

PENERAPAN ALGORITMA C4.5 UNTUK MEMPREDIKSI TREN WISATA DI KABUPATEN MAJALENGKA

Suhendri¹, Mita Karmila²

^{1,2}Program Studi Informatika, Fakultas Teknik, Universitas Majalengka

Email: suhendri@unma.ac.id

ABSTRACT

Majalengka Regency is one of the tourist areas in Indonesia which has quite large tourism potential with various tourist attractions, such as natural tourism objects, cultural heritage and other attractions. However, to effectively optimize the development of the tourism sector in Majalengka Regency, a deep understanding of existing tourism trends and how they develop over time is needed. Therefore, predicting tourist trends can help local governments and tourism stakeholders to make better decisions. This research aims to predict tourism trends in Majalengka Regency in order to evaluate the effectiveness of strategies to improve it by utilizing the C4.5 Algorithm. The initial stage to determine the causes and effects of this tourism trend is by using the fishbone methodology. From the results of RapidMiner testing using the C4.5 Algorithm with Random tree, Apply Data and Performance operators, tourist trends that are very popular, popular and less popular obtained an accuracy of 86.67%.

Keywords: Data Mining, Prediction, C4.5 Algorithm, Random Tree, Accuracy.

ABSTRAK

Kabupaten Majalengka adalah salah satu daerah wisata di Indonesia yang memiliki potensi pariwisata yang cukup besar dengan berbagai atraksi wisata, seperti obyek wisata alam, warisan budaya dan daya tarik lainnya. Namun, untuk mengoptimalkan pengembangan sektor pariwisata di Kabupaten Majalengka secara efektif, diperlukan pemahaman yang mendalam tentang tren wisata yang ada dan bagaimana mereka berkembang dari waktu ke waktu. Oleh karena itu, memprediksi tren wisatawan dapat membantu pemerintah daerah dan stakeholder pariwisata untuk mengambil keputusan yang lebih baik. Penelitian ini bertujuan untuk memprediksi tren wisata di Kabupaten Majalengka agar dapat mengevaluasi keefektifan strategi untuk meningkatkannya dengan memanfaatkan Algoritma C4.5. Tahapan awal untuk mengetahui sebab dan akibat tren wisata ini dengan menggunakan metodologi fishbone. Dari hasil pengujian RapidMiner menggunakan Algoritma C4.5 dengan operator Random tree, Apply Data dan Performace bahwa tren wisata sangat populer, populer dan kurang populer memperoleh akurasi sebesar 86,67%.

Kata Kunci: Data Mining, Prediksi, Algoritma C4.5, Random Tree, Akurasi.

Riwayat Artikel :

Tanggal diterima : 16-11-2023

Tanggal revisi : 27-11-2023

Tanggal terbit : 29-11-2023

DOI :

<https://doi.org/10.31949/infotech.v9i2.7393>

INFOTECH journal by Informatika UNMA is licensed under CC BY-SA 4.0

Copyright © 2023 By Author



1. PENDAHULUAN

Menurut undang-undang no. Nomor 10 Tahun 2009 Tentang Pariwisata adalah berbagai kegiatan pariwisata yang didukung oleh berbagai fasilitas dan pelayanan yang disediakan oleh masyarakat, pengusaha, pemerintah, dan pemerintah daerah. Hingga saat ini, Indonesia telah menjadi salah satu tujuan wisata utama dunia yang memberikan berbagai penawaran kepada wisatawan untuk memenuhi kebutuhannya. Hal ini khususnya berdampak pada destinasi pariwisata daerah yang semakin mengoptimalkan daerahnya masing-masing.

Indonesia merupakan destinasi yang menarik bagi para wisatawan internasional. Keunikan budaya dan kekayaan alam yang beragam menjadi daya tarik utama pariwisata di Indonesia. Tak hanya itu, tersedia juga akomodasi yang terjangkau, kuliner khas, dan tradisi unik. Kombinasi faktor-faktor tersebut menjadikan Indonesia sebagai tujuan populer yang banyak dikunjungi oleh wisatawan dari berbagai belahan dunia (*Mu'minin, Fauziah, & Gunaryati, 2022*).

Kabupaten Majalengka adalah salah satu daerah wisata di Indonesia yang memiliki potensi pariwisata yang cukup besar dengan berbagai atraksi wisata, seperti obyek wisata alam, warisan budaya dan daya tarik lainnya. Namun, untuk mengoptimalkan pengembangan sektor pariwisata di Kabupaten Majalengka secara efektif, diperlukan pemahaman yang mendalam tentang tren wisata yang ada dan bagaimana mereka berkembang dari waktu ke waktu. Oleh karena itu, memprediksi tren wisatawan dapat membantu pemerintah daerah dan *stakeholder* pariwisata untuk mengambil keputusan yang lebih baik.

Visi yang tercantum dalam Rencana Induk Pengembangan Pariwisata Daerah Kabupaten Majalengka Tahun 2010 hingga 2025 adalah untuk menjadikan Kabupaten Majalengka sebagai destinasi ekowisata alam argo dan seni budaya yang bersaing, berkelanjutan, dan berbasis masyarakat yang beriman dan bertaqwa pada tahun 2025. Untuk mewujudkan visi tersebut, Disparbud Kabupaten Majalengka menetapkan sejumlah misi, di antaranya:

- a. Membangun destinasi pariwisata alam argo dan seni budaya yang memiliki daya saing di tingkat internasional. Ini akan dicapai dengan cara melindungi, mengembangkan, dan mengelola sumber daya wisata alam pegunungan, pertanian, serta seni budaya tradisional dan kontemporer yang mendukung kepariwisataan Kabupaten Majalengka secara berkelanjutan.
- b. Memajukan industri pariwisata kreatif yang berpihak pada masyarakat lokal, memiliki daya saing, dan mencerminkan nilai budaya. Selain itu, mengembangkan jaringan industri pariwisata baik di tingkat nasional maupun internasional untuk mendukung perekonomian Kabupaten Majalengka.

