

# Implementasi Algoritma C4.5 Untuk Klasifikasi Data Insentif Karyawan PT. Adhi Cakra Utama Mulia

Putri Widia Sari<sup>1\*</sup>, Gilang Giwangga<sup>2</sup>, Novi Dian Fitriyanti<sup>3</sup>, Advinca Arlisdia Putra<sup>4</sup>, Hari Irawan<sup>5</sup>

Universitas Bina Sarana Informatika  
Jl. Kramat Raya No 98, Jakarta Pusat, Indonesia

e-mail korespondensi: putrwidia0414@gmail.com

Submit: 21-06-2024 | Revisi: 03-07-2024 | Terima: 05-07-2024 | Terbit online: 08-07-2024

**Abstrak** - Penentuan insentif karyawan yang akurat dan adil merupakan salah satu tantangan utama dalam manajemen sumber daya manusia. Penelitian ini bertujuan untuk mengimplementasikan algoritma C4.5 dalam proses klasifikasi data insentif karyawan guna meningkatkan ketepatan dan keadilan dalam penentuan insentif. Algoritma C4.5 dipilih karena kemampuannya dalam menangani data kontinu dan kategorik serta kemampuannya untuk menghasilkan pohon keputusan yang mudah dipahami. Data penelitian diperoleh dari PT Adhi Cakra Utama Mulia, yang mencakup atribut seperti nama karyawan, wilayah, nama toko, minimal penjualan, pencapaian dan keputusan. Proses penelitian dimulai dengan pengumpulan dan pra-pemrosesan data, yang meliputi pembersihan data dan normalisasi. Selanjutnya, algoritma C4.5 diterapkan untuk membangun model klasifikasi berdasarkan data yang telah diproses. Hasil implementasi menunjukkan bahwa algoritma C4.5 mampu mengklasifikasikan data insentif karyawan dengan tingkat akurasi yang tinggi. Selain itu, pohon keputusan yang dihasilkan memberikan wawasan yang jelas mengenai faktor-faktor yang paling berpengaruh dalam penentuan insentif. Penelitian ini diharapkan dapat membantu perusahaan dalam mengambil keputusan yang lebih objektif dan transparan terkait pemberian insentif kepada karyawan. Berdasarkan penelitian yang di lakukan dapat disimpulkan bahwa data mining yang digunakan dengan metode algoritma C45 dapat diterapkan dengan tingkat *accuracy* = 77.73%, dengan nilai entropy 0.829762 dan gain information terbesar 0.038638 sedangkan untuk recall = 35.67% dan precision = 61.76%.

Kata Kunci : Algoritma C4.5, Klasifikasi Data, Penentuan Insentif

**Abstract** - *Determining accurate and fair employee incentives is one of the main challenges in human resource management. This research aims to implement the C4.5 algorithm in the employee incentive data classification process in order to increase accuracy and fairness in determining incentives. The C4.5 algorithm was chosen because of its ability to handle continuous and categorical data and its ability to produce decision trees that are easy to understand. Research data was obtained from PT Adhi Cakra Utama Mulia, which includes attributes such as employee name, region, shop name, minimum sales, achievements and decisions. The research process begins with data collection and pre-processing, which includes data cleaning and normalization. Next, the C4.5 algorithm is applied to build a classification model based on the processed data. The implementation results show that the C4.5 algorithm is able to classify employee incentive data with a high level of accuracy. In addition, the resulting decision tree provides clear insight into the most influential factors in determining incentives. It is hoped that this research can help companies make more objective and transparent decisions regarding providing incentives to employees. Based on the research conducted, it can be concluded that data mining used with the C45 algorithm method can be applied with an accuracy level = 77.73%, with an entropy value of 0.82962 and the largest information gain of 0.038638, while for recall = 35.67% and precision = 61.76%.*

