

BAB II

KAJIAN PUSTAKA

2. 1. Tinjauan Pustaka

2.1.1 Konsep Dasar Pariwisata

Pariwisata atau *tour* dalam Bahasa Inggris, mempunyai makna perjalanan menuju tempat yang dinilai memiliki sisi estetis, sejarah maupun hiburan sesuai tujuan masing-masing wisatawan. Terdapat empat tujuan orang melakukan wisata yaitu (Rosita Butarbutar dkk., 2021):

1. Mencari sesuatu untuk dilihat.
2. Mencari sesuatu untuk dilakukan.
3. Mencari sesuatu untuk dibeli.
4. Mencari sesuatu untuk diketahui.

Saat ini, industri pariwisata telah berkembang menjadi salah satu industri terbesar di seluruh dunia. Di banyak negara berkembang, pariwisata memegang peranan penting dalam pengembangan ekonomi, terutama sebagai sumber devisa yang signifikan. Industri ini tidak hanya menciptakan peluang kerja, tetapi juga mendorong investasi dalam infrastruktur dan layanan, serta memperkenalkan budaya dan keindahan alam suatu negara kepada wisatawan internasional. Dengan pertumbuhan yang pesat, pariwisata menjadi pilar utama dalam strategi ekonomi banyak negara, berkontribusi pada peningkatan kesejahteraan masyarakat setempat dan memperkuat hubungan antarbangsa. (Lopes & Rodríguez-López, 2022).

Pariwisata merupakan salah satu sektor usaha jasa yang memiliki peluang besar untuk berkembang, terutama karena destinasi baru terus bermunculan di berbagai tempat (Eddyono, 2021). Dalam menghadapi banyaknya pilihan destinasi, pelaku wisata dihadapkan pada tantangan untuk memenuhi kebutuhan wisatawan yang mencari pengalaman memuaskan dalam waktu yang efektif. Wisatawan cenderung mempertimbangkan berbagai faktor seperti fasilitas, waktu, biaya, jarak, transportasi, dan keamanan sebelum memutuskan tujuan perjalanan mereka (Satria dkk., 2018). Selain itu, mereka juga dituntut untuk membuat keputusan cepat dalam memilih objek wisata yang menawarkan risiko rendah dan tingkat keramaian yang

sesuai dengan preferensi pribadi (Krakover & Corsale, 2021). Oleh karena itu, kemampuan untuk memahami dan menyesuaikan diri dengan preferensi dan kebutuhan wisatawan menjadi sangat penting bagi pelaku industri pariwisata yang ingin tetap kompetitif dan berhasil.

Menurut penelitian Naranjo dkk. (2022), salah satu faktor kunci dalam mengembangkan pariwisata adalah memperhatikan ketidakpuasan wisatawan, karena menjaga kepuasan wisatawan secara berkelanjutan sangat penting untuk keberhasilan jangka panjang. Kepuasan ini harus dijaga tidak hanya pada objek wisata yang dikelola oleh pemerintah, tetapi juga pada destinasi yang dikelola oleh pihak swasta. Destinasi wisata dapat dianggap sebagai paket yang terdiri dari berbagai layanan dan fasilitas yang saling bekerja sama untuk menarik wisatawan (Naranjo dkk., 2022).

Eddyono (2021) menyoroti pentingnya memahami Tourism Area Life Cycle (TALC) atau Siklus Hidup Destinasi Wisata, dengan tujuan utama mendeteksi ketidakpuasan wisatawan. Dengan mengetahui titik-titik ketidakpuasan ini, kebijakan dapat disusun untuk meningkatkan daya tarik dan kepuasan destinasi, sehingga tercipta pariwisata yang berkelanjutan dan mampu memenuhi harapan wisatawan (Eddyono, 2021).

2.1.2 Faktor-faktor Yang Mempengaruhi Pemilihan Destinasi Pariwisata

Pemilihan destinasi wisata dapat dilakukan dengan berbagai metode, salah satunya adalah menggunakan model multinomial logit. Penelitian yang dilakukan oleh Tang dkk. (2020) membuktikan bahwa faktor-faktor seperti umur, jenis kelamin, moda transportasi, rekan seperjalanan, dan lama waktu liburan mempengaruhi pemilihan destinasi wisata. Model ini memberikan wawasan tentang bagaimana karakteristik demografis dan preferensi perjalanan mempengaruhi keputusan wisatawan. Namun, meskipun bermanfaat, model ini memiliki keterbatasan dalam menghasilkan keputusan secara cepat dan berkelanjutan, yang dapat menjadi hambatan dalam konteks dinamis industri pariwisata. (Tang dkk., 2020).

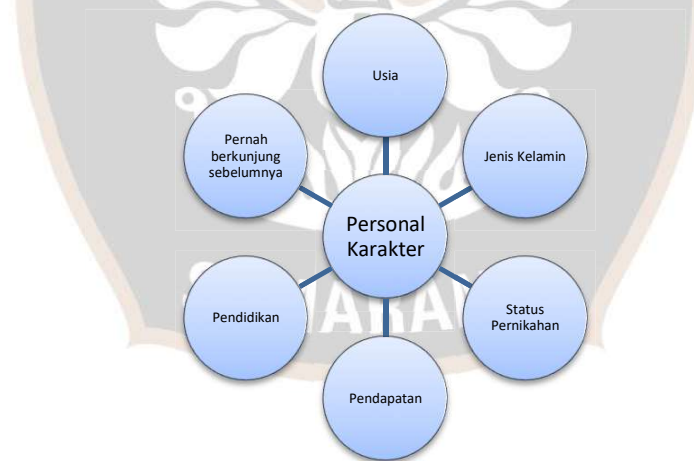
Selain faktor-faktor yang disebutkan sebelumnya, penelitian oleh Pestana dkk. dalam "*Motivations, emotions and satisfaction: The keys to a tourism destination choice*" menambahkan atribut seperti status pernikahan dan tingkat pendidikan

sebagai variabel penting dalam pemilihan destinasi wisata. Penelitian ini menunjukkan hasil yang dapat diterapkan di Eropa, tetapi belum terbukti efektif untuk diterapkan di wilayah luar Eropa (Pestana dkk., 2020).