- c. Mengembangkan strategi pemasaran dan promosi pariwisata yang bertanggung jawab dan terpadu, yang juga berintegrasi dengan sektor ekonomi lain di Kabupaten Majalengka. Hal ini juga melibatkan destinasi pariwisata lain di Jawa Barat dan tingkat nasional.
- d. Meningkatkan struktur kelembagaan pariwisata dan memberdayakan Sumber Daya Manusia yang memiliki nilai-nilai keimanan dan ketakwaan. Hal ini bertujuan untuk memperkuat keterlibatan masyarakat dalam upaya pembangunan sektor pariwisata di Kabupaten Majalengka.

Tabel 1. Data kunjungan obyek daya tarik wisata (ODTW) Kabupaten Majalengka Tahun 2019-2022 (Disparbud Kab. Majalengka, 2023)

Tahun	Jumlah Wisatawan
2019	701.335
2020	375.000
2021	472.906
2022	818.543

Berdasarkan tabel 1 data yang diperoleh dari Dinas Pariwisata dan Kebudayaan Kabupaten Majalengka, kunjungan wisatawan yang berkunjung di Kabupaten Majalengka cenderung mengalami perubahan setiap tahunnya. Namun pada tahun 2020 menurun, hal ini dikarenakan adanya pandemi Covid-19. Tapi di tahun 2021 kembali menunjukkan kenaikan.

Destinasi pariwisata perlu menyediakan semua fasilitas yang diperlukan wisatawan agar mereka dapat merasa puas dan nyaman selama kunjungannya. Fasilitas-fasilitas ini mencakup sarana transportasi, tempat menginap, agen perjalanan, atraksi (termasuk kegiatan budaya, rekreasi, dan hiburan), layanan makanan dan minuman, serta souvenir (*Fauziah, 2021*).

Salah satu aset penting dalam industri pariwisata Kabupaten Majalengka adalah keberagaman objek wisata alam dan daya tarik khusus yang menjadi fokus pengembangan pariwisata di wilayah tersebut. Data dari Dinas Pariwisata dan Kebudayaan Majalengka menunjukkan bahwa Kabupaten Majalengka memiliki beragam obyek wisata yang bervariasi.

Perubahan jumlah kunjungan wisatawan di Kabupaten Majalengka terus menyulitkan Pemerintah Kabupaten Majalengka untuk menentukan strategi selanjutnya dalam pengembangan industri pariwisata, seperti promosi wisata dan memberikan pelayanan yang baik kepada wisatawan yang berkunjung, yang nantinya akan menjadi tren wisata di Kabupaten Majalengka.

Perkembangan pariwisata di Kabupaten Majalengka sebenarnya sudah baik di mata wisatawan, tempat wisata terus melakukan inovasi, terhadap daya tarik pariwisatanya (*Pebriana, Mulyawan, & Sutrisno, 2021*). Namun perkembangan destinasi wisata di daerah tersebut masih belum seimbang, serta sarana

dan prasarana pendukung belum cukup sempurna. Dampaknya, beberapa destinasi wisata mengalami peningkatan jumlah wisatawan namun ada juga yang sepi pengunjung. Oleh karena itu, Pemerintah Daerah perlu memikirkan strategi untuk menjamin pemerataan pengembangan destinasi wisata.

Kurangnya analisis guna mengevaluasi kebijakan strategi pemasaran obyek wisata dapat menyebabkan kesulitan dalam menilai keefektifan strategi pemasaran yang ada dan membuat perbaikan yang diperlukan, serta belum adanya analisis kebijakan guna membuat rencana strategi pengembangan tren obyek wisata. Penting untuk melakukan evaluasi dan analisis yang teratur terhadap upaya pemasaran dan pengembangan tren obyek wisata di Kabupaten Majalengka. Berdasarkan permasalahan tersebut, perlu dilakukan penelitian untuk memprediksi tren wisata di Kabupaten Majalengka yang akurat untuk masa yang akan datang.

Dalam era perkembangan teknologi informasi dan kecerdasan buatan, penggunaan analisis data menjadi semakin penting dalam pengambilan keputusan yang berkaitan dengan sektor pariwisata. Dalam konteks ini, penerapan metode prediksi menjadi penting untuk memprediksi tren wisata di Kabupaten Majalengka. Salah satu metode yang dapat digunakan adalah algoritma C4.5. Algoritma C4.5 adalah teknik klasifikasi pohon keputusan yang sering digunakan karena memiliki berbagai keunggulan dibandingkan dengan algoritma lainnya. Algoritma C4.5 mampu menghasilkan pohon keputusan yang mudah diinterpretasikan, mempunyai tingkat akurasi yang dapat diterima, dan dapat menangani atribut baik yang bersifat diskret maupun numerik. (Muhamad, Windarto, & Suhada, 2019).

