Penentuan Destinasi Wisata Favorit di Yogyakarta dengan Menggunakan Algoritma Decision Tree

Abdul Hamid Wijaya[1], Andriany Vandaningtyas[2], dan Rinci Kembang Hapsari[3]*

[1,2,3]Teknin Informatika, Fakultas Teknik Elektro dan Teknologi Informasi, Institut Teknologi Adhi Tama Surabaya, Surabaya, Jawa Timur

**Penulis Korespondensi : rincikembang@itats.ac.id*

ABSTRACT

Living life cannot be separated from tourism activities. Through tourism, a person can experience a new atmosphere and enjoy the pleasures that he may have never experienced during his daily routine, so tourism activities become an important element in the lives of the community. With the spread of various tourist destinations in Indonesia, a real role is needed to optimize the services of tourism managers in terms of services and facilities. Yogyakarta is one of the choices of tourists from all over Indonesia, and the island of Java, in particular, is a favorite tourist destination in Indonesia. There are various tours in Yogyakarta, namely cultural tourism, nature tourism, culinary tourism, and shopping places. This study uses the Decision Tree Algorithm to determine the Favorite Tourist Destinations in Yogyakarta. The data used is a tourist destination dataset, secondary data downloaded from Kaggle.com. Based on the model produced and tested on the test data, an average accuracy value of 93.66%, an average precision value of 96.89%, and an average sensitivity or recall value of 76.84% were obtained.

Article History

Received : 05-03-2025
Revised : 11-03-2025
Accepted : 20-03-2025

Keywords

Destinasi wisata
Decision Tree
Akurasi
Presisi
Sensitivitas

ABSTRAK

Menjalani kehidupan tidak dapat dipisahkan dari kegiatan berwisata. Melalui berwisata, seseorang memiliki kesempatan untuk merasakan suasana baru dan menikmati kesenangan yang mungkin belum pernah dia alami selama menjalani rutinitas sehari-hari maka kegiatan berwisata menjadi suatu elemen penting dalam kehidupan Masyarakat. Dengan tersebarnya berbagai destinasi wisata di Indonesia maka perlu peran nyata untuk mengoptimalkan pelayanan pengelola wisata dari segi pelayanan maupun fasilitas. Yogyakarta menjadi salah satu pilihan wisatawan dari seluruh penjuru Indonesia umumnya atau pulau jawa khusunya tujuan wisata favorit di Indonesia. Berbagai wisata terdapat di Yogyakarta yaitu wisata budaya, wisata alam, wisata kuliner dan tempat berbelanja. Penelitian ini bertujuan untuk menentukan Destinasi Wisata Favorit di Yogyakarta menggunakan Algoritma Decision Tree. Data yang digunakan merupakan dataset destinasi wisata yang merupakan data sekunder yang diunduh dari Kaggle.com. Berdasarkan model yang telah dihasilkan den diujikan pada data pengujian nilai rata-rata untuk akurasi 93,66%, nilai rata-rata untuk presisi 96,89% dan nilai rata-rata untuk sensitivitas atau recall 76,84%.

PENDAHULUAN

Yogyakarta merupakan destinasi wisata yang populer di Indonesia, dikenal dengan kekayaan warisan budaya, pemandangan alam yang indah, dan beragamnya aktivitas yang dapat dilakukan. Seiring dengan upaya kota ini untuk bersaing dengan destinasi "10 Bali Baru", maka penting untuk memahami preferensi dan tingkat kepuasan wisatawan guna mengoptimalkan pengembangan sektor pariwisata. [1]. Objek wisata di Daerah Istimewa Yogyakarta menawarkan potensi yang menarik dengan berbagai destinasi di area seluas 3.185,80 km2. Untuk mengidentifikasi faktor-faktor yang mempengaruhi pilihan wisatawan terhadap destinasi wisata, penelitian ini bertujuan untuk mengimplementasikan algoritma Decision Tree guna menentukan destinasi wisata yang paling diminati di Yogyakarta. Terdapat 78 kecamatan yang terbagi dalam 4 kabupaten dan 1 kota. Dimana di Daerah Istimewa Yogyakarta terdapat 126 objek wisata yang tersebar terletak di 4 kabupaten dan 1 kota. Tidak mengherankan jika Daerah Istimewa Yogyakarta menjadi destinasi wisata bagi wisatawan domestik maupun mancanegara karena banyaknya objek wisata yang dapat dikunjungi. Penelitian ini bertujuan untuk mengimplementasikan algoritma Decision Tree guna mengidentifikasi faktor-faktor kunci yang mempengaruhi pemilihan destinasi wisata favorit di Yogyakarta. Salah satu

