

# Pengaruh Electronic Analisis Persepsi Dan Preferensi Destinasi Wisata Berdasarkan Sentimen Dan *Tourist Experience*: Studi Kasus Pada Negara Indonesia Dan Vietnam Menggunakan *Text Classification*

## *Analysis Of Perceptions And Preferences For Tourist Destinations Based On Sentiment And Tourist Experience: A Case Study Of Indonesia And Vietnam Using Text Classification*

Fathiyah Hasanah<sup>1</sup>, Andry Alamsyah<sup>2</sup>, Dian Puteri Ramadhani<sup>3</sup>

<sup>1</sup> Manajemen Bisnis Telekomunikasi & Informatika, Fakultas Ekonomi dan Bisnis, Universitas Telkom, Indonesia, fatyano@student.telkomuniversity.ac.id

<sup>2</sup> Manajemen Bisnis Telekomunikasi & Informatika, Fakultas Ekonomi dan Bisnis, Universitas Telkom, Indonesia, andrya@telkomuniversity.ac.id

<sup>3</sup> Manajemen Bisnis Telekomunikasi & Informatika, Fakultas Ekonomi dan Bisnis, Universitas Telkom, Indonesia, dianpramadhani@telkomuniversity.ac.id

### **Abstract**

*This study was conducted to analyze tourist experiences in relation to perceptions and choices of tourist destinations in Indonesia and Vietnam. The background of this research is the need to understand the factors influencing tourists' decisions in choosing specific destinations. The main objective of this study is to identify and classify the most significant dimensions of the tourism experience that shape tourists' perceptions and decisions. The research method used involves collecting secondary data through web scraping techniques from tourist reviews on TripAdvisor, which are then analyzed using the BERT model for sentiment analysis and text classification. The research results show that nature-based entertainment dimensions dominate the tourist experience in Indonesia, while cultural entertainment is the main factor in Vietnam. The main conclusion of this study is that differences in tourists' perceptions between the two countries can guide destination managers to enhance attractiveness and visitor satisfaction.*

*Keywords-tourist experience, text classification, sentiment analysis*

---

### **Abstrak**

Penelitian ini dilakukan untuk menganalisis pengalaman turis dalam hubungannya dengan persepsi dan pilihan destinasi wisata di Indonesia dan Vietnam. Latar belakang dari penelitian ini adalah kebutuhan untuk memahami faktor-faktor yang memengaruhi keputusan wisatawan dalam memilih destinasi wisata tertentu. Tujuan utama dari penelitian ini adalah untuk mengidentifikasi dan mengklasifikasikan dimensi pengalaman wisata yang paling signifikan dalam membentuk persepsi dan keputusan wisatawan. Metode penelitian yang digunakan melibatkan pengumpulan data sekunder melalui teknik web scraping dari ulasan wisatawan di TripAdvisor, yang kemudian dianalisis menggunakan model BERT untuk analisis sentimen dan klasifikasi teks. Hasil penelitian menunjukkan bahwa dimensi hiburan berbasis alam mendominasi pengalaman wisata di Indonesia, sementara hiburan budaya menjadi faktor utama di Vietnam. Kesimpulan utama dari penelitian ini adalah bahwa perbedaan dalam persepsi wisatawan antara kedua negara dapat menjadi panduan bagi pengelola destinasi wisata untuk meningkatkan daya tarik dan kepuasan pengunjung.

**Kata Kunci-pengalaman wisata, klasifikasi teks, analisis sentimen**

---

## **I. PENDAHULUAN**

Teknologi kian berkembang pesat dan mengubah cara wisatawan dalam bertukar informasi. Internet menjadi salah satu sumber informasi utama mereka. Penggunaan *mobile devices*, internet, dan media sosial memberikan wisatawan

kemudahan untuk bertukar informasi secara *real-time*. Saat berkunjung ke destinasi wisata, mereka akan membagikan kesan, pengalaman, atau mencari informasi yang lebih spesifik mengenai destinasi tersebut (Xu et al., 2021). Dengan akses yang semakin luas terhadap internet, wisatawan memiliki kontrol yang lebih besar terhadap perencanaan perjalanan mereka secara bebas disesuaikan dengan preferensi pribadi mereka. Kebebasan tersebut menimbulkan pola pergerakan wisata yang kompleks (Ramadhani et al., 2021).

Hal ini dapat menjadi tantangan bagi pelaku industri pariwisata dalam beradaptasi terhadap pola baru wisatawan sehingga dapat memenuhi kebutuhan wisatawan. Ulasan destinasi wisata kini menjadi hal yang penting semenjak penggunaan internet dan media sosial oleh para wisatawan dalam membuat keputusan berkunjung meningkat. Tripadvisor merupakan salah satu platform wisata berskala global yang memfasilitasi ulasan personal, rating, dan pendapat mengenai destinasi yang telah dikunjungi. Platform ini memainkan peran penting bagi wisatawan untuk bertukar informasi, menjadi ruang untuk sumber referensi dalam memahami daya tarik suatu destinasi wisata. Ketersediaan data dalam bentuk ulasan pengunjung menjadi aset yang penting untuk dapat mempelajari persepsi dari *tourist experience*. Berdasarkan latar belakang dan fenomena di atas, penelitian ini bertujuan untuk mengetahui persepsi dan preferensi destinasi wisata wisatawan dengan memanfaatkan data ulasan wisatawan secara online, sehingga dapat menyediakan kebutuhan pariwisata yang lebih tepat sasaran.

## II. TINJAUAN LITERATUR

### A. Pariwisata

Pariwisata adalah aktivitas individu yang melibatkan perjalanan sementara ke tempat tujuan di luar tempat tinggal dan tempat kerja mereka, serta kegiatan yang mereka lakukan selama berada di tempat tujuan tersebut, seperti perjalanan wisata, berlibur, atau tamasya (Nazwirman, 2019). Menurut Undang-Undang No 10 Tahun 2009 tentang kepariwisataan, pariwisata didefinisikan sebagai beragam kegiatan wisata yang didukung oleh berbagai fasilitas dan layanan yang disediakan oleh masyarakat, pelaku usaha, pemerintah, dan pemerintah daerah.

