

# PENERAPAN ALGORITMA *DECISION TREE* C4.5 UNTUK PREDIKSI CUACA DI KOTA SEMARANG

Ami Nafisatun Najah<sup>1,\*</sup>, Amira Nur Amalina<sup>2</sup>, Rahmat Hidayat<sup>3</sup>

<sup>1, 2, 3</sup> Program Studi Ilmu Komputer, Universitas Putra Bangsa Kebumen

Jl. Ronggowarsito No 18, Sudagaran, Kedawung, Kec. Pejagoan, Kab. Kebumen, Jawa Tengah 54361, Indonesia

e-mail: [aminafisa54@gmail.com](mailto:aminafisa54@gmail.com)<sup>1</sup>

(Naskah masuk : 18 Januari 2025 Diterima untuk diterbitkan : 4 Mei 2025)

## ABSTRAK

Cuaca adalah kondisi atmosfer yang dipengaruhi oleh berbagai faktor seperti suhu, tekanan, dan kelembapan. Untuk memastikan bahwa aktivitas manusia tidak terganggu untuk berbagai aspek kehidupan, termasuk pertanian dan transportasi, sangat penting untuk membuat prediksi cuaca yang akurat. Di Kota Semarang, prediksi cuaca membantu masyarakat dan pemerintah dalam pengambilan keputusan. Namun, ketepatan prediksi sering terhambat oleh banyaknya parameter yang terlibat. Tujuan dari penelitian ini untuk menganalisis klasifikasi curah hujan di Kota Semarang menggunakan algoritma *Decision Tree* C4.5 berbasis data cuaca dari BMKG periode 2016-2023. Algoritma C4.5 dipilih sebab kapasitasnya membentuk pohon keputusan yang bisa memberikan visualisasi hasil klasifikasi dengan akurasi tinggi. Data mining dilakukan menggunakan RapidMiner dengan membagi data ke dalam data training dan data testing yang memiliki rasio 90:10. Hasil studi menunjukkan tingkat akurasi sebesar 99,12%, yang ada pada klasifikasi sangat baik (*Excellent Classification*).

**Kata Kunci:** Algoritma C4.5, Data Mining, *Decision Tree*, Prediksi Cuaca, Semarang

## ABSTRACT

Weather is an atmospheric condition that is influenced by many factors such as temperature, pressure, and humidity. To ensure that human activities are not disrupted for various aspects of life, including agriculture and transportation, it is very important to make accurate weather predictions. In Semarang City, weather predictions help the community and government in decision making. However, the accuracy of predictions is often hampered by the many parameters involved. The purpose of this study was to analyze the classification of rainfall in Semarang City using the *Decision Tree* C4.5 algorithm based on weather data from BMKG for the period 2016-2023. The C4.5 algorithm was chosen because of its capacity to form a decision tree that can provide visualization of classification results with high accuracy. Data mining was carried out using RapidMiner by dividing the data into training data and testing data which had a ratio of 90:10. The results of the study showed an accuracy level of 99,12%, which is in the very good classification (*Excellent Classification*).

**Keywords:** C4.5 Algorithm, Data Mining, *Decision Tree*, Semarang, Weather Prediction

## I. PENDAHULUAN

Cuaca mengacu pada kondisi atmosfer di lokasi tertentu, yang dipengaruhi oleh suhu, tekanan udara, angin, awan, kelembapan, dan curah hujan [1]. Konferensi Iklim Dunia mendefinisikan cuaca sebagai keadaan atmosfer secara menyeluruh, meliputi variasi, perkembangan, dan hambatan terhadap suatu fenomena [2]. Banyak aspek kehidupan yang dipengaruhi oleh cuaca, seperti transportasi, industri, pertanian dan pariwisata. Oleh karena itu, memprediksi cuaca dengan akurat sangat penting dan dalam memprakirakan cuaca menjadi tahapan antisipasi dalam meminimalisir efek yang mungkin berlangsung. Supaya aktivitas manusia tidak terganggu, prediksi cuaca harus sangat akurat, hal ini berlaku untuk pertanian, perkebunan, dan penerbangan yang bergantung pada kondisi cuaca untuk menjalankan kegiatan mereka dengan lancar. Sistem informasi yang melacak keadaan alam di masa mendatang termasuk prakiraan cuaca dan iklim [3]. Pada dasarnya, sistem informasi berfungsi sebagai cara untuk memaksimalkan pemantauan dan dapat digunakan untuk evaluasi atau klasifikasi dalam memprediksi cuaca atau iklim [4].

