konteks penelitian ini. Harga penginapan adalah variabel kunci yang digunakan sebagai label dalam analisis dengan Algoritma AI. Jumlah keseluruhan label pada data yang didapat dalam penelitian ini adalah 447 kelas, dengan penyebaran yang tidak merata untuk masing-masing jenis penginapan.

Untuk mendapatkan hasil analisis yang akurat dan relevan, penting bahwa distribusi harga di seluruh kategori jenis penginapan tetap seimbang. Ini berarti bahwa label harga yang digunakan dalam analisis harus merata terdistribusi di antara berbagai jenis penginapan, seperti hotel, villa, resort, guest house, dan lain sebagainya.

Distribusi harga yang merata pada seluruh jenis penginapan adalah kunci untuk mencegah bias dalam pembentukan model AI. Jika terdapat ketidakseimbangan dalam distribusi harga, algoritma AI dapat cenderung untuk memberikan bobot lebih pada kategori harga yang lebih dominan, sementara mengabaikan kategori harga lainnya. Hal ini dapat mengakibatkan hasil yang tidak akurat dan potensial untuk pengambilan keputusan yang kurang efektif dalam menentukan harga penginapan terutama untuk *kelompok algoritma klasifikasi*.

Oleh karena itu, dalam penelitian ini, penting untuk mengetahui distribusi harga di seluruh jenis penginapan di Pulau Bali. Dengan demikian, model AI yang dihasilkan akan mampu memberikan prediksi harga yang lebih akurat dan dapat digunakan sebagai dasar yang kuat untuk penentuan harga penginapan yang wajar dan pantas. Berikut grafik distribusi harga untuk beberapa jenis penginapan.

Tabel 4.
 Distribusi harga berdasarkan jenis penginapan



B. Analisis Data

1) Analisis Statistik Geospasial

Data geografis yang telah dihasilkan dapat digunakan untuk melakukan analisis geospasial. Dalam penelitian ini, analisis geospasial meliputi:

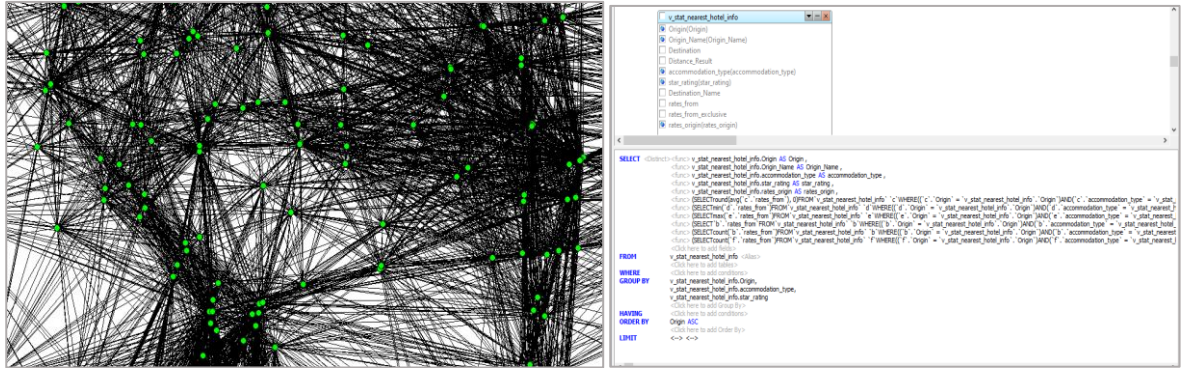
- Plotting data pada peta: memvisualisasikan lokasi penginapan dan fasilitas sekitarnya pada peta menggunakan QGIS dan MapInfo.
- Analisis jarak dan harga: menghitung jarak antara penginapan sejenis selanjutnya diikuti dengan mengitung harga rata-rata dan modus dalam radius yang telah ditentukan. Proses ini dilakukan menggunakan perangkat lunak GIS. Berikut adalah contoh Query yang digunakan untuk mencari penginapan dalam radius 3 km.