### B. *Tourist Experience*

*Tourist Experience* merupakan fenomena yang kompleks dan multifaset, yang memegang peranan penting dalam industri pariwisata. Pengalaman-pengalaman ini mencakup berbagai macam aktivitas dan interaksi yang dilakukan oleh wisatawan saat mengunjungi destinasi, termasuk wisata melihat-lihat, imersi budaya, eksplorasi kuliner, aktivitas petualangan, dan interaksi dengan komunitas lokal (Fernandes & Cruz, 2016). Dimensi-dimensi *tourist experience* yang dianalisis meliputi: *Functional Benefits*, berupa manfaat praktis dan kenyamanan yang dirasakan oleh wisatawan. *Entertainment*, berupa aktivitas rekreasi dan budaya yang menarik wisatawan. *Environment*, berupa keindahan dan suasana fisik dari tempat wisata. *Learning*, berupa kesempatan untuk mendapatkan pengetahuan baru. *Service Providers*, berupa kualitas layanan yang diberikan oleh staf dan penyedia layanan. *Trust*, berupa tingkat kepercayaan wisatawan terhadap penyedia layanan dan keselamatan di destinasi.

### C. Big Data

Di dalam buku yang berjudul *Business Intelligence: A managerial Perspective on Analytics* (Sharda et al., 2013) big data mengacu pada sekumpulan data yang sangat besar dan kompleks sehingga tidak dapat disimpan dalam sebuah satuan penyimpanan konvensional. Big data memiliki karakteristik utama yaitu volume, velocity, variety, dan veracity. Volume mengacu pada banyaknya data, kecepatan mengacu pada kecepatan pembuatan dan pemrosesan data, ketepatan mengacu pada keandalan dan keakuratan data, dan variasi mengacu pada berbagai jenis data yang dihasilkan.

### D. *Data Collection*

*Data Collection* menggunakan Web scraping atau bisa disebut web harvesting, ekstraksi data web, dan penambahan data web merupakan pengumpulan data secara otomatis. Berbagai data dapat diambil menggunakan teknik mengikis dan data yang diperoleh merupakan data terstruktur. Hasil dari scraping web dapat di export ke database (CSV, XML, Excel). Scraping web dipopulerkan dengan istilah “screen scraping” sebagai alat untuk mengekstrak data dari representasi visual (vanden Broucke & Baesens, 2018). Web scraping adalah proses pengumpulan data melalui cara yang bebas dengan menggunakan API. Hal ini dilakukan dengan menulis program otomatis yang menghubungi web server, meminta data, dan kemudian mem-parsing data tersebut untuk mengekstrak informasi yang dibutuhkan.

E. *Text Mining*

Text mining didefinisikan sebagai proses menggali suatu pengetahuan implisit dari data tesktual (Bater Makhabel et al., 2017). Pengetahuan implisit merupakan hasil keluaran dari text mining yang tidak diberikan penyimpanan, maka harus dibedakan dari informasi yang diambil dari penyimpanan. Text mining digunakan sebagai text classification, clustering dan association. Text Mining merupakan suatu proses dalam mengidentifikasi pola dan menggali pengetahuan implisit.

F. *Text Classification*

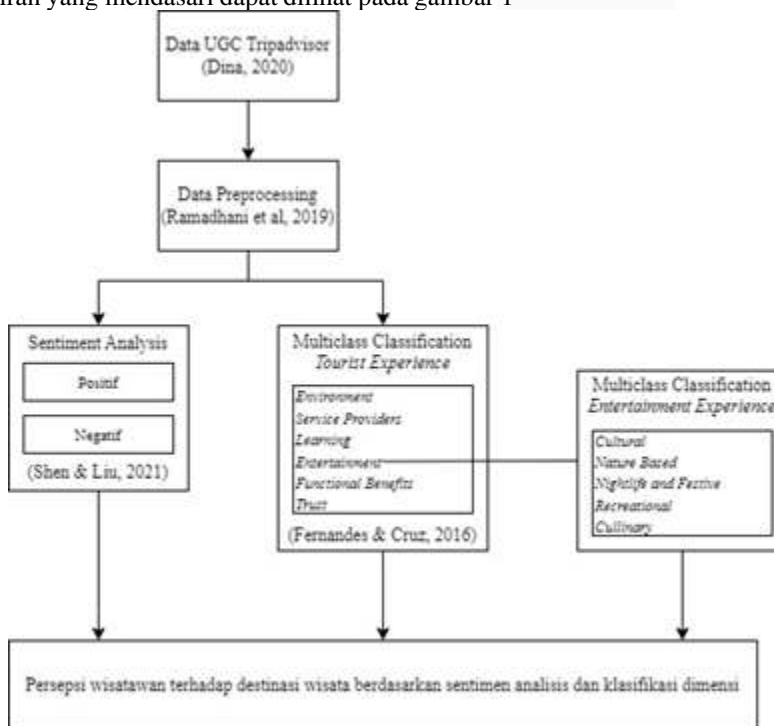
Klasifikasi teks atau text classification didefinisikan sebagai proses dalam menentukan suatu dokumen kedalam suatu kategori atau kelas tertentu. Dalam melakukan proses klasifikasi, terdapat beberapa algoritma yang bisa digunakan seperti Naïve Bayes, Support Vector Machine (SVM), k-Nearest Neighbor (KNN), Artificial Neural Network (ANN), dan Decision Tree (Hidayatullah & Maarif, 2016). Penelitian juga dapat menggunakan BERT (Bidirectional Encoder Representations from Transformers) untuk melakukan analisis sentimen dan multiclass classification.

G. *Sentiment Analysis*

Analisis sentimen (sentiment analysis) bisa disebut dengan opinion mining yang menjadi salah satu area penelitian paling aktif dalam pemrosesan bahasa sejak tahun 2000. Adapun tujuan dari analisis sentimen yaitu sebagai alat otomatis yang dapat mengekstrak informasi subjektif dari teks dalam bahasa alami, seperti opini dan sentimen, sehingga menciptakan pengetahuan yang terstruktur dan dapat digunakan oleh sistem dalam membuat keputusan. Analisis sentimen merupakan studi komputasi dari pendapat, sentimen, emosi, dan sikap dari individu. Analisis sentimen juga disebut opinion mining, yaitu bidang studi yang menganalisis pendapat, sentimen, penilaian, sikap, dan emosi individu terhadap entitas dan atributnya diungkapkan melalui teks tertulis. Analisis sentimen dikhususkan untuk memeriksa opini, sikap, dan emosi yang diekspresikan dalam teks (Shen & Liu, 2021).