Keywords: C4.5 Algorithm, Data Classification, Determining Incentives

## 1. Pendahuluan

Insentif adalah penghargaan berupa uang yang diberikan kepada orang yang dapat bekerja melebihi standar yang telah ditentukan [1]. Pemberian insentif bertujuan untuk memacu semangat kerja karyawan dalam mencapai prestasi dan memotivasi karyawan untuk memberikan kontribusi lebih kepada perusahaan [2]. Sistem Pendukung Keputusan (SPK) menjadi suatu sistem yang selalu dibutuhkan bagi perusahaan dalam menetapkan suatu keputusan, salah satunya dalam menentukan bonus tahunan bagi [3]. Pemberian insentif merupakan salah satu hal pokok yang harus diperhatikan oleh perusahaan. Semangat tidaknya karyawan bisa juga disebabkan oleh besar kecilnya insentif yang diterima. Apabila karyawan tidak mendapatkan insentif yang sesuai dengan besarnya



pengorbanan dalam bekerja, maka karyawan tersebut cenderung malas bekerja dan tidak bersemangat yang ada akhirnya mereka bekerja semaunya tanpa ada motivasi yang tinggi. Dengan adanya pemberian insentif yang tepat serta cara kerja yang baik, ke depannya, proses kerja organisasi dapat berjalan sesuai tujuan organisasi. Kinerja karyawan sangat erat kaitannya dengan gaji dan insentif yang mereka terima, karena dapat berdampak positif maupun negatif terhadap semangat kerjanya yang pada akhirnya akan meningkatkan kinerja karyawan dalam suatu organisasi [4].

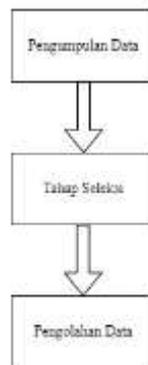
Dalam konteks pemberian insentif karyawan, penelitian ini menggunakan algoritma C4.5 untuk membuat keputusan yang tepat. Algoritma C4.5 merupakan salah satu solusi pemecahan kasus yang sering digunakan untuk membuat pohon keputusan dalam pemecahan masalah pada teknik klasifikasi yang memiliki karakteristik yaitu dengan 3 proses penentuan nilai entropy dan nilai gain. Suatu riset memaparkan bahwa Algoritma C4.5 ialah metode klasifikasi dengan memakai entropy serta keuntungan data selaku pemisah pada pohon keputusan [5].

Adapula yang berpendapat bahwa Algoritma C4.5 adalah sebuah metode klasifikasi yang digunakan untuk membentuk sebuah model prediksi berdasarkan data yang telah diorganisir ke dalam decision tree [6]. Decision Tree adalah struktur pohon di mana node pohon mewakili atribut yang diuji. Setiap cabang dari pohon ini merupakan distribusi hasil pengujian, dan simpul-simpul tersebut mewakili beberapa kelompok kelas [7]. Metode Decision Tree adalah bagian dari metode machine learning yang menghasilkan klasifikasi berupa pohon keputusan berdasarkan dataset yang diujikan. Proses klasifikasi dengan metode Decision Tree diawali dengan membaca dataset, kemudian proses preprocessing dan apply model [8].

Machine Learning (ML) merupakan bagian dari ilmu kecerdasan Buatan atau yang dikenal sebagai Artificial Intelligent (AI) [9]. Konsep pohon keputusan adalah mengubah data menjadi pohon keputusan dan aturan keputusan. Manfaat utama menggunakan pohon keputusan adalah kemampuan untuk menyederhanakan proses pengambilan keputusan yang kompleks sehingga pengambil keputusan dapat menafsirkan solusi untuk masalah [10].

Penelitian ini bertujuan untuk mengimplementasikan algoritma C4.5 dalam klasifikasi data insentif karyawan PT. Adhi Cakra Utama Mulia. Dengan memanfaatkan data historis karyawan, penelitian ini akan menganalisis faktor-faktor yang berpengaruh terhadap pemberian insentif dan mengembangkan model klasifikasi yang dapat diandalkan.

## 2. Metode Penelitian



Gambar 1. Alur Penelitian

### 2.1 Pengumpulan Data

Metode pengumpulan data yang digunakan oleh penulis memiliki tujuan untuk mengumpulkan data yang akurat dan terstruktur. Data yang digunakan adalah data perhitungan insentif karyawan di PT Adhi Cakra Utama Mulia. Jumlah data yang digunakan terdiri dari 11 file yang sudah diambil poin utamanya menjadi 229 data yang berisi beberapa atribut seperti nama karyawan, wilayah, nama toko, jumlah minimal, pencapaian dan target. Dalam proses pengambilan poin utama dilakukan secara manual di MS. Excel.

### 2.2 Tahap seleksi

Pada tahap seleksi ini melibatkan data cleaning yang dimana tugas utama data cleaning adalah menangani missing value. Data yang telah dikumpulkan tadi diseleksi untuk mencari missing value. Setelah missing value ditangani format data terpilih diubah kedalam format yang sesuai dengan prosedur data mining agar bisa dilakukan proses selanjutnya. Dalam proses ini penulis menggunakan software RapidMiner.