Kemunculan aplikasi TikTok juga memberikan pengaruh terhadap pemilihan destinasi wisata. Penelitian tersebut menggunakan model UTAUT untuk mengetahui faktor-faktor yang mempengaruhi pengguna tiktok dengan tujuan untuk menentukan tujuan wisata (Zhou dkk., 2023).

Untuk melakukan analisis sentiment, Sinha dkk (2018) menggunakan *multi label naïve bayes classifier*. Hasil dari penelitian tersebut pelaku usaha pariwisata dapat mengetahui retensi dan tingkat kepuasan pelanggan. Dari hasil uji model didapatkan bahwa isu sosial menjadi pertimbangan utama bagi para wisatawan dalam proses menentukan destinasi.

Seddighi dan Theocharous mengemukakan model pemilihan destinasi dengan menggunakan 3 kelompok variabel dijelaskan pada Gambar 2.1.

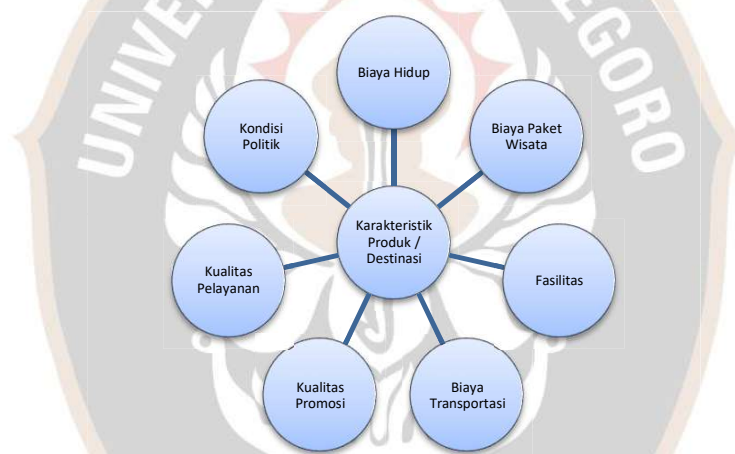


Gambar 2. 1. Variabel yang termasuk kelompok Personal Karakter

Gambar 2.1. menunjukkan bahwa kelompok pertama adalah personal karakter, didalamnya terdapat data terkait dengan pengunjung, seperti usia, jenis kelamin, status pernikahan, pendapatan, pendidikan, dan apakah wisatawan pernah berkunjung ke tempat tersebut sebelumnya. Data ini berfungsi untuk menggambarkan profil wisatawan dan membantu dalam memahami preferensi serta perilaku perjalanan mereka. Variabel-variabel tersebut penting untuk dianalisis

karena dapat mempengaruhi keputusan wisatawan dalam memilih destinasi yang sesuai dengan kebutuhan mereka.

Variabel berikutnya termasuk dalam kelompok karakteristik produk atau destinasi, yang berkaitan dengan berbagai aspek yang melekat pada produk wisata itu sendiri. Variabel-variabel ini mencakup elemen seperti biaya hidup, biaya paket wisata, fasilitas yang tersedia, biaya transportasi, kualitas promosi, kualitas pelayanan, dan kondisi politik di destinasi tersebut. Semua variabel ini berperan penting dalam membentuk persepsi wisatawan terhadap destinasi, karena langsung mempengaruhi kenyamanan, kepuasan, dan keputusan mereka untuk mengunjungi atau merekomendasikan destinasi tersebut kepada orang lain.



Gambar 2. 2. Variabel kelompok karakteristik produk

Pada gambar 2.2. menunjukkan 7 variabel yang dikelompokkan dalam karakteristik produk, yaitu, biaya hidup, biaya paket wisata, fasilitas, biaya transportasi, kualitas promosi, kualitas pelayanan, kondisi politik. Wisatawan mengemukakan persepsinya, penting tidaknya variabel tersebut bagi mereka.

Kelompok ini terdiri dari satu pertanyaan yang mewakili penilaian keseluruhan, yaitu apakah wisatawan akan mempertimbangkan untuk mengunjungi kembali destinasi wisata tersebut. Model yang diusulkan merupakan kombinasi dari pendekatan karakteristik Lancasterian dan model pemilihan transportasi Koppleman. Hasil analisis model ini ditampilkan pada Tabel 2.1, yang menggambarkan profil individu dengan kemungkinan tinggi untuk kembali berkunjung. Penggabungan kedua pendekatan ini memungkinkan model untuk

lebih akurat dalam memprediksi perilaku wisatawan, dengan mempertimbangkan berbagai aspek personal dan karakteristik destinasi. Model ini diharapkan dapat menjadi alat yang efektif bagi pengelola destinasi dalam meningkatkan tingkat kunjungan ulang wisatawan (Seddighi & Theocharous, 2002) .

Tabel 2. 1. Profil individu yang memiliki kemungkinan tinggi untuk revisit

Peluang Tertinggi (Lebih dari 90%)	
Personal Karakter	
Umur	Pertengahan
Jenis Kelamin	Perempuan
Status Pernikahan	Belum Menikah
Pendapatan	Rata-rata
Pendidikan	SMA dan Kuliah
Pernah Berkunjung Sebelumnya	Ya
Biaya Hidup	Penting
Fasilitas	Penting
Biaya Transportasi	Penting
Kualitas Layanan	Penting
Kestabilan Politik	Penting

Pengunjung yang mempunyai peluang tinggi untuk berkunjung kembali adalah responden dengan usia pertengahan, jenis kelamin perempuan, belum menikah, mempunyai pendapatan rata-rata, pendidikan lanjut. Wisatawan pernah mengunjungi objek wisata tersebut sebelumnya, menganggap bahwa biaya hidup, harga paket wisata, fasilitas biaya transportasi, kualitas promosi, pelayanan dan kondisi politik merupakan hal yang penting untuk dipertimbangkan.

Tang (2020) dalam penelitiannya menggunakan variabel-variabel yang ditunjukkan pada Gambar 2.3. Variabel-variabel tersebut mencakup faktor-faktor demografis seperti usia, jenis kelamin, dan pendapatan, serta karakteristik perjalanan seperti status kepemilikan mobil, status tempat tinggal, destinasi pilihan, lama liburan, dan rekan seperjalanan. Selain itu, variabel lain yang juga dipertimbangkan adalah apakah wisatawan ditemani orang tua, anak-anak, dan moda transportasi yang digunakan. Penelitian ini memberikan wawasan penting tentang bagaimana faktor personal dan preferensi perjalanan berperan dalam menentukan pilihan destinasi wisata yang dilakukan oleh wisatawan, sehingga

dapat membantu pengelola destinasi memahami lebih dalam perilaku dan kebutuhan wisatawan.