H. Kerangka Penelitian

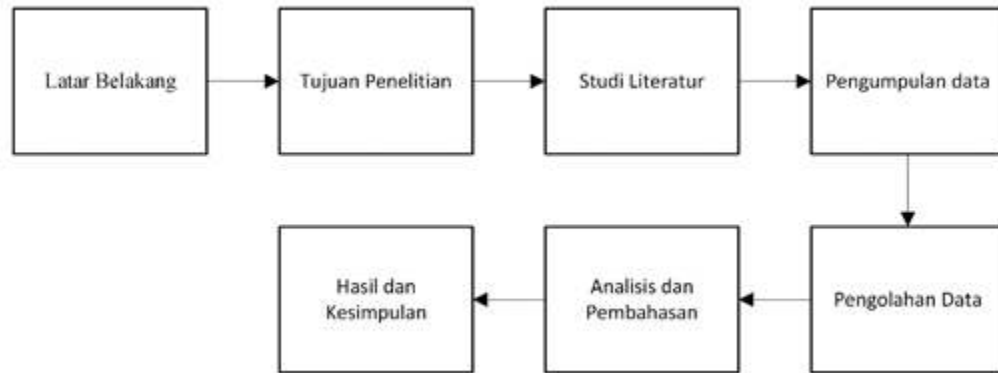
Kerangka pemikiran yang mendasari dapat dilihat pada gambar 1



Gambar.1.Kerangka Penelitian  
(Data Olahan Penulis, 2024)

### III. METODOLOGI PENELITIAN

Penelitian ini merupakan penelitian dengan metode kuantitatif yang menganalisis pengalaman turis terhadap destinasi wisata di kedua negara tersebut dengan menggunakan metode klasifikasi teks dan analisis sentimen. Penelitian ini menggunakan data sekunder yang diperoleh dari ulasan wisatawan di website TripAdvisor. Teknik pengumpulan data dilakukan melalui web scraping menggunakan ekstensi Web Scraper pada browser Google Chrome. Unit data yang diteliti adalah individu dikarenakan subjek pada penelitian ini adalah ulasan wisatawan. Data diambil selama Tripadvisor berjalan yaitu sekitar 2 dekade dari tahun 2004 hingga tahun 2023. Tahapan penelitian dapat dilihat pada gambar 2.



Gambar.2.Tahapan Penelitian  
(Data Olahan Penulis, 2024)

#### A. Latar Belakang

Pada latar belakang ini penulis melakukan identifikasi masalah fokus terhadap isu-isu yang memiliki dampak signifikan dan relevan bagi tujuan

#### B. Tujuan Penelitian

Penulis menentukan tujuan penelitian terkait fenomena masalah yang telah dirumuskan.

#### C. Melakukan Kajian Pustaka

Pada tahapan selanjutnya melakukan tinjau literatur yang relevan untuk memahami kerangka pemikiran dan temuan sebelumnya.

#### D. Pengumpulan data

Pengumpulan data dengan cara scraping data website tripadvisor menggunakan ekstensi google chrome Web Scraper. Data yang di scraping merupakan data Tripadvisor pada halaman Top Attraction in Indonesia dan Vietnam dengan mengambil masing-masing 25 destinasi terfavorit atau teratas.

#### E. Pengolahan data

Penulis melakukan tahap data *pre-processing* pada text. Tahapan *pre-processing text* terdiri dari *Transformation*, *Data Cleaning* dan *Labelling*. Serelah tahapan *pre-processing*, data dapat diklasifikasikan menggunakan *Text Clasification* dan diolah menggunakan *Sentiment Analysis* untuk menganalisis sentimen positif ataupun sentimen negatif.

#### F. Analisis dan pembahasan

Pada tahap ini, penulis akan melakukan identifikasi persepsi wisatawan menggunakan *Text Classification* dan *Sentiment Analysis* untuk mengetahui sentimen dan persepsi wisatawan berdasarkan *Tourist Experience* yang terbagi menjadi 6 dimensi

G. Hasil dan kesimpulan

Tahapan yang terakhir yaitu membuat hasil dari pengolahan data dan penjelasan terkait rumusan masalah dan tujuan penelitian yang dituliskan dalam format yang terstruktur serta membuat kesimpulan dan rekomendasi berdasarkan hasil penelitian.

IV. HASIL DAN PEMBAHASAN

A. Karakteristik data

Karakteristik dari data yang digunakan adalah data sekunder yang dikumpulkan dari website TripAdvisor. Proses pengumpulan data dilakukan melalui teknik web scraping, yang merupakan metodologi pengambilan data secara otomatis dari sebuah website. Data yang berhasil di-scrape terdiri dari beberapa elemen penting yang menjadi fokus utama dalam analisis ini, antara lain: nama pengguna, teks review, tanggal kunjungan, dan destinasi wisata yang dikunjungi. Nama pengguna merupakan identitas dari pengguna web TripAdvisor yang telah mengunjungi destinasi wisata. Tanggal kunjungan adalah kapan para pengguna TripAdvisor mengunjungi destinasi wisata tersebut. Destinasi wisata yang dikunjungi memberikan data lokasi spesifik yang dapat dikaitkan dengan karakteristik tertentu dari tempat tersebut. Nama pengguna dan tanggal kunjungan memberikan konteks mengenai waktu dan subjektivitas pengalaman yang diungkapkan dalam review. Teks review merupakan ulasan yang diberikan oleh para pengguna TripAdvisor yang telah mengunjungi destinasi wisata.

B. *Sentiment Analysis*

Analisis sentimen dilakukan dengan menggunakan model BERT. Model ini bertujuan untuk melakukan klasifikasi sentimen yang terbagi menjadi dua kategori, yaitu positif dan negatif. Hasil sentimen menggunakan BERT menunjukkan tingkat akurasi yang tinggi dalam memahami ulasan wisatawan.

Tabel 1 Hasil Analisis Sentimen

Negara	Positif	Negatif
Indonesia	77,18%	22.82%
Vietnam	77.96%	22.04%

Indonesia memiliki tingkat sentimen positif sebesar 77,18%, yang menunjukkan mayoritas wisatawan merasa puas dengan pengalaman mereka di destinasi wisata di Indonesia. Sentimen negatif yang sebesar 22,82% menunjukkan adanya beberapa aspek yang masih perlu diperbaiki untuk meningkatkan kepuasan wisatawan. Vietnam menunjukkan tingkat sentimen positif sebesar 77,96%, sedikit lebih tinggi dari Indonesia. Hal ini mengindikasikan bahwa mayoritas wisatawan juga merasa puas dengan pengalaman mereka di Vietnam, dengan sentimen negatif sebesar 22,04% yang menunjukkan adanya beberapa aspek yang masih perlu diperbaiki untuk meningkatkan kepuasan wisatawan.