Di Kota Semarang, prediksi cuaca yang akurat sangat penting dalam hal perencanaan dan pengambilan keputusan. Prediksi cuaca yang akurat dapat membantu masyarakat dan pemerintah kota dalam mengambil keputusan yang lebih baik. Ketepatan dan kecepatan prediksi cuaca kurang terpenuhi karena banyaknya parameter yang membentuknya. Cara yang efektif yaitu dengan menggunakan

algoritma klasifikasi. Di antara berbagai algoritma klasifikasi, peneliti memilih algoritma C4.5 atau *Decision Tree*, karena yang paling populer dan efektif. Para ilmuwan dapat meneliti iklim Kota Semarang dari tahun 2016 hingga 2023 menggunakan teknik data mining, yakni praktik menemukan dan mengekstrak informasi berharga dari basis data besar melalui teknik matematika, statistik, kecerdasan buatan, dan *machine learning* [5].

Evaluasi *dataset* yang digunakan oleh peneliti yaitu menggunakan *software* RapidMiner. Data mining data, *text mining*, dan prediksi adalah semua bidang yang memanfaatkan program RapidMiner *open source*. Algoritma *Decision Tree* C4.5 menggunakan analisis RapidMiner, yang menghasilkan pemrosesan dan visualisasi yang cukup akurat. *Decision Tree* C4.5 adalah salah satu algoritma yang dapat digunakan untuk membuat pohon Keputusan [6]. Berdasarkan *dataset* cuaca di Semarang pada tahun 2016-2023, algoritma C4.5 dapat membantu peneliti mendapatkan informasi yang mudah dipahami, dianalisis, dan diinterpretasikan.

Ada berbagai studi terdahulu dengan subjek yang sama, contohnya pada penelitian Ubaidillah et al., (2023) [7] yang berjudul “Implementasi *Data Mining* Pada *Dataset* Prakiraan Cuaca Menggunakan Algoritma C4.5”, mendapatkan hasil uji akurasi sebesar 81.94%, dimana termasuk pada *good classification* yaitu pada kisaran 0.80-0.90. Selanjutnya pada penelitian Prasetyo (2024) [8] yang berjudul “Pemanfaatan Algoritma *Decision Tree* C4.5 dalam Memprakirakan Hujan di Stasiun Meteorologi Kelas I Sultan Hasanuddin”, dengan mendapatkan hasil akurasi prakiraan periode 2023 yakni sejumlah 84% dan akurasi ini bisa dianggap sangat baik. Dari beberapa penelitian terdahulu yang sudah dilakukan dapat dikatakan algoritma C4.5 cukup baik dalam mengklasifikasi sebuah data.

Berdasarkan penelitian dan penjabaran tersebut, tujuan dari studi ini yaitu dalam memprediksi cuaca di Kota Semarang pada tahun berikutnya menggunakan algoritma C4.5, dan hasil analisis dapat berguna dalam menghadapi berbagai tantangan cuaca dan meminimalisir dampak negatif yang mungkin terjadi. Data untuk penelitian ini berasal dari data cuaca harian dari BMKG Kota Semarang.

## II. METODE PENELITIAN

### 2.1 Sumber Data

*Dataset* yang digunakan yaitu berasal dari data cuaca harian dari BMKG Kota Semarang.

### 2.2 Variabel

Variabel yang diterapkan dalam penelitian ini, meliputi:

- a. Variabel dependen (Y) adalah variabel yang terpengaruh oleh variabel lainnya. Dalam konteks penelitian ini, variabel tersebut yaitu status curah hujan yang diklasifikasikan menjadi 2 kelompok, yaitu hujan dan tidak hujan.
- b. Variable independent (X) adalah variabel yang berdampak pada variabel lainnya. Dalam penelitian ini, terdapat empat variabel, yakni kelembapan rata-rata, temperatur rata-rata, kecepatan angin rata-rata, dan lama penyinaran rata-rata.

### 2.3 Tahap Analisis Data

- Seleksi Data

Data yang diperoleh sebelum memulai proses, kemudian disimpan dalam file yang terpisah. Pada tahap ini dilakukan pemilihan data kelembapan rata-rata, temperatur rata-rata, kecepatan angin rata-rata, dan waktu penyinaran rata-rata yang diperlukan untuk pengolahan data.