Gambar 2. 3. Variabel yang digunakan Tang dkk

Akan tetapi karena dalam Gambar 2.3 pemilihan destinasi wisata dihubungkan dengan pemilihan moda transportasi maka variable yang ada tidak akan digunakan keseluruhan. Adapun metode yang digunakan dalam penelitian sebelumnya adalah multinominal logit(Tang dkk., 2020).

Tankovic dan rekan-rekannya meneliti pengaruh media sosial dalam menentukan destinasi wisata dengan menggunakan pendekatan PLS-SEM (Partial Least Squares Structural Equation Modeling) untuk menganalisis data yang diperoleh. Hasil penelitian menunjukkan bahwa media sosial berpengaruh signifikan terhadap keputusan individu dalam memilih destinasi wisata. Interaksi di media sosial, ulasan dari pengguna lain, dan konten promosi terbukti memainkan peran penting dalam membentuk preferensi wisatawan terhadap suatu destinasi.

Variabel yang digunakan dalam penelitian ini mencakup intensitas penggunaan media sosial, interaksi dengan konten destinasi, serta persepsi terhadap ulasan dan rekomendasi. Selain itu, faktor demografi seperti usia dan gaya hidup juga berfungsi sebagai moderator dalam hubungan antara paparan media sosial dan keputusan wisata. Temuan ini menyoroti pentingnya media sosial sebagai faktor utama yang mempengaruhi perilaku dan pilihan destinasi wisatawan, memberikan panduan bagi pelaku industri pariwisata untuk lebih memanfaatkan platform digital

dalam strategi pemasaran mereka.



Gambar 2. 4. Variabel yang digunakan dalam penelitian Tankovic dkk

Gambar 2.4 menjelaskan tentang pengaruh jejaring sosial terhadap keputusan untuk menentukan destinasi wisata. Adapun indikatornya adalah Penggunaan jejaring sosial, Komunikasi melalui jejaring sosial dan Informasi pada jejaring sosial. Dengan hasil bahwa ke 3 indikator tersebut mempunyai korelasi positif terhadap pemilihan destinasi wisata (Tanković dkk., 2022).

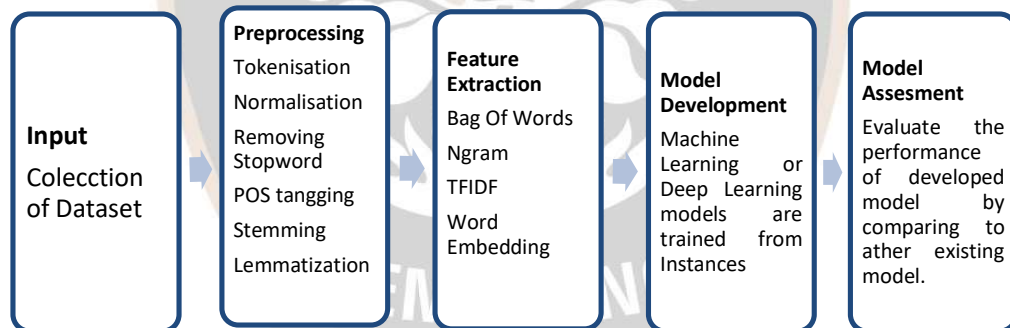
Dalam penelitian Alaei, menggunakan Teknik klasifikasi dan klusterisasi dalam mengolah data teks yang diperoleh. Hasilnya *analysis sentiment* dengan menggunakan Teknik klasifikasi memberikan hasil yang lebih baik dari pada Teknik klusterisasi. Dalam penelitian tersebut baru digunakan metode analisis sentimen dalam bidang pariwisata tanpa menggunakan atribut yang lain(Alaei dkk., 2019). Kemudian pada tahun 2019 Vu dkk juga melakukan prediksi rute perjalanan wisata di Australia dengan Teknik Sequential Rule mining. Mereka menggunakan data set berupa kumpulan foto dari pelaku pariwisata kemudian menggunakan teknik SRM untuk mengekstrak tujuan dari urutan perjalanan (Vu dkk., 2018).

2.1.3 Analisis Sentimen Dalam Bidang Pariwisata

Penerapan analisis sentimen dalam bidang pariwisata sudah banyak dilakukan karena merupakan komponen kunci penggunaan big data dalam penelitian bidang pariwisata. Karena analisis sentimen mampu menemukan emosi yang diungkapkan wisatawan berdasarkan pengalaman masing-masing(Irawan dkk., 2019). Sentimen

analisis juga mampu mendeteksi opini negative dan positif dalam jumlah yang besar (Sigala dkk., 2019). Moud dkk memaparkan penggunaan sentimen analisis untuk membuat sistem rekomendasi. Dalam penelitian tersebut, sistem analisis dikombinasikan dengan *semantic clustering*. Hasilnya sistem yang diusulkan tersebut lebih unggul dibandingkan dengan *F- Measure* (Abbasi-Moud dkk., 2021). Oleh karena itu teknik analisis sentimen sangat tepat untuk mendukung Pariwisata berkelanjutan. Karena selain berguna dari sisi wisatawan juga berguna bagi para pelaku usaha, yaitu untuk melakukan pemasaran (Irawan dkk., 2019).

Sentimen analisis dikenal juga sebagai penambangan opini, dengan analisis sentimen ini dapat diketahui persepsi penulis positif, netral atau negatif (Onyenwe dkk., 2020). Dalam penelitian Nandwani 2021 menunjukkan bahwa penerapan Lexicon Based mampu memberikan hasil yang baik dalam menganalisis sentimen penulis opini (Nandwani & Verma, 2021). Langkah dasar Analisis sentimen dapat dilihat pada gambar 2.5.



Gambar 2. 5. Tahapan Analisis Sentimen

Gambar 2.5 menunjukkan langkah-langkah analisis sentimen yang dimulai dari *Input Dataset* hingga mendapatkan hasil akhir yang merepresentasikan sentimen data yang dianalisis. Berikut adalah penjelasan setiap langkah dalam proses analisis sentimen:

1. *Input Dataset*

Dataset yang diambil dari media sosial ataupun web dimana data tersebut merupakan opini dari pelanggan.