Perbandingan ini menunjukkan bahwa meskipun kedua negara memiliki tingkat sentimen positif yang tinggi, Vietnam sedikit lebih unggul dalam hal kepuasan wisatawan. Kedua negara harus terus memperbaiki aspek-aspek yang mendapatkan sentimen negatif untuk meningkatkan pengalaman wisatawan secara keseluruhan.

C. *Multiclass Classification*

*Multiclass classification* adalah proses pengelompokan teks ke dalam beberapa kategori yang telah ditentukan sebelumnya. Ulasan wisatawan yang telah melalui proses tahap *pre-procesing* akan di klasifikasikan terhadap 6

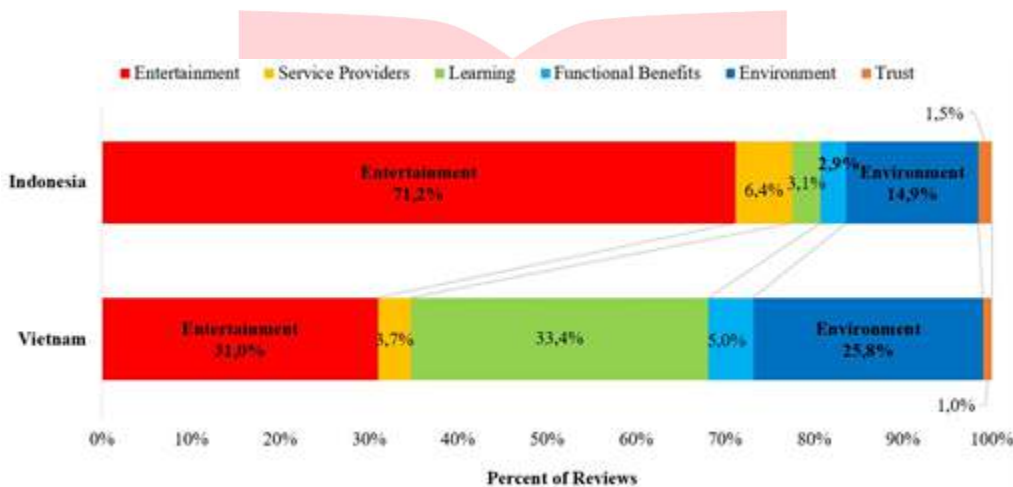
dimensi. Dalam konteks ini, klasifikasi dilakukan terhadap dimensi *Tourist Experience* yang meliputi *Entertainment, Service Providers, Learning, Functional Benefits, Environment, Trust*. *Training* dilakukan dengan menggunakan metode BERT dan didapatkan hasil diperoleh hasil seperti pada tabel 2

Tabel 2 Hasil Uji Performa

<i>Model</i>	<i>Accuracy</i>	<i>F1 Score</i>
BERT	76.75%	76.39%

Dari analisis data model pengalaman wisatawan menggunakan model BERT, akurasi model sebesar 76,75% dan skor F1 sebesar 76,39%, di mana akurasi model dapat diklasifikasikan sebagai cukup baik.

1. Hasil Proporsi Klasifikasi Dimensi *Tourist Experience* di Indonesia dan Vietnam



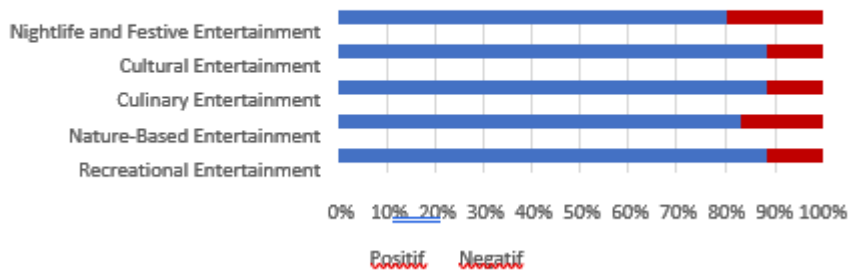
Gambar.3. Proporsi Klasifikasi Dimensi *Tourist Experience* di Indonesia (Data Olahan Penulis, 2024)

Dimensi hiburan mendominasi ulasan wisatawan di Indonesia, mencakup 71,2% dari total ulasan, yang menunjukkan bahwa aktivitas rekreasi dan budaya di negara ini sangat dihargai oleh pengunjung. Sebaliknya, di Vietnam, hiburan juga penting namun dengan porsi yang lebih rendah sebesar 31,0%, menandakan bahwa wisatawan di Indonesia lebih menikmati dan menghargai aktivitas hiburan yang ditawarkan. Sementara itu, kualitas layanan lebih sering dibahas di Indonesia, dengan 6,4% ulasan yang menyoroti interaksi dengan penyedia layanan, lebih tinggi dibandingkan dengan 3,7% di Vietnam, menunjukkan apresiasi yang lebih besar terhadap pelayanan yang diberikan di Indonesia.

Pembelajaran adalah dimensi yang sangat signifikan di Vietnam, di mana 33,4% ulasan menekankan pada aspek pendidikan dan budaya, jauh lebih tinggi dibandingkan dengan 3,1% di Indonesia, menunjukkan bahwa wisatawan di Vietnam lebih tertarik pada kegiatan edukatif dan budaya. Selain itu, manfaat fungsional dan keindahan lingkungan lebih sering disoroti oleh wisatawan di Vietnam dibandingkan dengan Indonesia, yang masing-masing mencapai 5,0% dan 25,8% ulasan. Kepercayaan, meskipun penting, menjadi dimensi yang paling jarang dibahas di kedua negara, dengan porsi yang sangat kecil, yakni 1,5% di Indonesia dan 1,0% di Vietnam, menunjukkan bahwa wisatawan lebih fokus pada aspek lain dari pengalaman mereka selama berada di destinasi wisata.