Menurut Herdianto dalam (Rohmawati, Rohman, & Mujilahwati, 2017) menyatakan bahwa “Prediksi merupakan suatu proses sistematis dalam memproyeksikan kemungkinan terjadinya suatu kejadian di masa depan, berlandaskan pada informasi dari masa lalu dan situasi saat ini. Tujuannya adalah untuk meminimalkan kesalahan, yaitu perbedaan antara hasil prediksi dengan kenyataan yang terjadi. Penting untuk diingat bahwa prediksi tidak selalu memberikan jawaban pasti mengenai kejadian di masa depan, melainkan berupaya mencari perkiraan yang paling mendekati kemungkinan yang akan terjadi”.

Dengan menggunakan algoritma C4.5, dapat dilakukan analisis yang lebih mendalam terhadap variabel-variabel yang mempengaruhi tren wisata di Kabupaten Majalengka. Algoritma ini mampu mengatasi masalah *overfitting* dan mampu menangani data yang kompleks dengan baik.

Penelitian ini bertujuan untuk menerapkan algoritma C4.5 dalam memprediksi tren wisata di Kabupaten Majalengka. Dengan mengidentifikasi faktor-faktor yang mempengaruhi kunjungan wisatawan, diharapkan penelitian ini dapat memberikan pemahaman yang lebih baik tentang tren wisata di

daerah tersebut. Hasil penelitian ini dapat menjadi dasar bagi pihak-pihak terkait, seperti Dinas Pariwisata, pengelola atraksi wisata dan pelaku bisnis di sektor pariwisata untuk mengambil keputusan yang lebih cerdas dalam pengembangan wisata di Kabupaten Majalengka.

Dengan mempertimbangkan latar belakang tersebut, penerapan Algoritma C4.5 dapat menjadi solusi yang baik untuk memprediksi tren wisata di Kabupaten Majalengka dengan memanfaatkan berbagai faktor yang berpengaruh, serta dapat memberikan informasi yang berharga bagi pengambil keputusan di sektor pariwisata.

Berdasarkan uraian diatas, maka peneliti mengusulkan untuk mengambil judul penelitian yaitu “PENERAPAN ALGORITMA C4.5 UNTUK MEMREDIKSI TRENN WISATA DI KABUPATEN MAJALENGKA”.

2. LANDASAN TEORI

2.1. Data Mining

Menurut Connolly dan Begg dalam (Siregar & Puspabhuana, 2017) mengemukakan bahwa “*Data Mining* adalah suatu proses ekstraksi atau penggalian data yang belum diketahui sebelumnya, namun dapat dipahami dan berguna dari database yang besar serta digunakan untuk membuat suatu keputusan bisnis yang sangat penting.”

2.2. Decision Tree

Metode *Decision Tree* adalah suatu teknik klasifikasi dan prediksi yang terbukti sangat efektif dan populer. Pendekatan ini mengelompokkan informasi yang besar ke dalam struktur pohon keputusan yang memvisualisasikan aturan-aturan. Aturan-aturan ini dapat diungkapkan dalam bahasa alami atau melalui bentuk basis data seperti SQL untuk menemukan data tertentu. *Decision tree* adalah kerangka yang memungkinkan pemisahan dataset yang besar menjadi subset-subset data yang lebih kecil dengan menerapkan serangkaian aturan keputusan. (Muzakir & Wulandari, 2016).

2.3. Algoritma C4.5

Algoritma C4.5 merupakan algoritma yang digunakan untuk melakukan proses klasifikasi data dengan menggunakan teknik pohon keputusan. Algoritma C4.5 merupakan ekstensi dari algoritma ID3 dan menggunakan prinsip decision tree yang mirip. Algoritma ini sudah sangat terkenal dan disukai karena memiliki banyak kelebihan. Kelebihan ini misalnya dapat mengolah data numerik dan diskret, dapat menangani nilai atribut yang hilang, menghasilkan aturan-aturan yang mudah diinterpretasikan dan performanya merupakan salah satu yang tercepat dibandingkan dengan algoritma lain (Ferdian Harryanto & Hansun, 2017).

Kelebihan utama Algoritma C4.5 dapat membuat pohon keputusan (*decision tree*) yang efisien menangani atribut tipe diskrit dan tipe diskrit numerik, mudah untuk diinterpretasikan dan

memiliki tingkat akurasi yang dapat diterima (Muhamad, dkk., 2019). Berdasarkan (Muhamad, dkk., 2019) Secara umum Algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

- a. Pilih atribut sebagai akar
- b. Buat cabang untuk masing-masing nilai
- c. Bagi kasus dalam cabang
- d. Ulangi proses untuk masing-masing cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Untuk memilih atribut sebagai akar, didasarkan pada nilai gain tertinggi dari atribut-atribut yang ada. Untuk menghitung gain digunakan rumus seperti yang tertera berikut :

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n p_i \cdot Entropy(S_i) \tag{1}$$

Keterangan :

S : Himpunan kasus

A : Atribut

n : Jumlah patisi atribut A

|Si| : Jumlah kasus pada partisi ke i

|S| : Jumlah kasus dalam S

Sebelum mendapat nilai Gain adalah dengan mencari nilai Entropi. Entropi digunakan untuk menentukan seberapa informatif sebuah masukan atribut untuk menghasilkan sebuah atribut.

Rumus dasar :

$$Entropy(S) = - \sum_{i=1}^n p_i \cdot \log_2 p_i \tag{2}$$

Keterangan :

S : Himpunan Kasus

n : Jumlah partisi S

pi : Proporsi dari Si terhadap S

2.4. Prediksi (Forecasting)

Menurut Herdianto dalam (Rohmawati, dkk., 2017) mengemukakan bahwa “Prediksi adalah suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa lalu dan sekarang yang dimiliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan dapat diperkecil. Prediksi tidak harus memberikan jawaban secara pasti kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin yang akan terjadi.”