2. *Preprocessing*

Merupakan proses pembersihan data, dengan menghilangkan data yang sekiranya mengurangi akurasi hasil dengan menggunakan metode tertentu.

3. *Feature Extraction*

Dalam tahap ini dokumen dipecah menjadi kalimat kemudian kalimat dipecah lagi menjadi kata-kata supaya dapat diproses lebih lanjut.

4. *Model Development*

Pada tahap pengembangan model, digunakan proses klasifikasi menggunakan algoritme naïve bayes dan SMOTE Over sampling.

5. *Model Assesment*

Model Assesment adalah tahapan untuk mengukur kinerja dari model yang digunakan.

Menurut Rintyarna dkk dalam penelitiannya mengemukakan bahwa Sentimen Analisis dapat dimanfaatkan untuk :

1. Menggali persepsi pelanggan terkait kualitas layanan.
2. Menyediakan *Framework* untuk pemodelan penilaian kualitas layanan.
3. Mengevaluasi kinerja/ tingkat akurasi analisis sentimen menggunakan model yang diajukan (Rintyarna dkk., 2022).

Dalam melakukan analisis sentiment, peneliti dapat menggunakan Algoritme – Algoritme sebagai berikut :

1. Algoritme Naïve Bayes dan *Support Vector Machine*

Watori dkk melakukan penelitian analisis sentimen dengan membandingkan Algoritme naïve bayes dan SVM untuk mengetahui persepsi masyarakat tentang pemindahan ibu kota negara. Hasilnya adalah Algoritme Naïve Bayes menghasilkan akurasi 78,39% sedangkan SVM menghasilkan akurasi 76,40%(Watori dkk., 2020). Algoritme yang sama juga terbukti mampu menganalisis persepsi pengguna mobile banking id Malaysia(Rahman dkk., 2022).

2. *Neural Network*

Algoritme Neural Network dapat digunakan untuk analysis sentiment dalam penelitian Apriliani dkk untuk meneliti pelayanan hotel di Indonesia, dengan tingkat akurasi 88,99%(Apriliani dkk., 2020).

3. *Logistic Regression, Decission Tree, Maximum Entropi, K-Nearest*

Neighbours (KNN)

Wankhade dalam surveinya mengemukakan bahwa *Logistic Regression, Decision Tree, Maximum Entropi, K-Nearest Neighbours (KNN)* dapat digunakan untuk analisis sentiment(Wankhade dkk., 2022).

Untuk mengukur akurasi Algoritme digunakan rumus sebagai berikut :

$$Precision = \frac{TP}{TP + FP} \quad (2.1)$$

Dimana

Precision = Rasio Ketepatan Prediksi

TP = Prediksi Benar

FP = Prediksi Salah

Dalam dunia pariwisata analisis sentimen merupakan komponen kunci dalam teknologi *Big Data*. Pendekatan Analisis yang sering digunakan yaitu *Lexicon Based* dan *Machine Learning*. Penggunaannya paling besar untuk menganalisis atraksi wisata, hotel dan restaurant. Tapi pada perkembangan berikutnya analisis sentiment dapat juga digunakan untuk mengumpulkan opini terkait layanan, seperti layanan pemandu tour (Sigala dkk., 2019).

2.2 Keaslian Penelitian

Penelitian ini berfokus pada pengembangan model pemilihan destinasi wisata yang lebih personal dan relevan dengan menggunakan analisis sentimen sebagai komponen utama, yang belum banyak diterapkan dalam studi-studi sebelumnya. Kajian literatur mengungkapkan bahwa penelitian terdahulu umumnya menggunakan pendekatan berbasis variabel demografis dan preferensi perjalanan, seperti yang dilakukan oleh Tang dkk. (2020), yang mengandalkan model multinomial logit untuk menganalisis faktor-faktor seperti usia, jenis kelamin, dan moda transportasi dalam pemilihan destinasi wisata. Selain itu, penelitian oleh Tankovic dkk. (2022) menunjukkan pengaruh media sosial terhadap keputusan wisatawan, namun belum sepenuhnya mengintegrasikan ulasan pengguna sebagai input yang dapat diolah menjadi rekomendasi destinasi. Keterbatasan-keterbatasan ini menjadi motivasi bagi penelitian saat ini untuk mengadopsi pendekatan yang lebih adaptif dan berbasis data persepsi dari wisatawan, sehingga memberikan rekomendasi yang lebih akurat dan sesuai dengan kebutuhan wisatawan.

Keunikan penelitian ini terletak pada penggunaan analisis sentimen yang memproses ulasan wisatawan untuk memahami persepsi positif, netral, atau negatif terhadap destinasi tertentu, yang kemudian digunakan sebagai input utama dalam model rekomendasi. Pendekatan ini berbeda dari penelitian sebelumnya yang sering kali tidak mempertimbangkan aspek emosional dan persepsi wisatawan yang terekam melalui media digital. Dengan mengintegrasikan algoritme Naïve Bayes untuk mengklasifikasikan ulasan berdasarkan sentimen, penelitian ini tidak hanya menawarkan model yang lebih canggih, tetapi juga mampu menangkap dinamika pengalaman wisatawan secara lebih mendalam. Tabel 2.2 merangkum persamaan dan perbedaan utama antara penelitian ini dengan studi-studi terdahulu, serta menjelaskan kontribusi unik yang diharapkan dapat memperkaya literatur mengenai pemilihan destinasi wisata.