2. Hasil Analisis sentiment *Entertainment Tourist Experience* di Indonesia

### Analisis Sentimen Entertainment Tourist Experience di Indonesia

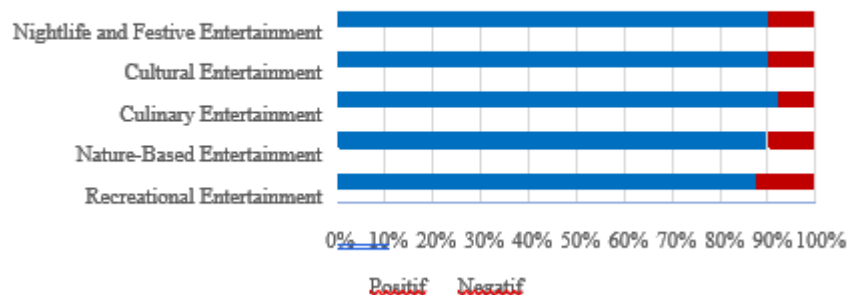


Gambar.4. Proporsi Analisis Sentimen *Tourist Experience* di Indonesia (Data Olahan Penulis, 2024)

Pada gambar 4 terlihat bahwa setiap dimensi hiburan di destinasi Indonesia mendapat persepsi sentimen positif yang dominan. Dimensi Cultural Entertainment menonjol dengan sentimen positif tertinggi sebesar 88,5%, menunjukkan apresiasi tinggi wisatawan terhadap kekayaan budaya dan acara budaya di Indonesia. Namun, dimensi Nightlife and Festive Entertainment memiliki sentimen negatif terbesar sebesar 19,3%, menunjukkan bahwa wisatawan sering mengeluhkan tentang keramaian, keamanan, atau kualitas acara malam hari di Indonesia.

### 3. Hasil Analisis sentiment *Entertainment Tourist Experience* di Vietnam

### Analisis Sentimen Entertainment Tourist Experience di Vietnam

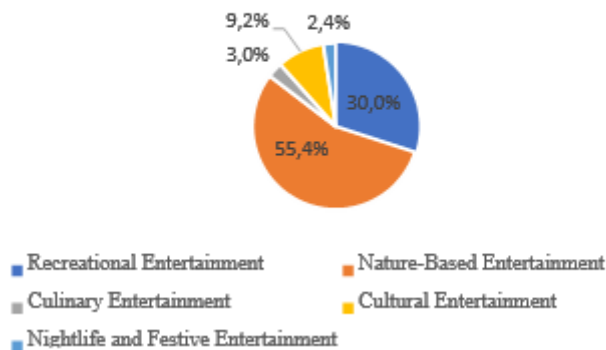


Gambar.4. Proporsi Analisis Sentimen *Tourist Experience* di Vietnam (Data Olahan Penulis, 2024)

Berdasarkan Tabel dan grafik analisis sentimen, terlihat bahwa setiap dimensi hiburan di Vietnam mendapat persepsi sentimen positif yang dominan. Dimensi Culinary Entertainment menonjol dengan sentimen positif tertinggi sebesar 92,1%, menunjukkan bahwa wisatawan sangat menghargai variasi dan kualitas makanan yang ditawarkan. Dimensi Nature-Based Entertainment dan Cultural Entertainment juga mendapatkan sentimen positif yang tinggi, masing-masing sebesar 90,0% dan 90,6%, menunjukkan apresiasi besar wisatawan terhadap keindahan alam dan kekayaan budaya Vietnam. Hiburan malam dan acara pesta (Nightlife and Festive Entertainment) juga menunjukkan kepuasan yang cukup tinggi dengan sentimen positif sebesar 89,1%, meskipun masih ada beberapa keluhan terkait kualitas acara malam. Dimensi Recreational Entertainment, meskipun memiliki sentimen positif sebesar 87,0%, mencatat ulasan negatif tertinggi sebesar 13,0%, menunjukkan adanya ruang untuk perbaikan dalam fasilitas rekreasi. Secara keseluruhan, wisatawan memiliki persepsi yang sangat positif terhadap berbagai dimensi hiburan di Vietnam, terutama pada aspek kuliner dan budaya, meskipun terdapat beberapa area yang memerlukan perhatian lebih, terutama dalam hiburan rekreasi.

4. Hasil Proporsi Klasifikasi Dimensi *Entertainment Tourist Experience* di Indonesia

**Klasifikasi Dimensi *Entertainment Tourist Experience* di Indonesia**



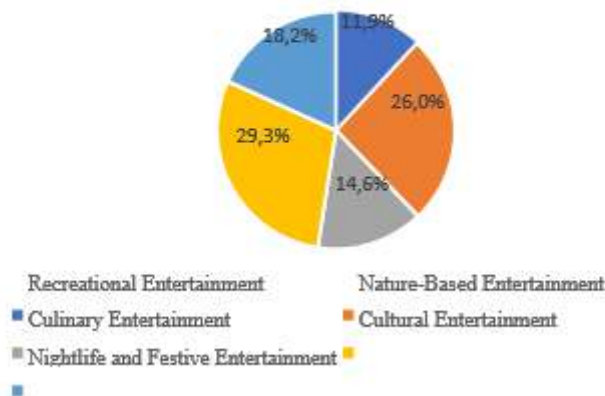
Gambar.4. Proporsi Klasifikasi Dimensi *Entertainment Tourist Experience* di Indonesia (Data Olahan Penulis, 2024)

Klasifikasi dimensi hiburan dari persepsi wisatawan di Indonesia menunjukkan bahwa Nature-based Entertainment mendominasi dengan proporsi sebesar 55,4%, menandakan bahwa keindahan alam Indonesia, seperti safari satwa liar dan ekowisata, merupakan daya tarik utama bagi wisatawan. Recreational Entertainment juga mendapatkan perhatian yang cukup besar, dengan proporsi 30,0%, menunjukkan bahwa aktivitas seperti taman hiburan dan olahraga petualangan sangat diminati oleh wisatawan yang mencari pengalaman menyenangkan dan memacu adrenalin. Sebaliknya, dimensi Culinary Entertainment hanya mendapat porsi ulasan sebesar 3,0%, yang menunjukkan bahwa aspek kuliner kurang disoroti dibandingkan dengan dimensi lainnya. Cultural Entertainment juga penting, dengan 9,2% ulasan yang mencerminkan apresiasi wisatawan terhadap kekayaan budaya Indonesia, meskipun masih lebih rendah dibandingkan dengan hiburan berbasis alam. Nightlife and Festive Entertainment mendapat perhatian paling kecil dengan hanya 2,4% ulasan, yang mungkin menunjukkan bahwa aspek ini kurang berkembang atau kurang menarik bagi wisatawan di Indonesia saat ini.