2.5. Pariwisata/Obyek Wisata

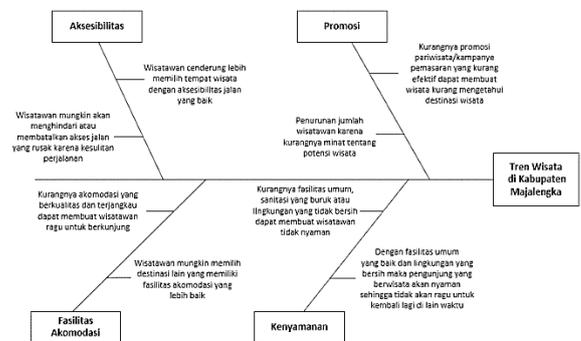
Menurut Mathieson & Wall dalam (Afandi, dkk., 2019) mengemukakan bahwa “Pariwisata merupakan serangkaian aktivitas yang berupa aktivitas perpindahan orang untuk sementara waktu ke suatu tujuan di luar tempat tinggal maupun tempat kerjanya, aktivitas yang dilakukannya selama tinggal di tempat tujuan tersebut dan kemudahan-kemudahan yang disediakan untuk memenuhi kebutuhannya baik selama dalam perjalanan maupun di lokasi tujuannya. Sedangkan pada UU No. 10 Tahun 2009, Pariwisata ialah berbagai macam kegiatan wisata dan didukung berbagai fasilitas serta layanan yang disediakan oleh masyarakat, pengusaha, pemerintah dan pemerintah daerah.”

Menurut Chafid Fandeli dalam (Fitriani, dkk., 2020) mengemukakan bahwa “Obyek wisata adalah perwujudan daripada ciptaan manusia, tata hidup, seni budaya serta sejarah bangsa dan tempat atau keadaan alam yang mempunyai daya tarik untuk dikunjungi wisatawan. Sedangkan obyek wisata alam adalah obyek wisata yang daya tariknya bersumber pada keindahan sumber daya alam dan tata lingkungannya.”

3. METODOLOGI PENELITIAN

3.1. Tahapan Penelitian

Tahap-tahap penelitian digunakan dengan tujuan untuk memberikan pemahaman yang lebih baik terkait rangkaian langkah yang dilakukan oleh peneliti dalam menjalankan penelitian pada tugas akhir ini. Penjelasan yang lebih rinci tentang langkah-langkah tersebut dapat ditemukan dalam diagram tahapan penelitian pada Gambar 1 di bawah ini :



Gambar 1. Metode fishbone

- a. Aksesibilitas
Aksesibilitas merupakan ukuran kemudahan lokasi untuk dijangkau dari lokasi lainnya melalui sistem transportasi. Aksesibilitas adalah faktor penting dalam menentukan popularitas dan potensi suatu destinasi wisata. Destinasi yang mudah diakses cenderung lebih menarik bagi wisatawan, wisatawan akan lebih memilih obyek wisata yang cenderung memiliki akses perjalanan yang baik dibandingkan perjalanan yang memiliki akses yang rusak karena akan cenderung kesulitan dalam perjalanan.
- b. Promosi
Promosi dalam konteks pariwisata mengacu pada serangkaian aktivitas dan strategi yang dilakukan untuk memasarkan, mempromosikan, dan mengkomunikasikan destinasi wisata atau layanan pariwisata kepada khalayak target. Tujuan utama dari promosi pariwisata adalah untuk menarik minat wisatawan potensial, membangun citra positif, dan mendorong kunjungan ke destinasi atau layanan yang dipromosikan. Kurangnya promosi pariwisata yang kurang efektif dapat membuat wisata kurang diketahui destinasi wisata tersebut akibatnya akan mengalami penurunan jumlah wisata karena kurangnya minat tentang potensi wisata.

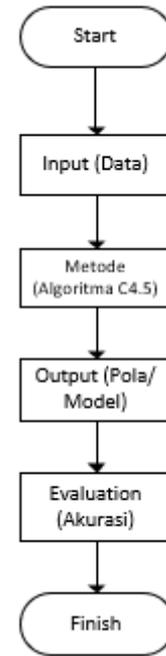
- c. Fasilitas Akomodasi
 Akomodasi adalah segala sesuatu yang disediakan untuk memenuhi kebutuhan atau keinginan seseorang saat berwisata. Akomodasi dalam dunia pariwisata merupakan segala hal yang disediakan oleh pihak penyedia jasa untuk kegiatan pariwisata baik dalam penyediaan bangunan untuk penginapan, tempat makan, tempat hiburan dan fasilitas lain yang dikelola secara komersial. Kurangnya akomodasi yang berkualitas dan terjangkau dapat membuat wisatawan ragu untuk berkunjung, wisatawan mungkin memilih destinasi lain yang memiliki fasilitas akomodasi yang lebih baik.
- d. Kenyamanan
 Kenyamanan adalah salah satu faktor yang dapat membuat pengalaman wisatawan menjadi lebih positif dan meningkatkan peluang untuk kembali atau merekomendasikan destinasi kepada orang lain. Dalam pengembangan pariwisata yang berkelanjutan, penting untuk memprioritaskan kenyamanan wisatawan dengan memahami preferensi dan kebutuhan mereka. Kurangnya fasilitas umum, sanitasi yang buruk atau lingkungan yang tidak bersih dapat membuat wisatawan tidak nyaman. Dengan fasilitas umum yang baik dan lingkungan yang bersih maka pengunjung yang berwisata akan nyaman sehingga tidak akan ragu untuk kembali lagi di lain waktu.