Tabel 2. 2. Penelitian Terdahulu tentang Pemilihan Destinasi Wisata

Judul	Peneliti	Tujuan Penelitian	Metode Penelitian	Hasil Penelitian	Persamaan	Perbedaan
A Model of Tourism destination choice: a theoretical and empirical analysis	H.R. Seddighi, A.L. Theocharous (2001)	Untuk mengetahui karakteristik turis dengan peluang terbesar untuk kembali mengunjungi Cyprus	Menggunakan model Koppelman dan Lanca7ster.	Karakteristik turis yang memutuskan untuk kembali mengunjungi Cyprus.	Variabel	Penelitian ini menggunakan machine learning, dengan menambahkan variabel lain. Tujuan penelitian adalah untuk memberikan rekomendasi tujuan wisata.
Social Networks Influence in Choosing a tourist Destination	Ana Cuic Tankovic, Ivana Bilic, Antonia Sohor.(2022)	Mengetahui pengaruh sosial media dalam menentukan tujuan wisata.	Data dikumpulkan menggunakan Google Form dan diolah dengan PLS-SEM, menggunakan variabel Penggunaan media sosial, Komunikasi dan Informasi di media sosial.	Penelitian tersebut menunjukkan bahwa sosial media memberikan pengaruh yang signifikan terhadap keputusan memilih destinasi wisata.	Variabel yang digunakan.	Mengambil variabel dari penelitian tankovic dkk ini untuk di modelkan Bersama variabel lain.

Tabel 2..2.Penelitian Terdahulu tentang Pemilihan Destinasi Wisata (Lanjutan)

Judul	Peneliti	Tujuan Penelitian	Metode Penelitian	Hasil Penelitian	Persamaan	Perbedaan
Choice Behaviour of Tourism Destination and Travel Mode : A Case Study of Local Resident in Hangzhou, China	Tang dkk (2020)	Penelitian bertujuan untuk mempelajari tujuan wisata dan perilaku pemilihan moda dan memberikan saran untuk memperbaiki sistem transportasi pariwisata	Menggunakan multinomial logit model.	.Variabel umur, tipe tempat tinggal, kepemilikan mobil, teman seperjalanan, dan lama liburan memberikan pengaruh yang signifikan terhadap pemilihan destinasi.	Variabel.	Penelitian Tang dkk, fokusnya ke moda transportasi ke destinasi tujuan wisata. Sedangkan dalam penelitian ini adalah untuk memilih destinasi wisata.
A Historical Tourism Recommendation System For The Elderly Tourist Using Natural Language Processing And The Ontology Technique.	Anongporn Salaiwarakul (2021)	Penelitian Salaiwarakul bertujuan untuk memberikan rekomendasi destinasi wisata sejarah kepada turis lansia.	Menggunakan Natural Language Processing dan Ontologi	Dapat memberikan saran yang sesuai untuk turis lansia.	Penggunaan analisis sentimen untuk rekomendasi tujuan	Penelitian Salaiwarakul hanya menggunakan NLP dan Ontologi untuk menghasilkan rekomendasi destinasi, sedangkan penelitian yang akan dilakukan menggunakan hasil analisis sentiment sebagai variabel dalam proses pemilihan destinasi wisata.
Selection of Tourism Destinations Priority using 6AsTD Framework and TOPSIS	Arif dkk (2019)	Memberikan rekomendasi destinasi wisata berdasarkan framework 6A	Menggunakan TOPSIS	Mengurutkan prioritas destinasi wisata di area Malang.	Memberikan rekomendasi Destinasi Wisata. Menggunakan framework 6A.	Dalam penelitian yang akan dilakukan, framework 6A digunakan untuk mengelompokkan destinasi wisata.

2.3 Landasan Teori

2.3.1 Ontologi

Ontologi merupakan prosedur yang digunakan untuk mendeteksi apakah review dari responden merupakan pendapat yang sebenarnya atau pendapat palsu (Vidanagama dkk., 2022). Albar dkk dalam penelitiannya melakukan pengelompokan kata-kata yang biasa digunakan untuk mengulas *smartphone*, kerangka ontology dibuat menggunakan software ontogen. Kemudian berhasil menemukan beberapa review yang tidak berhubungan dengan *smartphone* yang dimaksud (Akbar dkk., 2015). Penggunaan ontologi juga dilakukan oleh Salaiwarakul untuk memberikan rekomendasi destinasi wisata sejarah kepada turis lansia (Salaiwarakul, 2022).

2.3.2 Algoritme Naïve Bayes

Menurut Subashini dkk, Algoritme Naïve Bayes merupakan Algoritme yang paling sering digunakan dalam analisis sentimen (Subhashini dkk., 2021). Prinsip kerjanya adalah dengan mengelompokkan data kedalam kelas menggunakan probabilitas.

Keuntungan menggunakan Algoritme Naïve Bayes:

1. Algoritme yang sederhana untuk dipahami dan digunakan.
2. Lebih cepat untuk meprediksi kelas.
3. Dapat digunakan untuk dataset dengan jumlah kecil.

Kekurangan Algoritme Naïve Bayes:

1. Adanya *Zero Conditional Probability Problem* yang dapat menghapus semua probabilitas yang muncul.
2. Asumsi yang sangat kuat tentang fitur kelas independensi dapat mengurangi akurasi (Kaviani & Dhotre, 2017).

Rumusnya dapat dituliskan sebagai berikut (Raschka, 2014)

$$P(X|H) = \frac{P(X|H) * P(H)}{P(X)} \quad (2.2)$$

Keterangan

X = Bukti

H = Hipotesis

P (H|X) = Peluang Hipotesis H Benar untuk Bukti X

P(X|H) = Peluang Bukti X Benar untuk Hipotesis H

$P(H)$ = Probabilitas prior hipotesis H

$P(X)$ = Probabilitas prior bukti X

Dalam penelitian yang dilakukan oleh Subashini dkk. (2021), data diolah menggunakan algoritma Naïve Bayes untuk analisis sentimen. Penelitian ini menunjukkan bahwa Naïve Bayes memberikan akurasi tertinggi dibandingkan algoritma lainnya, seperti Support Vector Machine (SVM) dan Logistic Regression, dengan nilai akurasi mencapai 78,39%. Keunggulan ini dicapai karena pendekatan probabilistik Naïve Bayes yang sederhana namun efisien, terutama dalam mengolah data teks berbasis ulasan.

Naive Bayes adalah algoritma sederhana namun efisien, terutama dalam menangani data kategorikal seperti jenis destinasi wisata atau ulasan pengguna. Dalam sistem rekomendasi pariwisata, algoritma ini digunakan untuk mengklasifikasikan preferensi wisatawan berdasarkan dataset historis. Penelitian oleh Rizqi dkk (2024) menunjukkan bahwa penerapan Naive Bayes dalam sistem pendukung keputusan untuk rekomendasi wisata di Lamongan mencapai akurasi hingga 94%, menegaskan keandalan algoritma ini dalam memberikan rekomendasi yang sesuai dengan preferensi wisatawan. Keunggulan Naive Bayes terletak pada kecepatan dan efisiensinya, menjadikannya pilihan tepat untuk sistem rekomendasi yang memerlukan prediksi dalam waktu nyata (Rizqi dkk, 2024).