5. Hasil Proporsi Klasifikasi Dimensi *Entertainment Tourist Experience* di Vietnam

**Klasifikasi Dimensi *Entertainment Tourist Experience* di Vietnam**



Gambar.4. Proporsi Klasifikasi Dimensi *Entertainment Tourist Experience* di Indonesia (Data Olahan Penulis, 2024)

Klasifikasi dimensi hiburan dari persepsi wisatawan di Vietnam menunjukkan bahwa Hiburan Budaya menjadi daya tarik utama dengan proporsi sebesar 29,3%, yang mencerminkan apresiasi tinggi wisatawan terhadap aktivitas budaya seperti pertunjukan seni dan festival. Nature- based Entertainment juga menempati porsi yang signifikan sebesar 26,0%, menunjukkan bahwa keindahan alam Vietnam, termasuk wisata alam dan ekowisata, sangat dihargai oleh pengunjung. Hiburan kuliner mendapatkan perhatian yang cukup besar dengan 14,6%, menandakan bahwa wisatawan sangat menikmati pengalaman kuliner, seperti tur makanan dan festival kuliner. Hiburan malam dan acara pesta (Nightlife and Festive Entertainment) juga diminati dengan proporsi sebesar 18,2%, menunjukkan bahwa wisatawan menghargai kehidupan malam dan acara sosial di Vietnam. Sementara itu, Hiburan Rekreasi memiliki proporsi ulasan yang lebih kecil sebesar 11,9%, yang menunjukkan bahwa wisatawan di Vietnam kurang menyoroti aktivitas rekreasi dibandingkan dengan dimensi hiburan lainnya.

D. Hasil Pembahasan

Penelitian ini menggunakan data sekunder yang dikumpulkan dari website TripAdvisor melalui teknik web scraping dengan menggunakan ekstensi Web Scraper pada browser Google Chrome. Data yang berhasil di-scrape terdiri dari beberapa elemen penting yang menjadi fokus utama dalam analisis ini, termasuk nama pengguna, teks ulasan, tanggal kunjungan, dan destinasi wisata yang dikunjungi. Nama pengguna memberikan konteks mengenai siapa yang memberikan ulasan, sementara tanggal kunjungan memberikan informasi kapan ulasan tersebut dibuat. Destinasi wisata yang dikunjungi menyediakan data spesifik tentang lokasi wisata yang diulas. Data ini mencakup periode waktu dua dekade sesuai dengan waktu pertama kali TripAdvisor dibuat, dengan mengambil 25 destinasi wisata teratas di Indonesia dan Vietnam. Proses scraping dilakukan dengan mematuhi etika dan hukum yang berlaku terkait dengan penggunaan data publik untuk memastikan privasi pengguna tetap terjaga.

Data mentah yang diperoleh kemudian melalui tahap pre-processing untuk menjadi data yang lebih terstruktur. Tahapan pre-processing meliputi beberapa langkah penting, termasuk transformasi teks menjadi huruf kecil, penghapusan simbol dan emoji, serta penghilangan data yang tidak relevan atau duplikat. Tujuan dari pre-processing adalah untuk mengurangi inkonsistensi dalam data dan memastikan bahwa data yang digunakan untuk analisis lebih bersih dan konsisten. Setelah proses pembersihan, data siap untuk dianalisis lebih lanjut menggunakan berbagai metode analisis teks. Analisis sentimen dilakukan menggunakan model BERT (Bidirectional Encoder Representations from Transformers) untuk mengklasifikasikan sentimen wisatawan ke dalam dua kategori: positif dan negatif. BERT adalah model NLP yang mampu memahami konteks kata dalam teks ulasan dengan

mempertimbangkan kata-kata sebelum dan sesudahnya. Penggunaan analisis sentimen dalam penelitian ini juga dilakukan oleh Arianto dan Budi (2020) yang menyoroti pentingnya analisis sentimen berbasis aspek dalam memahami ulasan wisatawan di destinasi wisata Indonesia. Studi ini menekankan pada penggunaan data ulasan untuk mengidentifikasi kekuatan dan kelemahan destinasi wisata, sehingga pengelola dapat mengambil tindakan yang tepat untuk meningkatkan kualitas pengalaman wisatawan (Herawan et al., 2024).

Hasil analisis sentimen yang dihasilkan menunjukkan bahwa mayoritas wisatawan memberikan ulasan positif terhadap destinasi wisata di Indonesia dan Vietnam. Di Indonesia, tingkat sentimen positif mencapai 77,18%, sementara di Vietnam sedikit lebih tinggi dengan 77,96%. Sentimen negatif di Indonesia adalah 22,82% dan di Vietnam 22,04%. Ini menunjukkan bahwa wisatawan di kedua negara umumnya merasa puas dengan pengalaman mereka, meskipun ada beberapa aspek yang perlu diperbaiki. Penulis melakukan multiclass classification pada ulasan wisatawan terhadap enam dimensi Tourist Experience: Entertainment, Service Providers, Learning, Functional Benefits, Environment, dan Trust. Multiclass classification dilakukan menggunakan algoritma Neural Network dengan nilai akurasi 92,1%. Di Indonesia, dimensi Entertainment mendominasi dengan proporsi 71,2%, diikuti oleh Service Providers (64%) dan Environment (14,9%). Di Vietnam, dimensi Learning memiliki proporsi tertinggi (33,4%), menunjukkan bahwa wisatawan di Vietnam lebih tertarik pada aspek pendidikan dan budaya. Penelitian lebih lanjut dilakukan pada dimensi Entertainment, mengklasifikasikan dan menganalisis sentimen pada lima sub-dimensi: Recreational Entertainment, Nature-Based Entertainment, Culinary Entertainment, Cultural Entertainment, dan Nightlife and Festive Entertainment. Di Indonesia, Recreational Entertainment memiliki sentimen positif 88,1% dan negatif 11,9%, Nature-Based Entertainment memiliki sentimen positif 83,0% dan negatif 17,0%, Culinary Entertainment memiliki sentimen positif 87,8% dan negatif 12,2%, Cultural Entertainment memiliki sentimen positif 88,5% dan negatif 11,5%, dan Nightlife and Festive Entertainment memiliki sentimen positif 80,7% dan negatif 19,3%. Sementara di Vietnam, Recreational Entertainment memiliki sentimen positif 87,0% dan negatif 13,0%, Nature-Based Entertainment memiliki sentimen positif 90,0% dan negatif 10,0%, Culinary Entertainment memiliki sentimen positif 92,1% dan negatif 7,9%, Cultural Entertainment memiliki sentimen positif 90,6% dan negatif 9,4%, dan Nightlife and Festive Entertainment memiliki sentimen positif 89,1% dan negatif 10,9%.