- *Preprocessing*

Data yang dikumpulkan adalah penaksiran curah hujan. Data yang didapat di *cleaning* untuk menemukan data duplikat, data yang hilang dan *outlier*.

- Transformasi

Setelah data dibersihkan, data diubah sesuai dengan jenisnya dan dibagi menjadi kategori variabel target dan *predicator*. Perhatikan tabel dibawah ini.

**Table 1.** Atribut dan tipe data

No	Atribut	Tipe data	Keterangan
1	Tanggal	<i>Polynomial</i>	Tanggal, bulan, dan tahun
2	Temperatur	<i>Integer</i>	Ukuran tingkat temperatur dalam satuan <i>Celsius</i>
3	Kelembapan	<i>Integer</i>	Ukuran tingkat kelembapan rata-rata dalam satuan persen
4	Lama penyinaran	<i>Integer</i>	Ukuran tingkat lama penyinaran dalam satuan jam
5	Kecepatan angin	<i>Integer</i>	Pengukuran kecepatan angin rata-rata dalam satuan m/s
6	Curah hujan	<i>Integer</i>	Ukuran tingkat curah hujan dalam satuan mm

- *Data Mining*

Pada fase ini, dilakukan tahapan dalam menentukan metode data mining yang cocok menggunakan *Decision Tree C4.5*.

- Evaluasi

Tahapan ini dilakukan dalam menguji hasil estimasi dan memilih metode untuk melakukan klasifikasi data awal.

### 2.4 Algoritma C4.5

Menurut Nasrullah et al., (2024) [9] salah satu metode untuk membuat pohon keputusan adalah algoritma C4.5. Pohon keputusan merupakan alat yang mapan dan populer untuk membuat prediksi dan mengklasifikasikan data. Saat mempelajari data, pohon keputusan berguna untuk mengungkap korelasi tersembunyi antara kualitas atau variabel. Metode dapat dibuat dengan menyediakan sejumlah besar data pelatihan untuk menentukan nilai variabel target. *J. Ross Quinlan* menyempurnakan algoritma C4.5 dengan metode ID3-nya. Pohon keputusan dapat dibuat menggunakan algoritma ID3 [10]. Ada beberapa langkah yang terlibat dalam membuat pohon keputusan menggunakan algoritma C4.5, yaitu:

- Menyiapkan data *training*, biasanya diambil dari riwayat sebelumnya dan sekarang diorganisasikan ke dalam kategori tertentu [11].
- Menghitung akar pohon. Perhitungan nilai *gain* untuk setiap atribut akan digunakan untuk menentukan akar atribut yang akan dipilih. Akar dengan nilai *gain* tertinggi akan diambil terlebih dahulu. Tentukan nilai *entropy* sebelum beralih ke nilai *gain* setiap atribut [12].

Dalam mengevaluasi skor *entropy* bisa diterapkan rumus:

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

Keterangan:

- |                      |                                  |
|----------------------|----------------------------------|
| S : Himpunan Kasus   | A : Fitur                        |
| N : Jumlah partisi S | Pi : Proporsi dari Si terhadap S |

Setelah skor *entropy* tiap atribut didapatkan, bida ditentukan skor *gian* melalui rumus:

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

Keterangan:

- |                              |                                      |
|------------------------------|--------------------------------------|
| S : Himpunan Kasus           | A : Fitur                            |
| n : Jumlah partisi atribut A | Si  : Jumlah kasus pada partisi ke-i |
| S  : Jumlah kasus dalam S    |                                      |

## 2.5 Evaluasi Tools

Model evaluasi yang digunakan oleh peneliti yaitu *Confusion Matrix* yang menghitung nilai akurasi validasi algoritma terhadap dataset yang ada. Nilai akurasi dihitung dengan tahapan penilaian berikut:

**Tabel 2.** *Confusion Matrix*

Aktual	Prediksi	
	Positif	Negatif
Positif	<i>True Positive</i>	<i>False Positive</i>
	<i>False Positive</i>	<i>True Positive</i>
Negatif	<i>False Negative</i>	<i>True Negative</i>

Tingkat akurasi klasifikasi yaitu *Area Under Receiver Operating Characteristics (ROC)*, *Area Under curve (AUC)*, akurasi dan kesalahan. Untuk mengukur perbedaan kinerja menggunakan perkiraan AUC (*area under curve*). Nilai AUC pada *data mining* bisa diklasifikasikan ke dalam lima kelompok menurut Gorunescu (2011) [13]:

- Skor 0.90 - 1.00 menunjukkan *Excellent Classification*
- Skor 0.80 - 0.90 menunjukkan *Good Classification*
- Skor 0.70 - 0.80 menunjukkan *Fair Classification*
- Skor 0.60 - 0.70 menunjukkan *Poor Classification*
- Skor 0.50 - 0.60 menunjukkan *Error*

$$Akurasi = \frac{TP+TN}{TP+TN+FN+FP} \quad (3)$$

$$Error = \frac{FN+FP}{TP+TN+FN+FP} \quad (4)$$

Keterangan:

FP = *False Positive*

TP = *True Positive*

TN = *True Negative*

FN = *False Negative*

## 2.6 Rapid Miner

Rapid Miner adalah *tools data mining* yang dibuat menggunakan Java. Bahasa Java dianggap sebagai salah satu bahasa terbaik karena kesederhanaan, keamanannya, kekuatan, pengaruh, dan kemampuan berorientasi objek yang luar biasa [14].

## III. HASIL DAN PEMBAHASAN

### 3.1 Seleksi Data

Informasi yang dimanfaatkan dikumpulkan dari BMKG Semarang, dari tahun 2016 hingga tahun 2023. *Dataset* ini diakses melalui alamat *website* <https://dataonline.bmkg.go.id/> [15].

### 3.2 Preprocessing Data

Pada tahap ini dataset yang telah didapatkan dilakukan pengecekan dan ternyata terdapat data yang *missing value* sehingga harus dilakukan *cleaning* data yang hilang atau ganda. Berikut adalah tabel data yang telah dilakukan *cleaning*.

### 3.3 Transformasi Data

Dataset yang telah dilakukan *cleaning*, selanjutnya diklasifikasikan sesuai dengan variabel curah hujan yang dikategorikan menjadi 2 kelompok, yaitu hujan dan tidak hujan berdasarkan nilai ambang BMKG. Dataset yang akan diuji diimpor dan dilakukan perubahan pada atribut yang bersifat angka dengan jenis data yaitu Integer dan untuk jenis data kolom predicator yaitu label atau sebagai target. Berikut adalah tabel yang sudah dilakukan perubahan.

Tabel 3. Kategori Hujan

NO	Kategori Hujan	Intensitas Curah Hujan (mm)
1	Tidak Hujan	0
2	Hujan	$\geq 0.5$

Tanggal	Temperatur	Kelembapan	Lama Penyinaran	Kecepatan Angin	Curah Hujan
03-01-2016	28,5	79	3,1	2	0
04-01-2016	28,4	78	4,7	2	68,7
05-01-2016	27,9	78	6	3	0
08-01-2016	30,2	72	11,2	3	0
09-01-2016	29	76	9,8	2	24,2
10-01-2016	29,1	76	10,3	2	0
11-01-2016	28,6	79	8,8	2	2,5
14-01-2016	30,2	69	7,3	3	0
15-01-2016	31	66	8	2	0
17-01-2016	29,3	76	7,7	2	0
18-01-2016	28,7	79	7,5	2	23
19-01-2016	29,5	75	4,6	3	1,5
20-01-2016	27,2	87	8,5	2	29
24-01-2016	26,9	85	4,2	2	0
25-01-2016	28,3	77	0	4	41,9
27-01-2016	27,2	85	4,5	2	6,8
			.....		
29-12-2023	29,3	77	7	2	16,7
30-12-2023	29,2	76	9,3	2	114,4
31-12-2023	28,6	80	7,2	2	0