Radius = 3 km;

POINT(longitude, latitude) = Long/Lat Penginapan;

ALTER TABLE tb_penginapan ADD SPATIAL INDEX (location);

SELECT * FROM restaurants WHERE ST_DWithin(location, POINT(longitude, latitude), radius);



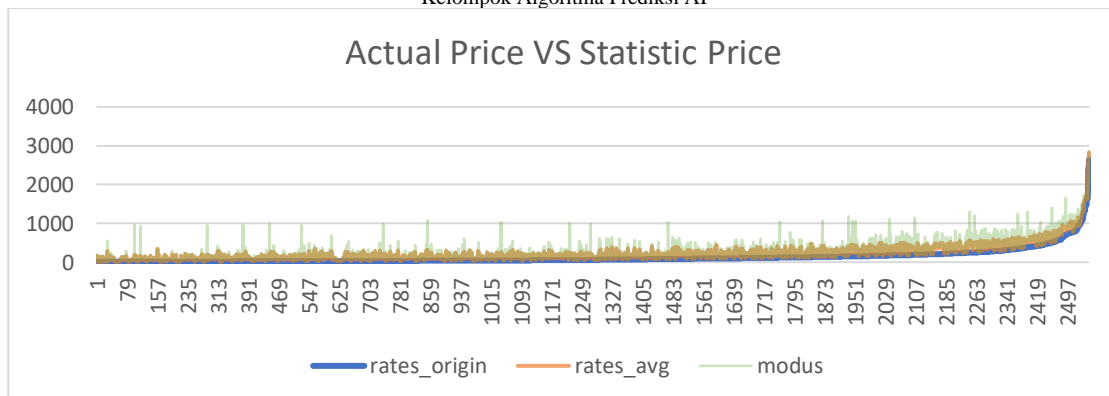
Gambar 2.
 Analisis Statistik Geospasial

2) Hasil Analisis Algoritma AI

Analisis dilakukan dengan menggunakan software Rapidminer, dengan menggunakan MySQL serta file Excel sebagai media penyimpanan data. Analisis dilakukan dengan menggunakan beberapa algoritma, pada data dengan radius yang ditentukan.

Row ... ↑	rates_from	prediction(rates_from)	star_rating	number_of_reviews	rating_average	JUMLAH_HOTEL	JUMLAH_MALL	JUMLAH_MART
1	150	147.483	100	22.005	84.160	4.350	1	8.816
2	28	27.725	60.400	16.552	79.210	40.219	69.538	81.763
3	89	70.074	80.200	17.031	79.210	10.630	16.231	34.868
4	47	45.372	80.200	27.349	81.190	29.332	77.154	86.974
5	265	274.872	100	15.855	87.130	84.601	16.231	29.658
6	104	140.230	80.200	13.710	80.200	82.275	23.846	21.842
7	71	60.327	60.400	8.940	77.230	92.417	23.846	29.658
8	36	38.384	60.400	10.635	80.200	51.849	39.077	60.921
9	35	23.754	40.600	16.839	78.220	28.588	77.154	86.974
10	247	293.563	100	7.751	84.160	86.095	23.846	24.447

Gambar 3.
 Kelompok Algoritma Prediksi AI



Gambar 4.
 Grafik perbandingan harga sebenarnya dan hasil perhitungan statistik geospasial

Tabel 7.
 Korelasi Antara Analisis Statistik Geospasial dengan Algoritma AI

Spatial Analisis	Decision Tree				
	1 km	2 km	3 km	5 lkm	10 km
rates_avg	0.036053	0.225061	0.325197	-0.012831	0.145118
modus	0.024185	0.149770	0.149815	-0.025436	0.113790
ID3					
	1 km	2 km	3 km	5 lkm	10 km
rates_avg	0.130579	0.131352	0.248958	0.218641	0.223841
modus	0.061446	0.111256	0.113117	0.149115	0.098952
Naïve Bayes					
	1 km	2 km	3 km	5 lkm	10 km
rates_avg	0.344167	0.352883	0.395337	0.311157	0.327027
modus	0.162425	0.131614	0.141715	0.120223	0.087862
Deep Learning					
	1 km	2 km	3 km	5 lkm	10 km
rates_avg	0.498557	0.502656	0.523876	0.519948	0.487918
modus	0.217354	0.199409	0.212407	0.204915	0.184177
Linear Regresion					
	1 km	2 km	3 km	5 lkm	10 km
rates_avg	0.563532	0.564108	0.565171	0.559854	0.484138
modus	0.203932	0.204294	0.203906	0.211989	0.163500
Neural Network					
	1 km	2 km	3 km	5 lkm	10 km
rates_avg	0.496912	0.491191	0.510023	0.497530	0.451399
modus	0.219422	0.212231	0.221245	0.222511	0.194026

Spatial Analisis	rates_avg	modus	Spatial Analisis	RATES
rates_avg	1.000000		rates_avg	0.340261
modus	0.396594	1.000000	modus	0.137333

Dari table tersebut diketahui analisis spasial harga rata-rata akan memiliki hubungan yang erat dengan Algoritma AI pada kelompok Prediksi yang terdiri dari *Deep Learning*, *Linier Regresion* dan *Neural Network* pada jarak/ radius 3 Km.