2.3.3 Multinomial Naïve Bayes

Multinomial Naïve Bayes adalah merupakan salah satu varian Naïve Bayes yang digunakan untuk mengelompokkan teks maupun dokumen. Prinsip kerjanya dengan menghitung kemungkinan / Probabilitas suatu kata masuk kedalam kelompok tertentu (D. Manning dkk., 2021).

$$P(c|d) = P(c) \prod_{k \leq n_d} P(t_k|c) \quad (2.3)$$

Keterangan

$P(c|d)$: Probabilitas dokumen d berada dikelas c

$P(c)$: Prior probability dokumen masuk kedalam kelas c

$P(t_k|c)$: Berapa banyak t_k berada dalam kelas c

$\{t_1, t_2, t_3, \dots, t_n\}$: Token dalam dokumen c

Keunggulan *Multinomial Naïve Bayes*

1. Mudah digunakan pada data kontinu dan diskrit
2. Itu dapat menangani kumpulan data besar
3. Itu dapat mengklasifikasikan data dengan beberapa label
4. Paling bagus digunakan untuk model *Natural Language Processing*(Shriram, n.d.)

2.3.4 Algoritme KNN (*K-Nearest Neighbors*)

K-Nearest Neighbor (KNN) adalah salah satu Algoritme pembelajaran mesin yang paling sederhana dan mudah dipahami. Algoritme ini digunakan untuk tugas-tugas klasifikasi dan regresi, yang bekerja berdasarkan prinsip kedekatan atau kemiripan antar data. KNN mengklasifikasikan objek berdasarkan mayoritas dari K tetangga terdekat, di mana kelas yang paling umum di antara K tetangga terdekat dipilih sebagai kelas dari objek tersebut(Sikri dkk., 2024).

Dalam penelitian yang dilakukan oleh Dhasaradhan Kaveripakam dan Jaichandran Ravichandran, KNN digunakan untuk mengidentifikasi penyakit diabetes dengan menggunakan dataset PIMA Indian Diabetic Dataset (PIMAIDD). Hasil dari penelitian menunjukkan bahwa KNN memiliki performa yang cukup baik dengan nilai akurasi yang kompetitif dibandingkan dengan Algoritme lain seperti *Logistic Regression* (LGR), *Support Vector Machine* (SVM), dan *Random Forest* (RF). Pengujian dilakukan dengan menggunakan berbagai teknik evaluasi seperti matriks kebingungan, akurasi (ACR), presisi (PCN), recall (RCL), f1-score (FSC), receiver operating characteristic (ROC), dan K-fold cross-validation(Kaveripakam & Ravichandran, 2025).

K-Nearest Neighbors (KNN) adalah algoritma yang efektif dalam mengelompokkan data berdasarkan kesamaan fitur, menjadikannya sangat relevan untuk sistem rekomendasi pariwisata. Algoritma ini bekerja dengan prinsip pencocokan pola, di mana rekomendasi destinasi diberikan berdasarkan kedekatan preferensi pengguna dengan data wisatawan sebelumnya. Penelitian oleh Anamisa dkk (2023) menunjukkan bahwa penerapan KNN dalam sistem rekomendasi objek wisata di Bangkalan mencapai akurasi hingga 78% pada nilai $k=1$, yang menunjukkan kemampuan KNN dalam memberikan rekomendasi sesuai dengan

kriteria wisatawan. Kemampuan algoritma ini untuk menangani data berbasis preferensi dan interaksi wisatawan membuatnya menjadi pilihan populer dalam model rekomendasi (Anamisa et al., 2023).

2.3.5 **Algoritme Decision Tree**

Decision Tree (DT) adalah salah satu teknik pembelajaran mesin yang termasuk dalam kategori supervised learning. Algoritme ini digunakan baik untuk tugas klasifikasi maupun regresi. *Decision Tree* memetakan fitur-fitur suatu data ke dalam struktur pohon di mana setiap cabang merepresentasikan keputusan yang diambil berdasarkan fitur tertentu, dan setiap daun merepresentasikan hasil atau kategori akhir dari data tersebut.

Secara visual, *Decision Tree* merepresentasikan data dalam bentuk diagram pohon, di mana setiap elemen memiliki peran spesifik. Node merepresentasikan fitur dalam dataset, berfungsi sebagai titik pengambilan keputusan. Cabang (branch) menggambarkan aturan atau keputusan yang diambil berdasarkan fitur tersebut. Sementara itu, daun (leaf) merepresentasikan hasil akhir atau klasifikasi yang dihasilkan dari proses keputusan. Struktur ini memudahkan interpretasi data dengan menggambarkan alur logika yang jelas dari fitur hingga hasil akhir (Sikri dkk., 2024).

Dalam penelitian yang berjudul "*Fuzzy Evaluation and Benchmarking Framework for Robust Machine Learning Model in RealTime Autism Triage Applications*," *Decision Tree* digunakan bersama Algoritme lain untuk membangun model prediksi dalam aplikasi triase autisme. *Decision Tree* dipilih karena kemampuannya dalam mengklasifikasikan data berdasarkan aturan-aturan yang jelas dan mudah diinterpretasikan. Algoritme ini, bersama dengan Algoritme lain seperti SVM, *AdaBoost*, dan *Random Forest*, diterapkan untuk mengevaluasi performa model dalam kondisi normal dan serangan adversarial, menunjukkan bahwa *Decision Tree* dapat menjadi pilihan yang baik (Shayea dkk., 2024).

Decision Tree adalah algoritma yang dikenal karena interpretabilitasnya, memungkinkan pengembang sistem untuk memahami faktor-faktor yang memengaruhi rekomendasi dengan lebih mudah. Dalam konteks pariwisata, algoritma ini sangat berguna untuk menangani data kompleks seperti demografi

wisatawan, ulasan destinasi, dan preferensi aktivitas. Penelitian oleh Larasati dan Sutrisno (2018) menunjukkan bahwa Decision Tree berhasil digunakan dalam sistem rekomendasi situs wisata di Jakarta berdasarkan ulasan dari platform web, menghasilkan model yang andal dan akurat. Kelebihan utama algoritma ini adalah kemampuannya untuk menangkap hubungan hierarkis antara fitur-fitur dalam dataset, yang penting dalam rekomendasi destinasi (Ajeng et al., 2018).