metode yang dapat digunakan untuk membantu pengambilan keputusan dalam pemilihan lokasi wisata adalah algoritma Decision Tree. Decision Tree merupakan salah satu teknik dalam machine learning yang digunakan untuk membuat model prediktif berdasarkan serangkaian aturan keputusan. Manfaat penggunaan algoritma Decision Tree antara lain kemampuannya dalam menangani data numerik maupun kategorik, interpretabilitasnya, serta kemampuannya dalam mengidentifikasi fitur-fitur penting. Penelitian yang ada telah mengeksplorasi penggunaan berbagai sistem pendukung keputusan dan metode pengambilan keputusan multikriteria untuk membantu pemilihan destinasi wisata. Algoritma ini dapat digunakan untuk mengklasifikasikan dan memprediksi destinasi wisata yang paling disukai berdasarkan berbagai faktor, seperti preferensi wisatawan, aksesibilitas, dan infrastruktur.

Beberapa penelitian sebelumnya telah membuktikan keefektifan algoritma Decision Tree di berbagai bidang. Algoritma Decision Tree yang digunakan untuk klasifikasi data telah menjadi dasar pengembangan model prediktif, termasuk dalam penelitian Setiawan (2020) yang menerapkan algoritma ini untuk merekomendasikan destinasi wisata berdasarkan ulasan pengguna [2]. Penelitian lain oleh Yahya dkk. menggunakan metode Fuzzy SAW untuk mengembangkan sistem pendukung keputusan pemilihan objek wisata di Kabupaten Berau [3]. Dalam literatur yang ada, penelitian telah memanfaatkan algoritma Pohon Keputusan untuk mengklasifikasikan dan merekomendasikan destinasi wisata pilihan kepada wisatawan berdasarkan preferensi mereka dan faktor-faktor relevan lainnya, yang menunjukkan potensi pendekatan ini dalam domain pariwisata.

Algoritma pohon keputusan adalah alat yang ampuh untuk klasifikasi dan pengambilan keputusan, karena dapat dengan cepat dan akurat mengidentifikasi fitur-fitur paling berpengaruh yang menentukan hasil tertentu [4]. Menyadari potensi algoritma ini, penelitian ini bertujuan untuk memanfaatkan algoritma Pohon Keputusan untuk menganalisis faktor-faktor utama yang membentuk preferensi wisatawan terhadap destinasi favorit mereka di Yogyakarta. Dengan menerapkan algoritma ini pada data tentang preferensi dan pengalaman wisatawan di Yogyakarta, penelitian ini akan memberikan wawasan yang dapat memandu pengembangan strategi dan infrastruktur pariwisata untuk memenuhi kebutuhan dan harapan pengunjung dengan lebih baik.

TINJAUAN PUSTAKA

Pariwisata

Pariwisata merupakan kegiatan yang melibatkan perpindahan orang dari satu tempat ke tempat lain untuk berbagai tujuan, seperti rekreasi, bisnis, atau pendidikan. Menurut Soekadijo, pariwisata memiliki beberapa karakteristik, yaitu:

1. Pariwisata tidak dapat disimpan atau diproduksi.
2. Pariwisata bersifat musiman.
3. Permintaan pariwisata bersifat elastis.
4. Pariwisata erat kaitannya dengan faktor-faktor eksternal.