Gambar 3. Data setelah dilakukan *cleaning*

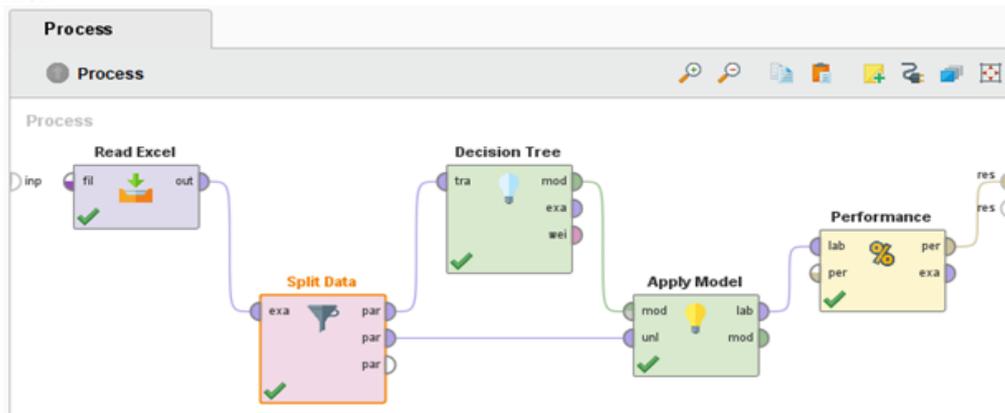
Row No.	Target	Tanggal	Temperatur	Kelembapan	Lama Peny...	Kecepatan ...	Curah Hujan
1	Tidak hujan	03-01-2016	29	79	3	2	0
2	Hujan	04-01-2016	28	78	5	2	69
3	Tidak hujan	05-01-2016	28	78	6	3	0
4	Tidak hujan	08-01-2016	30	72	11	3	0
5	Hujan	09-01-2016	29	76	10	2	24
6	Tidak hujan	10-01-2016	29	76	10	2	0
7	Hujan	11-01-2016	29	79	9	2	3
8	Tidak hujan	14-01-2016	30	69	7	3	0
9	Tidak hujan	15-01-2016	31	66	8	2	0
10	Tidak hujan	17-01-2016	29	76	8	2	0
11	Hujan	18-01-2016	29	79	8	2	23
12	Hujan	19-01-2016	30	75	5	3	2
13	Hujan	20-01-2016	27	87	9	2	29
14	Tidak hujan	24-01-2016	27	85	4	2	0
15	Hujan	25-01-2016	28	77	0	4	42

ExampleSet (2,275 examples, 1 special attribute, 6 regular attributes)

Gambar 4. Data sudah diklasifikasikan dan dilakukan perubahan jenis data

### 3.4 Data Mining

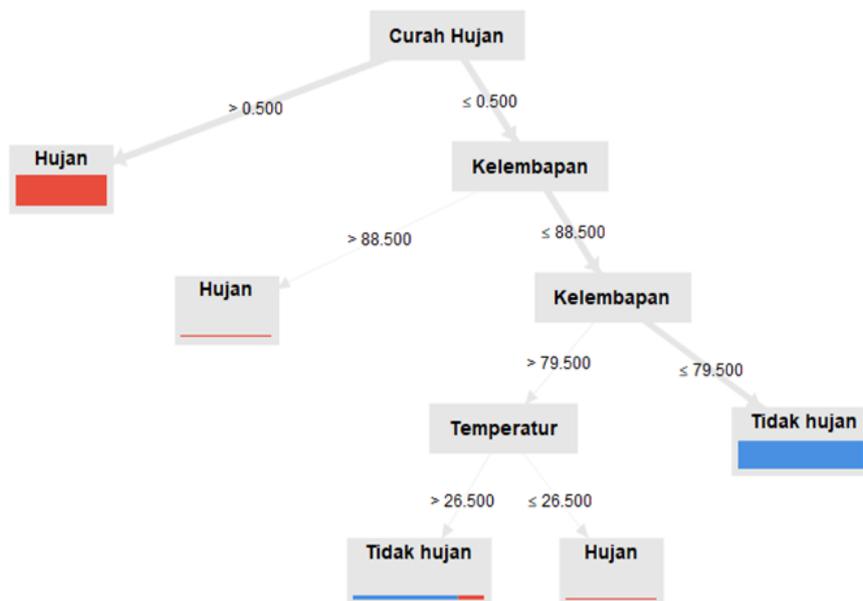
Data yang telah diimpor selanjutnya pembuatan pemodelan untuk klasifikasi menggunakan *Decision Tree* (C4.5) melalui *software* RapidMiner. Data yang akan diolah yaitu sebanyak 2275 dataset. Data yang telah dilakukan perubahan, perhatikan gambar 4 diatas. Peneliti telah melakukan pengujian dari perbandingan 50:50, 60:40, 70:30, 80:20, dan 90:10, peneliti memutuskan untuk menggunakan perbandingan 90:10. Hal ini dikarenakan perbandingan tersebut menghasilkan akurasi yang tertinggi dan tingkat kesalahan yang rendah, yaitu dengan 2.048 data latih dan 228 data uji. Kemudian dilakukan *cross validation* akurasi dan *error* pada proses klasifikasi. Berikut adalah gambar *modelling* pada RapidMiner.