2.3.6 Model Pemilihan Destinasi Pariwisata

Faktor – faktor yang melatarbelakangi wisatawan dalam menentukan destinasi sangat penting untuk dianalisis bagi para pelaku usaha wisata. Terkait dengan penyusunan strategi pemasaran wisata maupun pengembangan wisata (Yusendra, 2015). Dalam penelitian Yusendra tersebut motivasi personal memiliki rangking tertinggi dibandingkan dengan faktor yang lain seperti kebersihan, cuaca dan lain-lain. Selain untuk pelaku usaha wisata, pelaku wisata (wisatawan) juga memerlukan Sistem Pendukung Keputusan yang baik agar mendapatkan hasil sesuai dengan kriteria yang diinginkan (Aidah & Rasmita, 2018).

Dalam penelitian Tang dkk, menggunakan Teknik analisis Multinomial Logit, untuk mengetahui faktor wisatawan dalam menentukan destinasi/tujuan. Adapun variable yang digunakan adalah, sarana transportasi, Usia, Jenis Kelamin, Kepemilikan Tempat Tinggal, Kepemilikan Mobil, Rekan, Lama Liburan, Destinasi (Tang dkk., 2020).

Menurut Horner dan Swarbrooke, motivasi turis sangat dipengaruhi oleh pengalaman masa lalu mereka sebagai turis, baik yang bersifat positif maupun negatif. Pengalaman positif, seperti kunjungan yang menyenangkan dan layanan yang memuaskan, dapat meningkatkan keinginan turis untuk kembali mengunjungi destinasi yang sama atau menjelajahi destinasi baru dengan harapan mendapatkan pengalaman serupa. Sebaliknya, pengalaman negatif dapat menyebabkan turis enggan untuk kembali atau bahkan memberikan ulasan buruk yang mempengaruhi reputasi destinasi tersebut. Oleh karena itu, penting bagi pengelola destinasi untuk terus meningkatkan kualitas layanan dan fasilitas agar dapat memberikan pengalaman yang memuaskan dan berkesan bagi para pengunjung. Dengan cara ini, pengalaman masa lalu turis dapat dimanfaatkan sebagai alat yang efektif untuk

meningkatkan loyalitas dan menarik lebih banyak pengunjung di masa depan. Memahami motivasi yang dipengaruhi oleh pengalaman ini membantu industri pariwisata dalam merancang strategi pemasaran yang lebih tepat sasaran dan berkelanjutan (Horner & Swarbrooke, 2007).

2.3.7 Komponen Pengembangan Pariwisata

Dalam penelitian, Arief dkk (2019) menggunakan pendekatan 6A untuk melakukan rekomendasi prioritas tujuan pariwisata menggunakan TOPSIS. Pendekatan 6A merupakan kunci dari kesuksesan sebuah destinasi wisata yaitu :

1. Attractions /Atraksi

Atraksi pada sebuah destinasi wisata dapat berupa pemandangan alam, wisata buatan seperti taman hiburan, wisata budaya seperti pertunjukan musik (Buhalis & Amaranggana, 2013), juga event-event tertentu. Setiap destinasi memiliki ciri khas yang berbeda dengan daerah lain (Chaerunissa & Yuniningsih, 2020).

2. Accessibility / Akses

Akses berkaitan erat dengan tersedianya transportasi menuju dan keluar dari sebuah destinasi wisata. Kondisi jalan, dan jarak yang harus ditempuh untuk mengunjungi destinasi wisata masuk kedalam komponen ini.

3. Amenities

Termasuk dalam komponen amenitis adalah ketersediaan penginapan, tempat makan juga biro perjalanan.

4. Activities

Aktivitas merupakan kegiatan yang dapat diikuti oleh wisatawan di sebuah destinasi wisata, seperti ikut dalam kehidupan sehari-hari warga lokal. Menurut Brown dan Stange, dalam Chaerunnisa dan Yuniningsih mengemukakan bahwa setiap daerah memiliki kegiatan yang beda dengan daerah lain (Chaerunissa & Yuniningsih, 2020).

5. Ancillary Service

Sarana pendukung yang dimaksudkan adalah sarana seperti adanya ATM atau bisa juga pembayaran Nontunai, juga jaringan internet dan pusat informasi (Rosanto & Chainarta, 2021).

6. Available Packages

Paket wisata dari pengelola lokal dapat berupa layanan dalam satu paket pariwisata untuk ditawarkan kepada wisatawan (Tran dkk., 2017). Paket wisata juga dapat berupa gabungan dari berbagai layanan yang mencakup akomodasi, transportasi, makanan, dan aktivitas wisata di berbagai destinasi, yang dirancang untuk memberikan pengalaman perjalanan yang lengkap dan menyeluruh. Paket ini sering kali menggabungkan beberapa tujuan wisata yang berdekatan atau memiliki tema serupa, sehingga wisatawan dapat menikmati berbagai pengalaman dalam satu perjalanan. Selain itu, paket wisata ini biasanya ditawarkan dengan biaya yang lebih terjangkau dibandingkan jika layanan dan destinasi tersebut dipesan secara terpisah, memberikan nilai tambah dan kemudahan bagi wisatawan dalam merencanakan perjalanan mereka. (Cakici & Harman, 2007).

2.3.8 Algoritme Case Base Reasoning - Simple Matching Coefficient

a) Case Base Reasoning

Menurut Mulyana dalam Swari dkk (2020) CBR merupakan sebuah metode untuk memecahkan masalah berdasarkan kasus yang terjadi sebelumnya. (Swari & Firdaus, 2020). Aardmodt dan Plaza (1994) mengemukakan Case Based Reasoning menjadi 4 proses dan digambarkan secara melingkar (Agnar & Plaza, 1994). Proses-proses itu adalah

1. Retrieve

Retrieve adalah menemukan kasus-kasus yang mirip kemudian dibandingkan dengan kasus yang terjadi sebelumnya. Tahapan pada retrieve ini adalah, mengidentifikasi masalah, mencocokkan masalah, kemudian menyeleksi.