Pariwisata melibatkan berbagai pemangku kepentingan, seperti wisatawan, pemerintah, masyarakat lokal, dan industri pariwisata. Pengembangan pariwisata melibatkan dua aspek utama, yaitu pengembangan tujuan kawasan terkait pariwisata serta untuk pengembangan objek daya tarik wisata. Siklus evolusi daerah pariwisata bergantung pada peran pemangku kepentingan dan strategi pengembangan yang diimplementasikan. Dimana Siklus evolusi daerah pariwisata melalui beberapa tahap yang melibatkan proses perkembangan dan peningkatan aktivitas pariwisata: [1] [3]

1. Exploration (Penjajakan), Pengenalan produk pariwisata baru. Awal mula Wisatawan datang dalam jumlah sedikit.
2. *Involvement*, Mulai adanya keterlibatan masyarakat lokal dan pemerintah.
3. Development (pengembangan): Pembangunan infrastruktur untuk pariwisata dan peningkatan jumlah wisatawan.
4. Penggabungan Bijiksus, Daerah wisata sudah berkembang dengan baik.
5. Stagnation Penurunan jumlah wisatawan yang datang.

6. Decline atau Rejuvenation, Penurunan atau peremajaan produk pariwisata.

Dengan pemahaman ini, siklus evolusi daerah pariwisata mencerminkan perjalanan kompleks dari penjajakan hingga kemungkinan peremajaan untuk menjaga daya tarik pariwisata suatu daerah. Upaya pemasaran dalam pengembangan pariwisata melibatkan beberapa aspek penting, seperti pengembangan potret wisatawan dan potret pariwisata, mengembangkan kemitraan pemasaran, dan tidak berhenti untuk mempromosikan pariwisata. Kebijakan pariwisata, seperti yang dijelaskan oleh Hidayat, harus mencakup visi, misi, kebijakan, program, dan anggaran.

Algoritma Decision Tree

Algoritma Desision Tree atau Pohon Keputusan adalah salah satu metode klasifikasi dalam data mining yang digunakan untuk memprediksi nilai suatu variabel terikat berdasarkan nilai-nilai variabel bebas. Algoritma ini membentuk suatu struktur pohon yang terdiri dari node (simpul), edge (cabang), dan daun (leaf). Pada node internal, terdapat tes kondisi atribut, sedangkan daun merepresentasikan label kelas atau nilai prediksi[5]. Algoritma Decision Tree menggunakan pendekatan "divide and conquer" untuk membangun pohon keputusan.

Algoritma Decision Tree merupakan suatu pilihan teknik dalam Machine Learning yang dapat digunakan untuk memprediksi nilai target dengan membuat sebuah pohon keputusan berdasarkan aturan-aturan yang dihasilkan dari data. Tahapan-tahapan dalam pembuatan algoritma Decision Tree meliputi:

1. Pemilihan Fitur, identifikasi fitur-fitur dari data yang akan digunakan untuk menentukan aturan pembagian node. Pemilihan fitur dilakukan berdasarkan pada kriteria seperti kepentingan fitur terhadap prediksi, kemampuan fitur untuk membagi data dengan baik, dll.
2. Perhitungan Kriteria Pemilihan Node: Setelah fitur terpilih, selanjutnya menghitung nilai kriteria pemilihan node seperti Gini Index, Gain Ratio, Info Gain, untuk menentukan fitur terbaik yang akan digunakan sebagai node dalam pohon pada setiap langkahnya. Tujuannya adalah membagi data menjadi subset yang paling homogen mungkin dalam hal nilai target.
3. Pembuatan pohon keputusan : berdasarkan fitur terbaik, dibangun pohon keputusan dengan membelah node induk menjadi beberapa node anak. Proses membagi dataset berdasarkan pada kriteria yang telah ditentukan sebelumnya. Langkah ini dilakukan secara rekursif, dimulai dari node awal (atau root node) kemudian dipecah menjadi cabang-cabang yang lebih kecil berdasarkan fitur-fitur yang dipilih.

Proses pemilihan atribut akar pada Decision Tree mengacu kepada nilai Gain tertinggi dari beberapa atribut yang tersedia. Untuk menghitung nilai Gain, langkah pertama adalah menentukan nilai Entropy. Dimana untuk menghitung Entropy digunakan persamaan (1)

$$Entropy(S) = -\sum_{i=1}^{n} p_i * \log_2 p_i \quad (1)$$

Dimana:

S= Himpunan Kasus
n = Jumlah Partisi S
p_i= Proporsi dari S_i terhadap S

ISSN: 2774-4345, DOI: 10.31284/j.kernel.2023.v4i2.7473

Nilai Gain, didapatkan dengan menggunakan persamaan (2)

$$Gain\left(S,A\right) = Entropy(S) - \sum_{i=1}^{n} \frac{\left|S\right|}{\left|S_i\right|} * Entropy(S_i) \quad (2)$$

Dimana:

S= Himpunan Kasus
A = Atribut
n = Jumlah Partisi Atribut
$\left|S_i\right|$= Jumlah Kasus pada partisi ke-i
$\left|S\right|$ = Jumlah Kasus dalam S

Decision Tree (Pohon Keputusan) adalah bentuk pengklasifikasi sederhana yang didasarkan pada Classification and Regression Tree oleh Breiman et al. Setiap node internal di pohon keputusan yang merepresentasikan variabel atribut, sementara setiap cabang merepresentasikan keadaan dari variabel tersebut. Daun pada pohon menyatakan nilai yang diinginkan dari suatu kelas variabel yang menjadi tujuan prediksi. Dalam membangun pohon keputusan, kriteria pemisahan, termasuk pembuatan cabang dan kriteria penghentian, menjadi aspek penting. Pohon keputusan dibangun menggunakan data pembelajaran, sementara data uji digunakan untuk menguji model[6]. Keuntungan *Decision Tree* meliputi fleksibilitas dalam tugas data mining adalah:

1. Kejelasan dan kemudahan interpretasi
2. Penanganan berbagai jenis input data
3. Adaptabilitas terhadap dataset yang diperkirakan memiliki error atau nilai yang loss
4. Performa prediktif tinggi dengan kerja komputasi yang relatif kecil.

Rapid Miner

RapidMiner adalah perangkat lunak open source untuk analisis data dan text mining. RapidMiner menawarkan berbagai algoritma Machine Learning, seperti Decision Tree, Linear Regression, Support Vector Machine, dan lain-lain. RapidMiner memungkinkan pengguna untuk membangun dan mengeksekusi proses analitik dengan antarmuka drag-and-drop

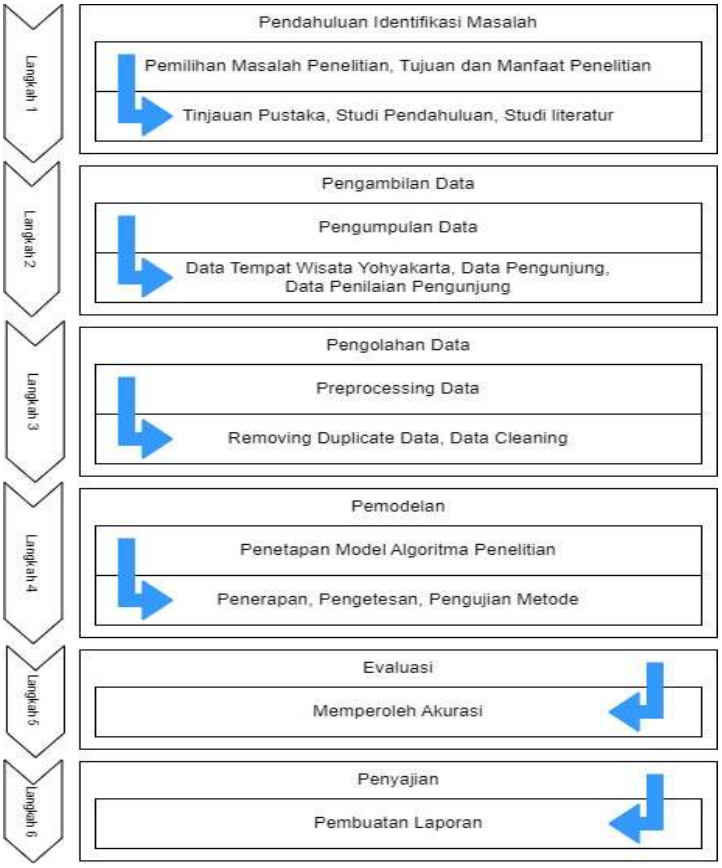
RapidMiner adalah perangkat lunak open source untuk analisis data dan text mining. RapidMiner menawarkan berbagai algoritma Machine Learning, seperti Decision Tree, Linear Regression, Support Vector Machine, dan lain-lain. RapidMiner memungkinkan pengguna untuk membangun dan mengeksekusi proses analitik dengan antarmuka drag-and-drop.