3.2. Pengumpulan Data

- a. Metode Observasi
 Observasi dilakukan dengan cara mendatangi Dinas Pariwisata dan Kebudayaan Kabupaten Majalengka untuk dimintai data serta mendatangi beberapa obyek wisata yang dikaji yang berada di Kabupaten Majalengka serta melakukan analisis permasalahan yang merupakan kebutuhan untuk mendapatkan informasi-informasi yang menunjang dalam penelitian ini.
- b. Metode Wawancara
 Metode ini dilakukan dengan cara bertanya jawab langsung dengan Staf yang berada di Dinas Pariwisata dan Kebudayaan Kabupaten Majalengka yang berkaitan dengan penelitian yaitu data kunjungan Obyek Daya Tarik Wisata (ODTW) Kabupaten Majalengka tahun 2019-2022.
- c. Metode Studi Literatur
 Dalam metode ini peneliti mengutip dari beberapa bacaan yang berkaitan dengan prediksi jumlah kunjungan wisata, algoritma C4.5 dan fishbone diagram. Beberapa hal yang dikutip dapat berupa teori ataupun pendapat dari berbagai artikel ilmiah dan jurna jurnal. Ini dimaksudkan untuk memberikan landasan teori yang kuat melalui buku-buku atau jurnal dan pengumpulan data dengan menggunakan fasilitas internet melalui mesin pencari (*search engine*).

3.3. Analisis Data Mining (Prediksi)

Pedekatan yang diusulkan untuk prosedur perhitungan untuk menetapkan prediksi tren wisata pada penerapan algoritma C4.5 dalam metodologi penelitian ini. Berikut adalah *flowchart* tahapan *data mining* (prediksi) :



Gambar 2. Flowchart proses data mining (prediksi)

4. PEMBAHASAN

4.1. Data Cleaning

Pembersihan Data dilakukan untuk menghilangkan data yang kurang relevan dalam dataset. Proses ini dilakukan menggunakan aplikasi *microsoft excel*. Pada penelitian ini menggunakan 4 atribut yaitu 3 atribut prediktor (Tahun, Jumlah Kunjungan, wisata dan jenis wisata) dan 1 atribut class yaitu Keterangan. Berikut adalah hasil dari data *cleaning* :

Tabel 2. Data cleaning

TAHUN	JML	WISATA	JENIS WISATA	KET
2019	31.442	CURUG MUARA JAYA	ALAM	POPULER
2019	58.115	TERASERING PANYAWEUYAN	ALAM	SANGAT POPULER
2019	75.594	PARALAYANG	ALAM	SANGAT POPULER
2019	26.839	SITU CIPANTEN	ALAM	POPULER
2019	7.844	TERASERING CIBOER PAS	ALAM	KURANG POPULER
2019	30.819	TALAGA PANCAR	ALAM	POPULER
2019	27.480	WATER BOOM TIRTA INDAH	REKREASI	POPULER
2019	54.921	PRABU SILIWANGI	BUDAYA	SANGAT POPULER
2019	26.299	CIKADONGDONG RIVER TUBING	ALAM	POPULER
2019	20.890	JEMBAR WATER PARK	REKREASI	POPULER

TAHUN	JML	WISATA	JENIS WISATA	KET
2019	4.753	TAMAN AIR RAJAWALI	REKREASI	KURANG POPULER
2019	30.422	BUPER PANTEN	ALAM	POPULER
2019	19.700	CURUG CIPEUTEUY	ALAM	POPULER
2019	15.300	SITU SANGIANG	ALAM	POPULER
2019	26.037	CADAS GANTUNG	ALAM	POPULER
2020	18.251	CURUG MUARA JAYA	ALAM	POPULER
2020	24.518	TERASERING PANYAWEUYAN	ALAM	POPULER
2020	19.943	PARALAYANG	ALAM	POPULER
2020	16.278	SITU CIPANTEN	ALAM	POPULER
2020	8.801	TERASERING CIBOER PAS	ALAM	KURANG POPULER
2020	12.505	TALAGA PANCAR	ALAM	POPULER
2020	16.626	WATER BOOM TIRTA INDAH	REKREASI	POPULER
2020	19.562	PRABU SILIWANGI	BUDAYA	POPULER
2020	14.195	CIKADONGDONG RIVER TUBING	ALAM	POPULER
2020	11.394	JEMBAR WATER PARK	REKREASI	POPULER
2020	8.841	TAMAN AIR RAJAWALI	REKREASI	KURANG POPULER
2020	5.545	BUPER PANTEN	ALAM	KURANG POPULER
2020	13.209	CURUG CIPEUTEUY	ALAM	POPULER
2020	11.090	SITU SANGIANG	ALAM	POPULER
2020	10.686	CADAS GANTUNG	ALAM	POPULER
2021	14.667	CURUG MUARA JAYA	ALAM	POPULER
2021	28.427	TERASERING PANYAWEUYAN	ALAM	POPULER
2021	20.991	PARALAYANG	ALAM	POPULER
2021	76.024	SITU CIPANTEN	ALAM	SANGAT POPULER
2021	24.542	TERASERING CIBOER PAS	ALAM	POPULER
2021	5.793	TALAGA PANCAR	ALAM	KURANG POPULER
2021	38.454	WATER BOOM TIRTA INDAH	REKREASI	POPULER
2021	22.162	PRABU SILIWANGI	BUDAYA	POPULER
2021	15.913	CIKADONGDONG RIVER TUBING	ALAM	POPULER
2021	35.171	JEMBAR WATER PARK	REKREASI	POPULER
2021	28.037	TAMAN AIR RAJAWALI	REKREASI	POPULER
2021	3.905	BUPER PANTEN	ALAM	KURANG POPULER
2021	12.347	CURUG CIPEUTEUY	ALAM	POPULER
2021	4.889	SITU SANGIANG	ALAM	KURANG POPULER
2021	1.982	CADAS GANTUNG	ALAM	KURANG POPULER
2022	26.315	CURUG MUARA JAYA	ALAM	POPULER
2022	109.980	TERASERING PANYAWEUYAN	ALAM	SANGAT POPULER