Analisis ini menunjukkan bahwa setiap sub-dimensi memiliki tingkat kepuasan yang berbeda-beda, dengan Culinary Entertainment di Vietnam memiliki sentimen positif tertinggi. Penulis melakukan pembagian klasifikasi dan analisis sentimen pada dimensi Entertainment karena hasil multiclass classification menunjukkan bahwa dimensi Entertainment memiliki aspek yang cukup dominan dibandingkan dengan dimensi yang lain, terutama di Indonesia yang memiliki proporsi dimensi Entertainment sebanyak 71,3%. Seperti halnya penelitian dari Luo, Fan, dan Shang (2022), yang menemukan bahwa faktor hiburan dan pengalaman wisata yang menarik memainkan peran penting dalam meningkatkan kepuasan wisatawan di destinasi wisata. Mereka menekankan bahwa hiburan yang terencana dengan baik dapat meningkatkan niat kunjungan kembali wisatawan ke destinasi tersebut (Herawan et al., 2024).

Klasifikasi pada dimensi Entertainment menunjukkan bahwa wisatawan di Indonesia lebih menyoroti Nature-Based Entertainment (55,4%) dan Recreational Entertainment (30%). Aktivitas seperti safari satwa liar, jalur alam, dan ekowisata sangat dihargai di Indonesia. Sedangkan di Vietnam, Cultural Entertainment (29,3%) dan Nature-Based Entertainment (26%) menjadi fokus utama. Aktivitas budaya seperti pertunjukan seni, musik tradisional, dan festival budaya sangat dihargai oleh wisatawan di Vietnam. Hasil penelitian menunjukkan bahwa wisatawan umumnya memiliki persepsi positif terhadap berbagai dimensi hiburan di kedua negara, dengan Vietnam sedikit lebih unggul dalam beberapa aspek. Nature-Based Entertainment dan Cultural Entertainment mendominasi di kedua negara, menunjukkan bahwa wisatawan sangat menghargai keindahan alam dan kekayaan budaya. Namun, ada beberapa aspek yang masih memerlukan perbaikan, terutama pada dimensi Nightlife and Festive Entertainment di Indonesia dan Recreational Entertainment di Vietnam. Peningkatan kualitas dan promosi yang lebih baik pada aspek-aspek ini diharapkan dapat meningkatkan kepuasan wisatawan secara keseluruhan di kedua negara. Analisis ini memberikan wawasan berharga bagi pengelola destinasi wisata untuk memahami dan meningkatkan pengalaman wisatawan sesuai dengan persepsi mereka.

## V. KESIMPULAN DAN SARAN

### A. Kesimpulan

Berdasarkan hasil penelitian, beberapa kesimpulan utama dapat diambil sebagai berikut:

1. Ulasan wisatawan di Indonesia dan Vietnam umumnya memiliki sentimen positif, dengan Indonesia mencatat 77.18% sentimen positif dan Vietnam sedikit lebih tinggi dengan 77.96%. Meskipun wisatawan cenderung puas, masih ada beberapa aspek yang memerlukan perbaikan di kedua negara.
2. Persepsi wisatawan diklasifikasikan dalam enam dimensi kognitif: Entertainment, Service Providers, Learning, Functional Benefits, Environment, dan Trust. Di Indonesia, dimensi Entertainment mendominasi (71.2%), sedangkan di Vietnam, Learning menjadi yang paling menonjol (33.4%), menunjukkan ketertarikan wisatawan Vietnam pada aspek pendidikan dan budaya.
3. Pada dimensi Entertainment, wisatawan di Indonesia lebih menyoroti Nature-Based Entertainment (55.4%) dan Recreational Entertainment (30%), dengan mayoritas sentimen positif. Di Vietnam, Cultural Entertainment (29.3%) dan Nature-Based Entertainment (26%) menjadi fokus utama, dengan sentimen positif tertinggi pada Cultural Entertainment (90.6%). Meskipun secara keseluruhan persepsi wisatawan terhadap hiburan sangat positif, beberapa aspek seperti Nightlife and Festive Entertainment di Indonesia dan Recreational Entertainment di Vietnam masih memerlukan peningkatan.

## B. Saran

### 1. Aspek Teoritis:

- a. Studi ini mencakup Indonesia dan Vietnam, tetapi penelitian selanjutnya disarankan untuk melibatkan lebih banyak negara guna mendapatkan pemahaman yang lebih komprehensif tentang persepsi wisatawan lintas budaya.
- b. Meskipun penelitian ini menggunakan model BERT untuk analisis sentimen, penelitian lanjutan dapat mengeksplorasi model yang lebih kompleks atau menggabungkan teknik NLP lainnya untuk meningkatkan akurasi analisis.
- c. Penelitian mendatang dapat memperdalam analisis citra destinasi dengan menambahkan dimensi afektif, untuk memahami bagaimana perasaan wisatawan mempengaruhi pengalaman mereka.
- d. Disarankan untuk menganalisis pola pergerakan wisatawan menggunakan data seperti geotagging untuk memahami preferensi destinasi secara lebih mendalam.

### 2. Aspek Praktis:

- a. Peningkatan kualitas infrastruktur dan aksesibilitas, terutama di Indonesia, perlu difokuskan untuk mengurangi sentimen negatif terkait aspek tersebut.
- b. Pengelola destinasi wisata di Indonesia dapat meningkatkan promosi budaya lokal, meniru kesuksesan Vietnam dalam menarik wisatawan melalui aspek Cultural Entertainment.
- c. Mengembangkan program edukatif dan interaktif, seperti workshop budaya atau tur pendidikan, dapat meningkatkan pengalaman wisatawan di kedua negara.
- d. Penting bagi pengelola destinasi untuk aktif merespons ulasan wisatawan dan menangani masalah yang diangkat, guna meningkatkan reputasi dan kepercayaan wisatawan.