1. Mengimpor Data: Pada tahap ini, data yang akan dianalisis diimport ke RapidMiner.
2. Preprocessing Data: Proses membersihkan dan mempersiapkan data agar siap digunakan untuk pemodelan.
3. Pemodelan: Pada tahap ini, algoritma Machine Learning seperti Decision Tree diterapkan pada data untuk membangun model prediktif.
4. Evaluasi Model: Hasil model dievaluasi menggunakan metrik seperti akurasi, F1-score, dll untuk menilai kinerjanya.
5. Deployment: Model yang telah dibangun dapat disimpan dan digunakan untuk membuat prediksi pada data baru

ISSN: 2774-4345, DOI: 10.31284/j.kernel.2023.v4i2.7473

METODE

Dalam penelitian ini terdiri dari 6 langkah, yaitu identifikasi maslah, pengambilan data, pengolahan data, pemodelan, evaluasi dan penyajian. Detail dari setiap langkah ditunjukkan pada Gambar 1.



Gambar 1. Kerangka penelitian

Pengumpulan data

Data yang digunakan dalam penelitian ini merupakan data sekunder, yang bersumber dari Kaggle dengan link dataset adalah https://www.kaggle.com/datasets/athreal/destinasi-wisata-dataset. Dataset tersebut terdiri dari 2.870 record, dan 15 atribut, Seperti yang ditunjukkan pada Gambar 2.

Preprocessing Data

Preprocessing data merupakan salah satu tahap penting dalam data mining yang bertujuan untuk meningkatkan kualitas data sebelum dilakukan analisis lebih lanjut. Data yang belum diproses sering kali mengandung noise, data yang hilang, inkonsistensi, atau format yang tidak sesuai, sehingga diperlukan langkah-langkah preprocessing untuk memastikan akurasi dan reliabilitas hasil analisis. Dalam penelitian ini preprocessing data yang digunakan dengan melakukan pembersihan data (*data cleaning*). Proses ini bertujuan untuk menghapus data yang tidak lengkap, menangani nilai yang hilang, menghilangkan data duplikat, pembersihan data ikat, serta mengatasi inkonsistensi dalam data.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Place_I	Place_I	Descrip	Catego	City	Price	Rating	Time_N	Coordir	Latitude	Longitu	User_Ic	Dari Kota/Location	Umur	Place_Ratings
2	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	2	Bekasi, Jawa Barat	21	Bagus
3	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	23	Jakarta Barat, DKI Jakarta	37	Bagus
4	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	25	Lampung, Sumatera Selatan	40	Bagus
5	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	39	Cilacap, Jawa Tengah	32	Bagus
6	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	43	Cirebon, Jawa Barat	33	Bagus
7	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	52	Karawang, Jawa Barat	18	Bagus
8	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	60	Surabaya, Jawa Timur	25	Bagus
9	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	68	Kota Gede, DIY	37	Bagus
10	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	72	Tanggerang, Banten	19	Bagus
11	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	74	Semarang, Jawa Tengah	30	Bagus
12	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	88	Jakarta Timur, DKI Jakarta	27	Bagus
13	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	93	Lampung, Sumatera Selatan	18	Bagus
14	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	104	Jakarta Selatan, DKI Jakarta	27	Bagus
15	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	132	Ponorogo, Jawa Timur	31	Bagus
16	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	140	Surabaya, Jawa Timur	25	Bagus
17	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	157	Yogyakarta, DIY	28	Bagus
18	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	159	Depok, Jawa Barat	20	Bagus
19	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	165	Cirebon, Jawa Barat	23	Bagus
20	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	182	Solo, Jawa Tengah	34	Bagus
21	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	233	Solo, Jawa Tengah	23	Bagus
22	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	247	Bekasi, Jawa Barat	28	Bagus
23	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	262	Bogor, Jawa Barat	38	Bagus
24	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	269	Cirebon, Jawa Barat	36	Bagus
25	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	276	Lampung, Sumatera Selatan	39	Bagus
26	85	Taman Pir	Taman Pir	Taman Hit	Yogyakart	6000	4,5	120	{'lat': -7,80	-7,80067	110,368	282	Yogyakarta, DIY	39	Bagus

Gambar 2. Dataset yang terdiri dari 15 atribut

HASIL DAN PEMBAHASAN

Sebelum melakukan pembuatan model dengan *decision tree,* dilakukan proses *preprocessing* data. Dalam proses ini dilakukan untuk menghapus data yang duplikat. Dengan menggunakan RapidMiner untuk menghapus data yang duplikat digunakan operator "remove duplicate", yang ditunjukkan pada Gambar 3.