TAHUN	JML	WISATA	JENIS WISATA	KET
2022	31.031	PARALAYANG	ALAM	POPULER
2022	167.798	SITU CIPANTEN	ALAM	SANGAT POPULER
2022	24.542	TERASERING CIBOER PAS	ALAM	POPULER
2022	3.775	TALAGA PANCAR	ALAM	KURANG POPULER
2022	39.631	WATER BOOM TIRTA INDAH	REKREASI	POPULER
2022	35.573	PRABU SILIWANGI	BUDAYA	POPULER
2022	22.618	CIKADONGDONG RIVER TUBING	ALAM	POPULER
2022	79.954	JEMBAR WATER PARK	REKREASI	SANGAT POPULER
2022	28.616	TAMAN AIR RAJAWALI	REKREASI	POPULER
2022	28.674	BUPER PANTEN	ALAM	POPULER
2022	20.964	CURUG CIPEUTEUY	ALAM	POPULER
2022	16.781	SITU SANGIANG	ALAM	POPULER
2022	988	CADAS GANTUNG	ALAM	KURANG POPULER

Tabel 2. diatas merupakan data *cleaning* yang nantinya akan dilakukan pengujian akurasi dan model. Dari tabel tersebut dinyatakan atribut jumlah kunjungan mempengaruhi *class* keterangan yaitu total jumlah kunjungan dibagi menjadi (<10.000, 10.000-50.000, >50.000) untuk kategori popularitas. Jika jumlah kunjungan wisata <10.000 maka dinyatakan kurang populer, jika jumlah kunjungan wisata 10.000-50.000 maka dinyatakan populer dan jika jumlah kunjungan >50.000 maka dinyatakan sangat populer dari wisata lain.

4.2. Penerapan Algoritma

Perhitungan algoritma C4.5 dimulai dengan terlebih dahulu memilih atribut akar dengan mencari total kasus keseluruhan, total kasus keputusan yang sangat populer, jumlah kasus keputusan populer dan jumlah kasus keputusan kurang populer. Menghitung *entropy* dari semua kasus yang dibagi berdasarkan tahun, jumlah kunjungan, dan jenis wisata. Setelah itu dilakukan perhitungan *gain* untuk masing-masing atribut. Hasil perhitungan ditunjukkan oleh tabel berikut ini.

Tabel 3. Perhitungan *entropy* dan *gain*

Atribut	Nilai	Jml (S)	Sangat Populer (S1)	Populer (S2)	Kurang Populer (S3)	Entropy	Gain
Total		60	32	17	11	1,447880327	
Thn							0,318270075
	2019	15	11	2	2	1,103307409	
	2020	15	1	11	3	1,052981575	
	2021	15	8	3	4	1,456564762	
	2022	15	12	1	2	0,905587262	
Jml Kunj							1,447880327
	< 10.000	11	0	0	11	0	
	10.000-50.000	42	0	42	0	0	
	>50.000	7	7	0	0	0	
Jenis Wisata							0,117479049
	Alam	44	22	14	8	1,472830152	
	Budaya	4	3	1	0	0	
	Rekreasi	12	8	2	2	1,251629167	

Dari hasil perhitungan pada tabel 3 didapatkan hasil *gain* tertinggi yaitu 1,447880327 untuk atribut Jumlah Kunjungan. Atribut jumlah kunjungan memiliki kemungkinan besar untuk menjadi akar atau simpul dari pohon keputusan karena jumlah kunjungan adalah atribut yang lebih informatif dan menghasilkan pengurangan ketidakpastian yang lebih besar.

Tabel 4. *Entropy* dan *gain* total

Sangat Populer	Populer	Kurang Populer	Total
32	17	11	60

Menghitung *Entropy* Total, Diketahui bahwa data jumlah kunjungan obyek daya tarik wisata adalah 60, dengan keterangan 32 sangat populer, 17 populer, 11 kurang populer.

$$Entropy (Total) = \left(-\left(\frac{32}{60}\right) \times \log_2\left(\frac{32}{60}\right)\right) + \left(-\left(\frac{17}{60}\right) \times \log_2\left(\frac{17}{60}\right)\right) + \left(-\left(\frac{11}{60}\right) \times \log_2\left(\frac{11}{60}\right)\right)$$

$$Entropy (Total) = 1,447880327$$

Dari total data jumlah kunjungan obyek daya tarik wisata didapatkan hasil nilai *Entropy* sebesar 1,447880327.

Selanjutnya menghitung nilai *entropy* dan *gain* dari setiap atribut.