Gambar 5. *Modelling* RapidMiner dengan Algoritma *Decision Tree*

### 3.5 Hasil Perhitungan Klasifikasi dengan RapidMiner

Setelah dilakukan pengujian dataset, algoritma C4.5 yang termasuk *decision tree*, menghasilkan sebuah pohon keputusan yang bisa diamati di gambar 6.



Gambar 6. Hasil Pohon Keputusan

Gambar 6 dapat diketahui terdapat 3 atribut yang diterapkan dalam merancang pohon keputusan. Hasil pohon keputusan bisa dikatakan bahwa suatu curah hujan tertentu dengan dan temperatur dapat diketahui cuaca pada suatu hari tertentu. Berdasarkan pada gambar 6 menghasilkan beberapa rule dari pohon keputusan, dapat dilihat dari tabel berikut ini.

**Tabel 4.** Aturan/ Rule

No	Jika	Maka
1	Curah hujan > 0.500	Hujan
2	Curah Hujan ≤ 0.500 dan Kelembapan > 88.500	Hujan
3	Curah Hujan ≤ 0.500, Kelembapan ≤ 88.500, dan Kelembapan > 79.500, serta Temperatur > 26.500	Tidak Hujan
4	Curah Hujan ≤ 0.500, Kelembapan ≤ 88.500, dan Kelembapan > 79.500, serta Temperatur ≤ 26.500	Hujan
5	Curah Hujan ≤ 0.500 dan Kelembapan ≤ 79.500	Tidak Hujan

Setelah dilakukan modelling pada Rapid Miner, seperti pada gambar 5 diatas, dengan metode *data mining* yang dipakai yaitu *Decision Tree / C4.5*. Gambar berikut menunjukkan hasil akurasi data.

accuracy: 99.12%

	true Tidak hujan	true Hujan	class precision
pred. Tidak hujan	107	2	98.17%
pred. Hujan	0	118	100.00%
class recall	100.00%	98.33%	

**Gambar 7.** Hasil Akurasi

- Data awal tidak hujan dan diperkirakan tidak hujan yaitu 107
- Data awal hujan dan diperkirakan hujan yaitu 118
- Data awal tidak hujan dan diperkirakan hujan yaitu 0
- Data awal hujan dan diperkirakan tidak hujan yaitu 2

Evaluasi metode menggunakan nilai akurasi *error*:

$$Akurasi = \frac{107 + 118}{107 + 118 + 0 + 2} = 99,12\%$$

$$Error = \frac{0 + 2}{107 + 118 + 0 + 2} = 0,88\%$$

Berdasarkan evaluasi akurasi *error* dari eksekusi data pelatihan dan data pengujian dievaluasi melalui pohon keputusan menciptakan performa akurasi sebesar 99,12% dengan presisi prediksi tidak hujan sebesar 98,17%, presisi prediksi hujan sebesar 100%. *Class recall* yang diciptakan dari kinerja akurasi yaitu prediksi true tidak hujan sebesar 100% dan prediksi *true* hujan 98,33%. Performa akurasi yang dihasilkan pada penelitian ini berada pada klasifikasi *Excellent Classification* dengan nilai akurasi berkisar antara 0.90 - 1.00.

#### IV. KESIMPULAN

Menurut temuan dari studi ini, *Decision Tree C4.5* dapat mengklasifikasikan data cuaca di Kota Semarang dengan akurasi tinggi, mencapai 99,12%. Dengan memanfaatkan atribut seperti curah hujan, temperatur, lama penyinaran, kecepatan angin, dan kelembapan, model ini memberikan hasil klasifikasi yang andal. Temuan ini menjabarkan bahwa algoritma C4.5 efektif diterapkan dalam melangsungkan prediksi cuaca dan dapat diterapkan pada skenario lain seperti pembuatan aplikasi prediksi cuaca kota Semarang, membantu perencanaan dan mitigasi risiko terkait cuaca. Model ini diharapkan dapat menjadi alat pendukung yang bermanfaat bagi masyarakat dan pemerintah daerah dalam pengambilan keputusan terkait cuaca.