2. Reuse

Adalah mengambil kasus lama untuk men dijabarkanyeleaikan kasus baru. Untuk melakukan reuse pada kasus lama ada 2 cara yaitu,

- Menggunakan solusi dari kasus lama
- Menggunakan metode dari kasus lama untuk diterapkan ke kasus baru.

3. Revise

Melakukan perubahan terhadap pemecahan masalah yang sudah pernah dilakukan sebelumnya. Langkahnya adalah dengan mengevaluasi Solusi, dan memperbaiki kesalahan atau kekurangan dari pemecahan masalah sebelumnya.

4. *Retain*

Dalam *retain* ini terdapat 3 tahapan extract, index dan integrate. Pada intinya menggunakan pemecahan masalah kasus sebelumnya, dan bisa juga digabungkan dengan metode yang baru.

b) *Simple Matching Coefficient*

Dalam penelitian Verma dan Aggarwal (2020), mengemukakan bahwa Han dkk mendeskripsikan SMC sebagai pengukuran statistik dengan menghitung kesamaan pada data biner (Verma & Aggarwal, 2019).

$$SMC = \frac{\text{Jumlah Atribut yang sama}}{\text{Total Atribut}} \quad (2.4)$$

Dijelaskan sebagai berikut

$$SMC = \frac{N_{00} + N_{01}}{N_{00} + N_{01} + N_{11} + N_{10}} \quad (2.5)$$

Keterangan

SMC = Nilai kesamaan

N_{00} = Jumlah Nilai lama 0 dan nilai baru 0

N_{11} = Jumlah Nilai lama 1 dan nilai baru 1

N_{01} = Jumlah Nilai Lama 0 dan Nilai baru 1

N_{10} = Jumlah Nilai Lama 1 dan Nilai baru 0

Sari dkk (2019) melakukan penelitian menggunakan SMC untuk melakukan pemilihan resep masakan. Penelitian menggunakan 20 data dan berhasil mendapatkan angka 89% untuk similaritas tertinggi sehingga dapat disimpulkan bahwa Algoritme *Simple Matching Coefficient* berhasil memberikan rekomendasi yang sesuai (Sari dkk., 2019).

2.3.9 Aplikasi Rekomendasi Pariwisata Sebagai *Decision Support System* bagi wisatawan

Decision Support System merupakan sistem informasi yang intreraktif, menyediakan informasi, pemodelan data dan alat analitis untuk membantu pengambilan keputusan (Ali dkk., 2023). Efektivitas DSS terletak pada kemampuannya untuk menyediakan informasi yang akurat, tepat waktu, dan relevan, sehingga mengurangi ketidakpastian dan meningkatkan kualitas keputusan. DSS memungkinkan organisasi untuk merespons perubahan dengan lebih cepat, mengoptimalkan alokasi sumber daya, dan meningkatkan produktivitas

secara keseluruhan. Dengan memanfaatkan DSS, bisnis dapat mencapai efisiensi operasional yang lebih baik dan perencanaan strategis yang lebih baik, yang pada akhirnya meningkatkan hasil dan keunggulan kompetitif. Hal ini terbukti dalam penelitian Dhanaraju yang membuat Sistem Pendukung Keputusan/DSS dalam bidang agrikultur(Dhanaraju dkk., 2022).

Dalam membangun aplikasi berbasis web yang menggunakan algoritme *machine learning*, bahasa pemrograman Python merupakan pilihan yang sangat tepat. Python menyediakan berbagai *library* dan *framework* yang mendukung pengembangan machine learning secara efisien(Raschka dkk., 2020). Untuk membangun aplikasi ini, beberapa komponen penting yang diperlukan antara lain:

1. Software Python

Penggunaan Python memungkinkan penulisan dan menjalankan kode program machine learning dengan efisien. Pustaka seperti TensorFlow, Scikit-learn, dan PyTorch sangat berguna untuk berbagai kebutuhan machine learning (Sundaram dkk., 2023).

2. Visual Studio Code (VS Code) Editor

VS Code merupakan lingkungan pengembangan terintegrasi (IDE) yang populer dan kaya fitur. VS Code sangat membantu dalam menulis, mengedit, dan debugging kode Python. Fitur debugging yang kuat dan berbagai ekstensi yang tersedia meningkatkan produktivitas dalam pengembangan aplikasi.

3. Heroku

Heroku adalah platform cloud yang memungkinkan penyebaran aplikasi web dengan mudah. Heroku mendukung berbagai bahasa pemrograman dan framework, termasuk Python, serta menyediakan layanan hosting yang skalabel dan andal. Dengan Heroku, aplikasi web yang dikembangkan dapat diakses oleh pengguna melalui internet dengan mudah dan cepat(Aluka dkk., 2023).

4. Open Street Map

Aplikasi Mapsbox aplikasi peta yang dapat diunduh melalui laman www.openstreetmap.org . OpenStreetMap (OSM) merupakan proyek global yang dikelola oleh komunitas sukarelawan dengan tujuan menciptakan dan memelihara basis data dan peta dunia yang dapat diakses dan diedit secara bebas. Proyek ini

telah berhasil mengumpulkan hampir 7,5 miliar titik data dari sekitar 1,8 juta pengguna hingga Maret 2022, menjadikannya contoh utama dari proyek geoinformasi bersumber dari kerumunan. Selain menarik perhatian industri, pemerintah, dan organisasi kemanusiaan, OSM juga mendapat minat besar dari komunitas akademik yang memanfaatkan data OSM untuk berbagai studi lintas disiplin. OSM terus berkembang sebagai sumber data penting yang memberikan kontribusi signifikan dalam berbagai penelitian dan aplikasi praktis (Grinberger dkk., 2022).

Penggunaan keempat komponen tersebut sangat mempermudah proses pengembangan dan penyebaran aplikasi berbasis web yang mengimplementasikan algoritme machine learning. Python memberikan fleksibilitas dalam pengembangan model, VS Code menyediakan lingkungan pengembangan yang efisien, dan Heroku memastikan aplikasi dapat diakses dengan mudah oleh pengguna akhir, menjadikan proses penyebaran lebih cepat dan efektif.