Tabel 5. *Entropy* dan *Gain* Tahun

Tahun	Sangat Populer	Populer	Kurang Populer	Total
2019	11	2	2	15
2020	1	11	3	15
2021	8	3	4	15
2022	12	1	2	15
Total				60

Menghitung nilai *entropy* untuk tahun 2019 sebagai berikut.

$$Entropy (Total) = \left(-\left(\frac{11}{15}\right) \times \log_2\left(\frac{11}{15}\right)\right) + \left(-\left(\frac{2}{15}\right) \times \log_2\left(\frac{2}{15}\right)\right) + \left(-\left(\frac{2}{15}\right) \times \log_2\left(\frac{2}{15}\right)\right)$$

$$Entropy (Total) = 1,103307409$$

Menghitung nilai *entropy* untuk tahun 2020 sebagai berikut.

$$Entropy (Total) = \left(-\left(\frac{1}{15}\right) \times \log_2\left(\frac{1}{15}\right)\right) + \left(-\left(\frac{11}{15}\right) \times \log_2\left(\frac{11}{15}\right)\right) + \left(-\left(\frac{3}{15}\right) \times \log_2\left(\frac{3}{15}\right)\right)$$

$$Entropy (Total) = 1,052981575$$

Menghitung *entropy* nilai untuk tahun 2021 sebagai berikut.

$$Entropy (Total) = \left(-\left(\frac{8}{15}\right) \times \log_2\left(\frac{8}{15}\right)\right) + \left(-\left(\frac{3}{15}\right) \times \log_2\left(\frac{3}{15}\right)\right) + \left(-\left(\frac{4}{15}\right) \times \log_2\left(\frac{4}{15}\right)\right)$$

$$Entropy (Total) = 1,456564762$$

Menghitung nilai *entropy* untuk tahun 2022 sebagai berikut.

$$Entropy (Total) = \left(-\left(\frac{12}{15}\right) \times \log_2\left(\frac{12}{15}\right)\right) + \left(-\left(\frac{1}{15}\right) \times \log_2\left(\frac{1}{15}\right)\right) + \left(-\left(\frac{2}{15}\right) \times \log_2\left(\frac{2}{15}\right)\right)$$

$$Entropy (Total) = 0,905587262$$

Menghitung nilai *Gain* atribut tahun sebagai berikut:

$$Gain = 1,447880327 - \left(\left(\frac{15}{60}\right) \times 1,103307409\right) - \left(\left(\frac{15}{60}\right) \times 1,052981575\right) - \left(\left(\frac{15}{60}\right) \times 1,456564762\right) - \left(\left(\frac{15}{60}\right) \times 0,905587262\right)$$

$$= 0,318270075$$

Tabel 6. Entropy dan gain jumlah kunjungan

Jumlah Kunjungan	Sangat Populer	Populer	Kurang Populer	Total
<10.000	0	0	11	11
10.000-50.000	0	42	0	42
>50.000	7	0	0	7
Total				60

Menghitung nilai *entropy* untuk tahun jumlah kunjungan <10.000 sebagai berikut.

$$Entropy (Total) = \left(-\left(\frac{0}{11}\right) \times \log_2\left(\frac{0}{11}\right)\right) + \left(-\left(\frac{0}{42}\right) \times \log_2\left(\frac{0}{42}\right)\right) + \left(-\left(\frac{11}{11}\right) \times \log_2\left(\frac{11}{11}\right)\right)$$

$$Entropy (Total) = 0$$

Menghitung nilai *entropy* untuk tahun jumlah kunjungan 10.000-50.000 sebagai berikut.

$$Entropy (Total) = \left(-\left(\frac{0}{42}\right) \times \log_2\left(\frac{0}{42}\right)\right) + \left(-\left(\frac{42}{42}\right) \times \log_2\left(\frac{42}{42}\right)\right) + \left(-\left(\frac{42}{42}\right) \times \log_2\left(\frac{42}{42}\right)\right)$$

$$Entropy (Total) = 0$$

Menghitung nilai *entropy* untuk tahun jumlah kunjungan >50.000 sebagai berikut.

$$Entropy (Total) = \left(-\left(\frac{7}{7}\right) \times \log_2\left(\frac{7}{7}\right)\right) + \left(-\left(\frac{0}{7}\right) \times \log_2\left(\frac{0}{7}\right)\right) + \left(-\left(\frac{0}{7}\right) \times \log_2\left(\frac{0}{7}\right)\right)$$

$$Entropy (Total) = 0$$

Menghitung Total *Gain* atribut Jumlah Kunjungan sebagai berikut :

$$Gain = 1,447880327 - \left(\left(\frac{11}{60}\right) \times 0\right) - \left(\left(\frac{42}{60}\right) \times 0\right) - \left(\left(\frac{7}{60}\right) \times 0\right)$$

$$= 1,447880327$$

Tabel 7. Entropy dan gain jenis wisata

Jenis Wisata	Sangat Populer	Populer	Kurang Populer	Total
Alam	22	14	8	44
Budaya	3	1	0	4
Rekreasi	8	2	2	12
Total				60

Menghitung nilai *entropy* untuk jenis wisata alam sebagai berikut.

$$Entropy (Total) = \left(-\left(\frac{22}{44}\right) \times \log_2\left(\frac{22}{44}\right)\right) + \left(-\left(\frac{14}{44}\right) \times \log_2\left(\frac{14}{44}\right)\right) + \left(-\left(\frac{8}{44}\right) \times \log_2\left(\frac{8}{44}\right)\right)$$

$$Entropy (Total) = 1,472830152$$

Menghitung nilai *entropy* untuk jenis wisata budaya sebagai berikut.

$$Entropy (Total) = \left(-\left(\frac{3}{4}\right) \times \log_2\left(\frac{3}{4}\right)\right) + \left(-\left(\frac{1}{4}\right) \times \log_2\left(\frac{1}{4}\right)\right) + \left(-\left(\frac{0}{4}\right) \times \log_2\left(\frac{0}{4}\right)\right)$$

$$Entropy (Total) = 0$$

Menghitung nilai *entropy* untuk jenis wisata rekreasi sebagai berikut.