Gambar 3. Langkah preprocessing data dengan RapidMiner

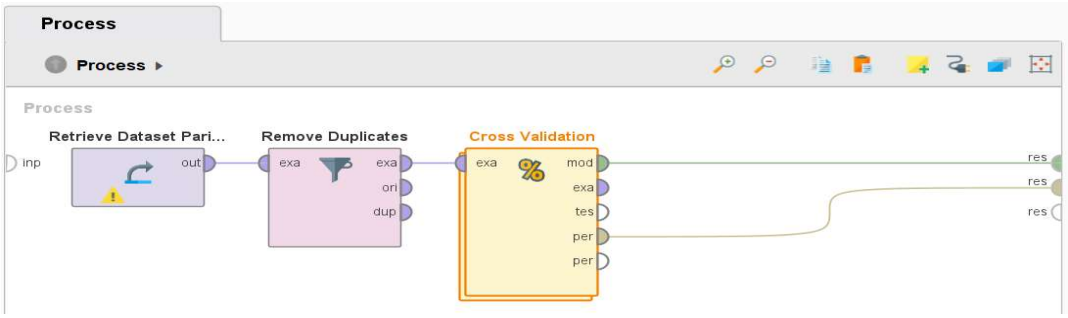
Hasil dari preprocessing data dari dataset ditunjukkan pada Gambar 4.

Row No.	Place_Ratin...	Place_Id	Place_Name	Description	Category	City	Price	Rating	Time
1	Bagus	85	Taman Pintar...	Taman Pintar...	Taman Hiburan	Yogyakarta	6000	4.500	120
2	Bagus	85	Taman Pintar...	Taman Pintar...	Taman Hiburan	Yogyakarta	6000	4.500	120
3	Bagus	85	Taman Pintar...	Taman Pintar...	Taman Hiburan	Yogyakarta	6000	4.500	120
4	Bagus	85	Taman Pintar...	Taman Pintar...	Taman Hiburan	Yogyakarta	6000	4.500	120
5	Bagus	85	Taman Pintar...	Taman Pintar...	Taman Hiburan	Yogyakarta	6000	4.500	120
6	Bagus	85	Taman Pintar...	Taman Pintar...	Taman Hiburan	Yogyakarta	6000	4.500	120
7	Bagus	85	Taman Pintar...	Taman Pintar...	Taman Hiburan	Yogyakarta	6000	4.500	120
8	Bagus	85	Taman Pintar...	Taman Pintar...	Taman Hiburan	Yogyakarta	6000	4.500	120
9	Bagus	85	Taman Pintar...	Taman Pintar...	Taman Hiburan	Yogyakarta	6000	4.500	120
10	Bagus	85	Taman Pintar...	Taman Pintar...	Taman Hiburan	Yogyakarta	6000	4.500	120
11	Bagus	85	Taman Pintar...	Taman Pintar...	Taman Hiburan	Yogyakarta	6000	4.500	120
12	Bagus	85	Taman Pintar...	Taman Pintar...	Taman Hiburan	Yogyakarta	6000	4.500	120
13	Bagus	85	Taman Pintar...	Taman Pintar...	Taman Hiburan	Yogyakarta	6000	4.500	120

ExampleSet (2,870 examples,1 special attribute,19 regular attributes)