V. KESIMPULAN

A. Analisis Statistik Geospasial

1. Nilai analisis statistik geospasial yang memiliki hubungan terbaik dengan seluruh algoritma AI yang digunakan adalah nilai rata-rata (mean) harga penginapan sejenis.

B. Hubungan Analisis Statistik Geospasial dengan Algoritma AI

1. Algoritma AI yang paling sesuai dengan harga hasil analisis statistik geospasial adalah kelompok algoritma prediksi yang terdiri dari *Regresi*, *Deep Learning* dan *Neural Network* pada radius 3 km.
2. Nilai terbaik kelompok algoritma prediksi yang paling sesuai dengan analisis statistik geospasial adalah Regresi, dengan nilai korelasi 0.565171.
3. Nilai terbaik kelompok algoritma klasifikasi yang paling sesuai dengan analisis statistik geospasial adalah Naïve Bayes, dengan nilai korelasi 0.395337. Algoritma klasifikasi memperlakukan harga sebagai label nominal, sehingga nilai yang dihasilkan adalah nilai label yang ada pada data pelatihan yang maksimal berjumlah 447 kelas. Dengan penyebaran 477 kelas yang tidak terdistribusi merata, maka algoritma klasifikasi tidak dapat digunakan untuk analisis kewajaran harga penginapan dengan akurat.

C. Radius Paling Berpengaruh

1. Radius paling mempengaruhi harga kewajaran penginapan di Pulau Bali adalah 3 km dari lokasi penginapan, dengan faktor yang mempengaruhi berupa penginapan lain sejenis disekitarnya tanpa memperhitungkan factor lainnya.

Berdasarkan hasil penelitian ini, dapat disimpulkan bahwa kedua metode, yaitu analisis statistik geospasial dan metode AI, memiliki nilai dan keunggulan masing-masing dalam menentukan harga penginapan yang pantas di Pulau Bali. Pemahaman faktor-faktor geospasial yang diintegrasikan dengan hasil AI dapat memberikan kerangka kerja yang lebih lengkap dan akurat dalam penentuan harga penginapan.

Kesimpulan ini dapat menjadi dasar bagi pemilik dan pengelola penginapan, serta pemangku kepentingan lainnya, untuk membuat keputusan yang lebih baik terkait harga penginapan di Pulau Bali. Selain itu, penelitian ini juga membuka pintu bagi penelitian lebih lanjut dalam menggali potensi integrasi antara analisis geospasial dan kecerdasan buatan dalam konteks industri penginapan dan pariwisata.

VI. SARAN

Dari hasil analisis, penelitian ini kemungkinan kurang akurat akibat adanya distorsi data yang berasal dari beberapa penginapan yang memasang harga tidak wajar. Fenomena ini dapat memiliki dampak signifikan terhadap hasil analisis. Oleh karena itu, kehati-

hatian perlu diambil dalam menafsirkan hasil penelitian ini, dan langkah-langkah lanjutan diperlukan untuk mengidentifikasi dan mengatasi distorsi yang berasal dari praktik harga yang tidak wajar tersebut.

Untuk meningkatkan akurasi penelitian dan mengatasi distorsi yang disebabkan oleh praktik harga tidak wajar, diperlukan koordinasi dan konsultasi lebih lanjut dengan pihak berwenang, ahli pariwisata, dan ekonomi. Kerjasama tersebut diharapkan dapat memberikan landasan yang lebih kokoh dan jelas dalam merumuskan harga penginapan yang berdasarkan perhitungan baku dan sesuai dengan kondisi pasar yang sebenarnya. Keterlibatan pemerintah dan pemangku kepentingan lainnya akan memastikan bahwa rumusan perhitungan kewajaran dan kepatasan harga mencerminkan kondisi yang lebih realistis, dan solusi yang lebih efektif untuk mendukung pengembangan industri pariwisata di Pulau Bali.