$$Entropy (Total) = \left(-\left(\frac{8}{12}\right) \times \log_2\left(\frac{8}{12}\right)\right) + \left(-\left(\frac{2}{12}\right) \times \log_2\left(\frac{2}{12}\right)\right) + \left(-\left(\frac{2}{12}\right) \times \log_2\left(\frac{2}{12}\right)\right)$$

$$Entropy (Total) = 1,251629167$$

Menghitung Total *Gain* atribut Jenis wisata sebagai berikut :

$$Gain = 1,447880327 - \left(\left(\frac{44}{60}\right) \times 1,4728301523\right) - \left(\left(\frac{4}{60}\right) \times 0\right) - \left(\left(\frac{12}{60}\right) \times 1,251629167\right)$$

$$= 0,117479049$$

5. KESIMPULAN

Dari hasil uji data prediksi jumlah kunjungan wisata yang dapat mempengaruhi tren wisata di Kabupaten Majalengka dengan menggunakan algoritma C4.5. Maka kesimpulan yang diperoleh dari penelitian ini adalah :

1. Penerapan Algoritma C4.5 dapat dilakukan untuk memprediksi tren wisata di Kabupaten Majalengka. Setelah dilakukan pengujian bahwa tren wisata lebih dari 10.000 maka dinyatakan sangat populer yang artinya wisata tersebut sering menjadi tren kalangan wisata lain, jika tren wisata <10.000 dinyatakan kurang populer, jika tren wisata 10.000-50.000 dinyatakan populer dan jika tren wisata >50.000 maka wisata tersebut dinyatakan sangat populer. Pengujian menggunakan *rapidminer* dengan *operators random tree, apply model* dan *performance* memperoleh akurasi sebesar 86.67%. Hasil prediksi tersebut bisa digunakan untuk mengevaluasi strategi dan promosi untuk meningkatkan tren wisata.
2. Untuk meningkatkan evaluasi strategi pemasaran dan promosi dalam rangka meningkatkan tren wisata di Kabupaten Majalengka, yaitu setelah dilakukan pengujian menggunakan *rapidminer* maka selanjutnya analisis SWOT digunakan untuk memperoleh pandangan dasar tentang strategi yang diperlukan untuk mencapai tujuan tertentu, dalam hal ini untuk menilai inisiatif apa saja

yang memungkinkan alternatif dalam mengelola dan mengembangkan strategi.

PUSTAKA

- Afandi, A., Indrawan, A. K., & Ratnawardhani, E. A. (2019). ANALISIS MODEL PENILAIAN KOTA IDEAL DESTINASI WISATA. *Barista : Jurnal Kajian Bahasa Dan Pariwisata*, 6(1), 1–7. <https://doi.org/10.34013/barista.v6i1.166>.
- Eddyono, Fauziah. (2021). *Pengelolaan Destinasi Pariwisata*. Jawa Timur: Uwais Inspirasi Indonesia.
- Ferdian Harryanto, F., & Hansun, S. (2017). Penerapan Algoritma C4.5 untuk Memprediksi Penerimaan Calon Pegawai Baru di PT WISE. In *Maret* (Vol. 3, Issue 2).
- Fitriani, D., Masri, A. R., & Tajibu, K. (2020). STRATEGI KOMUNIKASI DALAM PENGEMBANGAN OBYEK WISATA PANTAI MANDALA RIA DI KABUPATEN BULUKUMBA. In *Jurnal Washiyah* (Vol. 1, Issue 1).
- Kementrian Pariwisata. (2009). Undang-Undang Republik Indonesia No. 10 Tahun 2009 Tentang Kepariwisata.
- Muhamad, M., Windarto, A. P., & Suhada, S. (2019). PENERAPAN ALGORITMA C4.5 PADA KLASIFIKASI POTENSI SISWA DROP OUT. *KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer)*, 3(1). <https://doi.org/10.30865/komik.v3i1.1688>.
- Muminin, F., & Gunaryati, A. (2022). Prediksi Kunjungan Wisatawan Mancanegara Melalui Pintu Udara Menggunakan ARIMA, Glmnet, dan Prophet Prediction of Foreign Tourist Visits via Airline using ARIMA, Glmnet, and Prophet. In *Februari* (Vol. 21, Issue 1).
- Muzakir, A., & Wulandari, R. A. (2016). Model Data Mining sebagai Prediksi Penyakit Hipertensi Kehamilan dengan Teknik Decision Tree. *Scientific Journal of Informatics*, 3(1). <http://journal.unnes.ac.id/nju/index.php/sji>.
- Pebriana, F., Mulyawan, R., & Sutrisno, B. (2021). Strategi Pemerintah Daerah Dalam Pengembangan Pariwisata Untuk Meningkatkan Pendapatan Asli Daerah. *Janitra (Jurnal Administrasi Pemerintahan)*, 1 (April), 11–22.
- Rohmawati, F., Rohman, G., & Mujilahwati, S. (2017). SISTEM PREDIKSI JUMLAH PENGUNJUNG WISATA WEGO KEC.SUGIO KAB.LAMONGAN MENGGUNAKAN METODE FUZZY TIME SERIES. In *JOUTICLA* (Vol. 3, Issue 2).
- Siregar, Amril Mutoi & Puspabhuana, Adam. (2017). *Data Mining : Pengolahan Data Menjadi Informasi dengan RapidMiner*. Surakarta : CV. Kekata Group.