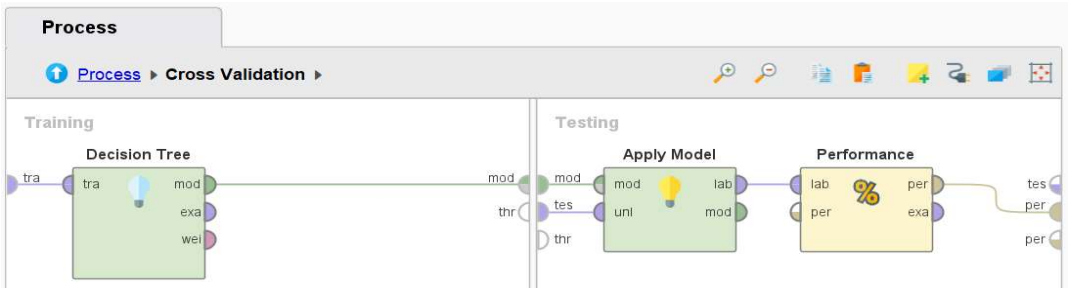
Gambar 4. Hasil Preprocessing dengan Rapid Miner

Hasil *preprocessing* data operator "*remove duplicates*" sehingga tidak ada lagi data ganda. Selanjutnya diterapkan pengujian menggunakan operator 10-*fold cross-validation* dan operator Algoritma *Decesion Tree*. Gambar 5 menunjukkan penggunaan operator "Cross Validation" yang digunakan untuk melakukan split data pengujian dengan 10-*fold cross-validation.*



Gambar 5. Pengujian dengan operator 10-*fold cross-validation*

Setelah dataset displit untuk data training dan data testing, sistem dilanjutkan untuk mengimplementasikan algoritma *decision tree* dengan menggunakan operator "Decision Tree". Hasil pemodelan dari tahap training digunakan untuk melakukan pengujian dan pengukuran kinerja dari algoritma tersebut. Dimana prosesnya ditunjukkan pada Gambar 6.



Gambar 6. Penerapan Algoritma *Decesion Tree*

Hasil model yang didapat dari proses training, didapatkan pohon keputusan yang ditunjukkan pada Gambar 7.



Gambar 7. Hasil pohon keputusan algoritma *Decesion Tree*

Hasil Model *decision tree* yang dihasilkan dapat dijelaskan bahwa: dimana atribut akarnya adalah Rating Pengunjung, dan Rating tersebut memiliki tiga kelompok kelas atau range yaitu range

1. dengan nilai <= 4.25 range 2. dengan nilai < 4.25 - 4.55 dan range 3. dengan nilai > 4.55. Dari tiga kelompok kelas tersebut, kelompok kelas rating <= 4.25 mendefinisikan kasus kedalam pilihan rating Cukup Bagus, sedangkan untuk kelompok kelas dengan rating < 4.25 - 4.55 mendefinisikan kasus kedalam pilihan Bagus dan kelompok kelas dengan rating > 4.55 mendefinisikan kasus kedalam pilihan Sangat Bagus.

Analisa pengukuran kinerja algoritma terhadap dataset berdasarkan 10-*fold cross-validation* dengan menggunakan aplikasi RapidMiner, sehingga didapatkan Confution Matrix yang ditunjukkan pada Gambar 8.

	true Bagus	true Sangat Bagus	true Cukup Bagus
pred. Bagus	1767	0	182
pred. Sangat Bagus	0	841	0
pred. Cukup Bagus	0	0	80

Gambar 8 Confution Matrix Akurasi Decesion Tree dengan Rapid Miner

Berdasarkan confution matrix yang dihasilkan digunakan untuk mengukur kinerja algoritma, dengan menghitung nilai akurasi, presisi, dan recal atau sensitifitas. Dimana perhitungan akurasi, didapatkan dengan :

$$Accuracy = \frac{1767+841+80}{1767+0+0+0+841+0+182+0+80}$$
$$= \frac{2688}{2879} = 0,9366$$

Nilai presisi didapatkan dengan :

$$\Pr ecesion_1 = \frac{1767}{1767+0+182}$$
$$= \frac{1767}{1949} = 0,9066$$
$$\Pr ecesion_2 = \frac{841}{0+841+0}$$
$$= \frac{841}{841} = 1$$
$$\Pr ecesion_3 = \frac{80}{0+0+80}$$
$$= \frac{80}{80} = 1$$