DAFTAR PUSTAKA

- [1] A. A. A. N. S. R. Gorda and K. J. A. Sudharma, "Legalisasi Standar Tarif Hotel Dalam Ekosistem 'New Normal' Terintegrasi Bagi Pariwisata Bali Dampak Covid-19," *J. Pembang. Huk. Indones.*, vol. 5, no. 1, pp. 172–185, 2023, doi: 10.14710/jphi.v5i1.172-185.
- [2] Moch Sambas, Shanti Pujilestari, Listijono Setyopratinjo, and Rina Kurniawati, "Analysis of Lodging and Competition on the Island of Bali during Covid-19 with Big Data," *Int. J. Travel. Hosp. Events*, vol. 1, no. 3, pp. 214–228, 2022, doi: 10.56743/ijthe.v1i3.172.
- [3] C. H. Ku, Y. C. Chang, Y. Wang, C. H. Chen, and S. H. Hsiao, "Artificial intelligence and visual analytics: A deep-learning approach to analyze hotel reviews & responses," *Proc. Annu. Hawaii Int. Conf. Syst. Sci.*, vol. 2019-Janua, pp. 5268–5277, 2019, doi: 10.24251/hicss.2019.634.
- [4] O. Pomortseva, S. Kobzan, O. Voronkov, and A. Yevdokimov, "Geospatial modeling of the infrastructure facility optimal location," *E3S Web Conf.*, vol. 280, 2021, doi: 10.1051/e3sconf/202128011013.
- [5] A. Tariq *et al.*, "Forest fire monitoring using spatial-statistical and Geo-spatial analysis of factors determining forest fire in Margalla Hills, Islamabad, Pakistan," *Geomatics, Nat. Hazards Risk*, vol. 12, no. 1, pp. 1212–1233, 2021, doi: 10.1080/19475705.2021.1920477.
- [6] F. Yalcin, "Determination of hedonic hotel room prices with spatial effect in Antalya," *Econ. Soc. y Territ.*, vol. xviii, pp. 697–734, 2018, doi: 10.22136/est20181228.
- [7] L. Balyen and T. Peto, "Promising artificial intelligence–machine learning–deep learning algorithms in ophthalmology," *Asia-Pacific J. Ophthalmol.*, vol. 8, no. 3, pp. 264–272, 2019, doi: 10.22608/APO.2018479.
- [8] T. Hu and H. Song, "Analysis of Influencing Factors and Distribution Simulation of Budget Hotel Room Pricing Based on Big Data and Machine Learning from a Spatial Perspective," *Sustain.*, vol. 15, no. 1, pp. 0–18, 2023, doi: 10.3390/su15010617.
- [9] Z. Zhang, Q. Ye, and R. Law, "Determinants of hotel room price: An exploration of travelers' hierarchy of accommodation needs," *Int. J. Contemp. Hosp. Manag.*, vol. 23, no. 7, pp. 972–981, 2011, doi: 10.1108/095961111111167551.
- [10] S. K. Punia, M. Kumar, T. Stephan, G. G. Deverajan, and R. Patan, "Performance analysis of machine learning algorithms for big data classification: ML and ai-based algorithms for big data analysis," *Int. J. E-Health Med. Commun.*, vol. 12, no. 4, pp. 60–75, 2021, doi: 10.4018/IJEHMC.20210701.oa4.
- [11] M. Ivić, "Artificial intelligence and geospatial analysis in disaster management," *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.*, vol. 42, no. 3/W8, pp. 161–166, 2019, doi: 10.5194/isprs-archives-XLII-3-W8-161-2019.
- [12] J. Yu, "Exploring the role of healthy green spaces, psychological resilience, attitude, brand attachment, and price reasonableness in increasing hotel guest retention," *Int. J. Environ. Res. Public Health*, vol. 17, no. 1, 2020, doi: 10.3390/ijerph17010133.
- [13] K. A. Fachrudin, D. L. Tarigan, and M. F. Iman, "Analisis Rating dan Harga Kamar Hotel Bintang Lima di Indonesia," *J. Akuntansi, Keuangan, dan Manaj.*, vol. 3, no. 3, pp. 237–252, 2022, doi: 10.35912/jakman.v3i3.1107.
- [14] S. Singh, "Comparative Study Id3 , Cart and C4 . 5 Decision Tree Algorithm : a Survey," *Int. J. Adv. Inf. Sci. Technol.*, vol. 27, no. 27, pp. 97–103, 2014, doi: 10.15693/ijaist/2014.v3i7.47-52.
- [15] M. Al Shehhi and A. Karathanasopoulos, "Forecasting hotel room prices in selected GCC cities using deep learning," *J. Hosp. Tour. Manag.*, vol. 42, no. April, pp. 40–50, 2020, doi: 10.1016/j.jhtm.2019.11.003.
- [16] B. Pirouz, H. J. Nejad, G. Violini, and B. Pirouz, "The role of artificial intelligence, MLR and statistical analysis in investigations about the correlation of swab tests and stress on health care systems by COVID-19," *Inf.*, vol. 11, no. 9, 2020, doi: 10.3390/info11090454.