## DAFTAR PUSTAKA

- [1] Sugiyanti, S. S. Siregar, and A. E. Fahrudin, “Aplikasi Jaringan Syaraf Tiruan (JST) Backpropagation untuk Prediksi Cuaca Harian di Wilayah Banjarbaru,” vol. 9, no. 2, 2012.
- [2] A. Novandya and I. Oktria, “Penerapan Algoritma Klasifikasi Data Mining C4.5 Pada Dataset Cuaca Wilayah Bekasi,” vol. 6, 2017.
- [3] A. N. Sallamah, A. Sunandar, A. M. Rizka, and C. Rozikin, “Prakiraan Cuaca Kota Bandung di Tahun Berikutnya dengan Menggunakan Metode Naïve Bayes,” *J. Comput. Inf. Technol.*, vol. 8, no. 1, pp. 27–32, 2024.
- [4] A. M. Siregar, S. Faisal, Y. Cahyana, and B. Priyatna, “Perbandingan Algoritme Klasifikasi Untuk Prediksi Cuaca,” *J. Account. Inf. Syst. AIMS*, vol. 3, no. 1, pp. 15–24, Apr. 2020, doi: 10.32627/aims.v3i1.280.
- [5] F. I. Harida and N. Khazizah, “ANALISIS CUACA DI KOTA JAKARTA BULAN JANUARI TAHUN 2018 MENGGUNAKAN ALGORITMA DECISION TREE,” vol. 14, no. 1, 2022.
- [6] P. Bimo Nugroho Setio, D. Retno Sari Saputro, and B. Winarno, “Klasifikasi dengan Pohon Keputusan Berbasis Algoritme C4.5,” *PRISMA Pros. Semin. Nas. Mat.*, vol. 3, pp. 64–71, 2020, doi: <https://journal.unnes.ac.id/sju/index.php/prisma/>.
- [7] F. Ubaidillah, M. R. Ramadhan, and P. Rosyani, “IMPLEMENTASI DATA MINING PADA DATASET PRAKIRAAN CUACA MENGGUNAKAN ALGORITMA C4.5,” *ALKHAWARIZMI J. Mat. Algoritma Dan Sains*, vol. 1, no. 1, pp. 131–140, 2023.
- [8] A. Prasetyo, “Pemanfaatan Algoritma Decision Tree C4.5 dalam Memprakirakan Hujan di Stasiun Meteorologi Kelas I Sultan Hasanuddin,” *Bul. GAW Bariri*, vol. 5, no. 1, pp. 47–55, Jun. 2024, doi: 10.31172/bgb.v5i1.120.
- [9] M. F. Nasrullah, Rd. R. Saedudin, and F. Hamami, “PERBANDINGAN AKURASI ALGORITMA C4.5 DAN K-NEAREST NEIGHBORS UNTUK KLASIFIKASI CURAH HUJAN BERDASARKAN IKLIM INDONESIA,” *JUPI J. Ilm. Penelit. Dan Pembelajaran Inform.*, vol. 9, no. 2, pp. 628–638, May 2024, doi: 10.29100/jipi.v9i2.4655.
- [10] J. R. Quinlan, *C4.5: Programs for Machine Learning*. San Mateo, 1993.
- [11] M. W. Prihatmono and A. F. Watratan, “IMPLEMENTASI ALGORITMA C4.5 MENGGUNAKAN PYTHON UNTUK KLASIFIKASI KEPUASAN KONSUMEN,” *J. Inform. Progres*, vol. 11, no. 2, pp. 49–55, Sep. 2019.
- [12] A. Shiddiq, R. K. Niswatin, and I. N. Farida, “Analisa Kepuasan Konsumen Menggunakan Klasifikasi Decision Tree Di Restoran Dapur Solo (Cabang Kediri),” *Gener. J.*, vol. 2, no. 1, 2018.
- [13] F. Gorunescu, *Data Mining: Concepts, models and techniques*, vol. 12. Springer Science & Business Media, 2011.
- [14] M. Al-Batah, B. Zaqaibeh, S. A. Alomari, and M. S. Alzboon, “Gene Microarray Cancer Classification using Correlation Based Feature Selection Algorithm and Rules Classifiers,” *Int. J. Online Biomed. Eng. IJOE*, vol. 15, no. 8, 2019.
- [15] “BMKG Dataonline,” BMKG Dataonline. Accessed: Jan. 03, 2025. [Online]. Available: <https://dataonline.bmkg.go.id>