Sehingga rata-rata precision didapatkan :

$$\Pr ecesion = \frac{0,9066+1+1}{3} = \frac{2,9066}{3} = 0,9689$$

Sedangkan nilai sensitifitas atau *Recall,* didapatkan dengan :

$$Sensitivity_1 = \frac{1767}{1767+0+0} = \frac{1767}{1767} = 1$$

ISSN: 2774-4345, DOI: 10.31284/j.kernel.2023.v4i2.7473

$$Sensitivity_2 = \frac{841}{0+841+0} = \frac{841}{841} = 1$$
$$Sensitivity_3 = \frac{80}{182+0+80} = \frac{80}{262} = 0,3053$$

Sehingga rata-rata nilai sensitifitas atau *recall* adalah :

$$Sensitivity = \frac{1+1+0,3053}{3} = \frac{2,3053}{3} = 0,7684$$

Analisa hasil pengolahan data dengan menggunakan algoritma *Decesion Tree* terhadap dataset, ditampilkan pada tabel 1.

Tabel 1. Analisa pengukuran kinerja algoritma *Decesion Tree*

	Akurasi	**Presisi**	**Sensitivitas/Recall**
Min	0,9103	0,9066	0,3053
Max	1	1	1
Average	93,66%	0,9689	0,7684

KESIMPULAN

Dari hasil pengujian digunakan tes dengan implementasi operator 10-fold *Cros Validation dan Confusion Matrix.* Algoritma Decesion Tree menghasilkan nilai rata-rata Akurasi sebesar 93,66%, nilai rata-rata Presisi sebesar 96,89%, dan nilai rata-rata Sensitivitas atau *Recall* sebesar 76,84%. Dengan demikian Algoritma Decesion Tree merupakan algoritma dan teknik yang baik untuk memprediksi Destinasi Wisata Favorit di Yogyakarta sehingga dapat memberikan gambaran kepada calon wisatawan serta pengelola wisata untuk meningkatkan kualitas pada tempat wisata yang dikelola.

DAFTAR PUSTAKA

[1] T. Tantra, H. Millanyani, and I. M. Yefri, "Analisis Kepuasan Pengujung Berdasarkan Tourist Experience ke Daerah Istimewa Yogyakarta," Nov. 30, 2021. doi: 10.47065/ekuitas.v3i2.1122.
[2] V. Damayanti, A. M. Ekasari, and E. Syaodih, "The Determinants of Cultural Tourism Attractions Based on Tourist Satisfaction in the Sundapolis Area, Bandung City," Jan. 01, 2020. doi: 10.2991/assehr.k.200225.117.
[3] M. W. Suhantoro, "Sistem Pendukung Keputusan Rekomendasi Objek Wisata Di Kabupaten Berau Menggunakan Metode Fuzzy SAW," Nov. 30, 2021. doi: 10.47233/jsit.v1i2.114.
[4] A. Susanty, S. Nugroho, and A. Adyan, "OPTIMASI PENGEMBANGAN KAWASAN WISATA DI SEMARANG DENGAN MENGGUNAKAN METODEANALYTICAL HIERARCHY PROCESS, ANALISIS SWOT, DAN MULTI-ATTRIBUTE UTILITY THEORY," Jun. 01, 2015, Diponegoro University. doi: 10.12777/jati.10.2.77-84.
[5] B. A. R. P. Wahyu, A. F. Farozi, C. P. Mahendra, and R. K. Hapsari, 'Classification of Diabetes Sufferers Based on Decision Tree', INTEGER, vol. 8, no. 1, pp. 80–89, 2023, doi: https://doi.org/10.31284/j.integer.2023.v8i1.4423.
[6] P. K. Handayani, "Penerapan Principal Component Analysis untuk Peningkatan Kinerja Algoritma Decision Tree pada Iris Dataset," Jun. 30, 2020. doi: 10.24176/ijtis.v1i2.4939.
[7] M. A. Subarkah, W. M. Rahmawati, S. R. Wardhana, and R. K. Hapsari, 'Analisis Sentimen Terhadap Video Ulasan Produk Menggunakan Metode Support Vector Machine Dengan Sequential Minimal Optimization', KERNEL J. Ris. Inov. Bid. Inform. dan Pendidik. Inform., vol. 3, no. 2, pp. 17–24, 2022.