## REFERENSI

- Arabadzhieva - Kalcheva, N., & Kovachev, I. (2022). Comparison of BERT and XLNet accuracy with classical methods and algorithms in text classification. *2021 International Conference on Biomedical Innovations and Applications (BIA)*, 74–76. <https://doi.org/10.1109/BIA52594.2022.9831281>
- Bater Makhabel, Pradepta Mishra, Nathan Danneman, & Richard Heimann. (2017). *R: Mining spatial, text, web, and social media data*. Packt Publishing Ltd.
- Camilleri, M. A. (2017). The Tourism Industry: An Overview. *MRN: Courses*. <https://api.semanticscholar.org/CorpusID:168772122>
- Carr, A., Ruhanen, L., & Whitford, M. (2016). Indigenous peoples and tourism: the challenges and opportunities for sustainable tourism. *Journal of Sustainable Tourism*, 24(8–9), 1067–1079. <https://doi.org/10.1080/09669582.2016.1206112>
- Dina, N. Z. (2020). Tourist sentiment analysis on TripAdvisor using text mining: A case study using hotels in Ubud, Bali. In *African Journal of Hospitality, Tourism and Leisure* (Vol. 9, Issue 2). <http://www.ajhtl.com>
- Dwityas, N. A. (2016). Komunikasi dan Pariwisata: Peran User Generated Content bagi Traveler dalam Media Sosial. *Jurnal Simbolika*, 2(1). <https://www.researchgate.net/publication/327248941>

- Eastern Daylight Time. (2022). *he World Is Moving Beyond Big Data, According to Ocient Survey of 500 Data and Technology Leaders*. A Berkshire Hathaway Company.
- Fernandes, T., & Cruz, M. (2016). Dimensions and outcomes of experience quality in tourism: The case of Port wine cellars. *Journal of Retailing and Consumer Services*, 31, 371–379. <https://doi.org/10.1016/j.jretconser.2016.05.002>
- Herawan, T., Anggraini, F. D., Ihalauw, J. J. O. I., Hendratono, T., Damiasih, D., & Suhendroyono, S. (2024). *Visualizing Trends in Tourism Entertainment Researches: Bibliometric Analysis Using the Scopus Database* (pp. 498–504). [https://doi.org/10.1007/978-3-031-48465-0\\_67](https://doi.org/10.1007/978-3-031-48465-0_67)
- Hidayatullah, A. F., & Maarif, M. R. (2016). Penerapan Text Mining dalam Klasifikasi Judul Skripsi. *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)*.
- Indrawati. (2018). *Metode penelitian kualitatif: manajemen dan bisnis konvergensi teknologi informasi dan komunikasi* (Rachmi, Ed.; 1st ed.). Refika Aditama.
- Khotimah, N. (2017). PENGEMBANGAN PARIWISATA ALAM BERBASIS LINGKUNGAN. *Geomedia: Majalah Ilmiah Dan Informasi Kegeografian*, 6(2). <https://doi.org/10.21831/gm.v6i2.15416>
- Losekoot, E., Lasten, E., & Cuong, T. P. (2019). *Tourism Education in Vietnam* (pp. 125–142). [https://doi.org/10.1007/978-981-13-2613-4\\_8](https://doi.org/10.1007/978-981-13-2613-4_8)
- Nazwirman, N. (2019). ANALISIS NIAT PERILAKU MELALUI PENGALAMAN WISATAWAN STUDI PADA TAMAN MINI INDONESIA INDAH JAKARTA. *ISOQUANT : Jurnal Ekonomi, Manajemen Dan Akuntansi*, 3(2), 55. <https://doi.org/10.24269/iso.v3i2.288>
- Pine, J., & Gilmore, J. H. (1998). Welcome to the Experience Economy. *Economic Cycles And Trends*, 97–105.
- Ramadhani, D. P., Alamsyah, A., Atmaja, M. N., & Panjaitan, J. N. C. G. (2021). Mapping Complex Tourist Destination Preferences: Network Perspectives. *2021 9th International Conference on Information and Communication Technology (ICoICT)*, 219–224. <https://doi.org/10.1109/ICoICT52021.2021.9527445>
- Ramadhani, D. P., Alamsyah, A., Febrianta, M. Y., & Damayanti, L. Z. A. (2024). Exploring Tourists' Behavioral Patterns in Bali's Top-Rated Destinations: Perception and Mobility. *Journal of Theoretical and Applied Electronic Commerce Research*, 19(2), 743–773. <https://doi.org/10.3390/jtaer19020040>
- Ramadhani, D. P., Alamsyah, A., & Wicaksono, M. B. (2019). EKSPLORASI PEMIMPIN OPINI UNTUK ALTERNATIF PENDUKUNG PEMASARAN PT. NET MEDIATAMA INDONESIA MENGGUNAKAN METODE ANALISIS JEJARING SOSIAL DI TWITTER. *SOSIOHUMANITAS*, 21(1), 73–78. <https://doi.org/10.36555/sosiohumanitas.v21i1.1002>
- Rogerson, C. M., & van der Merwe, C. D. (2016). Heritage tourism in the global South: Development impacts of the Cradle of Humankind World Heritage Site, South Africa. *Local Economy: The Journal of the Local Economy Policy Unit*, 31(1–2), 234–248. <https://doi.org/10.1177/0269094215614270>
- Sajib, S. M. S. al, Nicolli, F., & Alietti, A. (2022). Problematizing tourism for conservation: An eco-cultural critique on sustainability. *European Journal of Cultural Management and Policy*, 12. <https://doi.org/10.3389/ejcmp.2022.11094>
- Sekaran, U., & Bougie, R. (2016). *Research Methods For Business: A Skill Building Approach* (7th ed.). John Wiley & Sons.
- Sharda, R., Delen, D., & King, D. (2013). *Business Intelligence: A Managerial Perspective on Analytics* (3rd ed.). Pearson Education.
- Shen, Y., & Liu, J. (2021). Comparison of Text Sentiment Analysis based on Bert and Word2vec. *2021 IEEE 3rd International Conference on Frontiers Technology of Information and Computer (ICFTIC)*, 144–147. <https://doi.org/10.1109/ICFTIC54370.2021.9647258>
- Tripadvisor. (2019). *Tripadvisor*. <https://tripadvisor.mediaroom.com/ID-about-us>
- World Bank Group. (2023). *International tourism, number of arrivals*. World Tourism Organization. <https://data.worldbank.org/indicator/ST.INT.ARVL?end=2024&start=2012>
- Xu, H., Lovett, J., Cheung, L. T. O., Duan, X., Pei, Q., & Liang, D. (2021). Adapting to social media: the influence of online reviews on tourist behaviour at a world heritage site in China. *Asia Pacific Journal of Tourism Research*, 26(10), 1125–1138. <https://doi.org/10.1080/10941665.2021.1952460>

Zahra Syahputri, A., Della Fallenia, F., Syafitri, R., Lubis, R. N., Wulan, S., & Lubis, D. (2023). Tarbiyah: Jurnal Ilmu Pendidikan dan Pengajaran. *Tarbiyah: Jurnal Ilmu Pendidikan Dan Pengajaran*, 2(1), 160–164. [https://jurnal.diklinko.id/index.php/tarbiyah/](https://jurnal.diklinko.id/index.php/tarbiyah)