Tabel 1. Atribut Data Terpilih

No	Atribut Terpilih	Keterangan
1.	Nama Karyawan	Atribut Fitur
2.	Wilayah	Atribut Fitur

No	Atribut Terpilih	Keterangan
3.	Nama Toko	Atribut Fitur
4.	Minimal	Atribut Fitur
5.	Pencapaian	Atribut Fitur
6.	Keterangan Target	Atribut Label

Atribut terpilih akan diolah menggunakan algoritma C4.5 dan *software* RapidMiner dan akan menghasilkan pohon keputusan.

### 2.3 Pengolahan Data

Pada tahap ini data yang telah diseleksi sesuai dengan keperluan penelitian kemudian di proses dengan cara di impor ke aplikasi rapidminer. Data tersebut kemudian diolah menggunakan algoritma C.4 5 untuk menghasilkan pohon keputusan.

Tabel 2. Dataset Insentif Karyawan

Nama Karyawan	Wilayah	Nama Toko	Minimal	Pencapaian	Target
Kevin	Surabaya	Hanura	654	810	Y
Adit	Jakarta	Tunas Pinang	3500	3960	Y
Marjuki	Bogor	Citra Bangunan	175	120	N
Dendi	Cikarang	Usaha Mandiri	1000	1656	Y
Novi	Tangerang	Mahardika	900	710	N
Ardrian	Cikarang	Usaha Baru	682	908	Y
Mallorca	Makasar	Jatimulia	685	454	N
Ria	Tangerang	Murah Jaya	6000	7866	Y
Kevin	Surabaya	Mitra Jaya	580	575	N
Ridwan	Tangerang	Sinar Surya	1200	1576	Y

Keterangan :

- Nama Karyawan : Atribut yang digunakan untuk mengidentifikasi seorang karyawan.
- Wilayah : Atribut yang mengidentifikasi area penjualan.
- Nama Toko : Atribut yang mengidentifikasi nama toko dilokasi tersebut.
- Minimal : Atribut minimal penjualan per-periode dalam satuan KG (Kilo Gram).
- Pencapaian : Atribut pencapaian penjualan per-periode dalam satuan KG (Kilo Gram).
- Target : Atribut label dengan keterangan nilai Y (yes / ya mencapai target) dan N (no / tidak mencapai target).

Tabel 3. Nilai Atribut Dataset Insentif Karyawan

No	Atribut	Keterangan
1.	Nama Karyawan	Kevin, Adit, Marjuki, dst.
2.	Wilayah	Surabaya, Jakarta, Bogor, dst.
3.	Nama Toko	Hanura, Tunas Pinang, Citra Bangunan, dst.
4.	Minimal	654, 3500, 175, dst.
5.	Pencapaian	810, 3960, 120, dst.
6.	Keterangan Target	Y (yes), N (no)

## 3. Hasil dan Pembahasan

### 3.1. Analisa Masalah

Penelitian ini membahas implementasi algoritma C4.5 untuk klasifikasi data insentif karyawan di PT. Adhi Cakra Utama Mulia, yang menghadapi tantangan kompleksitas data, efisiensi waktu, dan kebutuhan akan objektivitas dalam penentuan insentif. Algoritma C4.5 digunakan untuk membangun model pohon keputusan yang dapat mengklasifikasikan data karyawan berdasarkan kinerja, kehadiran, dan kontribusi mereka. Dengan menggunakan data yang dikumpulkan dan diproses, penelitian ini mengevaluasi akurasi dan efisiensi model, bertujuan untuk meningkatkan keadilan dan konsistensi dalam proses penentuan insentif, serta mengurangi kesalahan manusia dan waktu yang dibutuhkan untuk evaluasi karyawan.

### 3.2. Penerapan Metode C4.5

Setelah dilakukan proses pengumpulan data maka selanjutnya dilakukan proses pengujian terhadap data yang digunakan pada penelitian. Pengujian bertujuan untuk mendapatkan hasil yang digunakan untuk proses pengambilan keputusan [11]. Adapun rumus untuk menghitung data yang tertera pada tabel 2 adalah

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

Keterangan:

$S$  = Himpunan kasus

$n$  = Jumlah partisi  $S$

$p_i$  = Proporsi  $S_i$  terhadap  $S$

Untuk menghitung nilai Gain dapat menggunakan rumus

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

Tabel 3. Nilai Entropy Total Seluruh Data

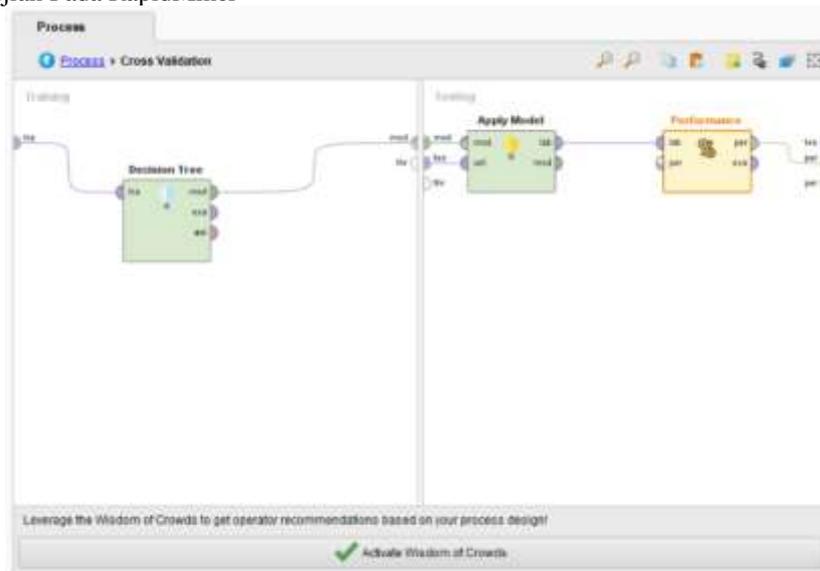
Jumlah Kasus	Main (Y)	Main (N)	Entropy
229	169	60	0.829762

Tabel 4. Nilai Entropy Untuk Masing-Masing Atribut

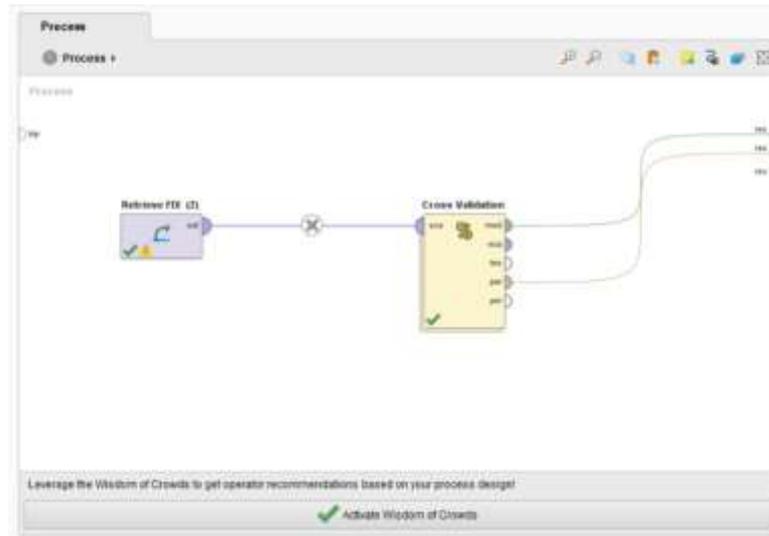
Atribut	Keterangan	Jumlah Kasus	Main (Y)	Main (N)	Entropy	Info Gain	Split Info	Gain Ratio
Nama Karyawan	Kevin	43	34	9	0.740147	0.038638	3.046518	0.012683
	Adit	41	28	13	0.90117			
	Marjuki	36	28	8	0.764205			
	Dendi	21	10	11	0.998364			
	Novi	17	14	3	0.672295			
	Ridwan	17	11	6	0.936667			
	Ardrian	23	18	5	0.755375			
	Mallorca	13	10	3	0.77935			
Wilayah	Ria	18	16	2	0.503258	0.014323	2.494881	0.005741
	Surabaya	43	34	9	0.740147			
	Jakarta	41	28	13	0.90117			
	Bogor	36	28	8	0.764205			
	Cikarang	44	28	16	0.94566			
	Tangerang	52	41	11	0.744413			
	Makasar	13	10	3	0.77935			
	Mallorca	43	34	9	0.740147			

Berdasarkan perhitungan Gain pada Tabel 4 diperoleh dari Gain tertinggi yaitu attribute Nama Karyawan, sehingga Nama Karyawan menjadi node akar awal pembentukan pohon keputusan.

### 3.3. Proses Pengujian Pada RapidMiner

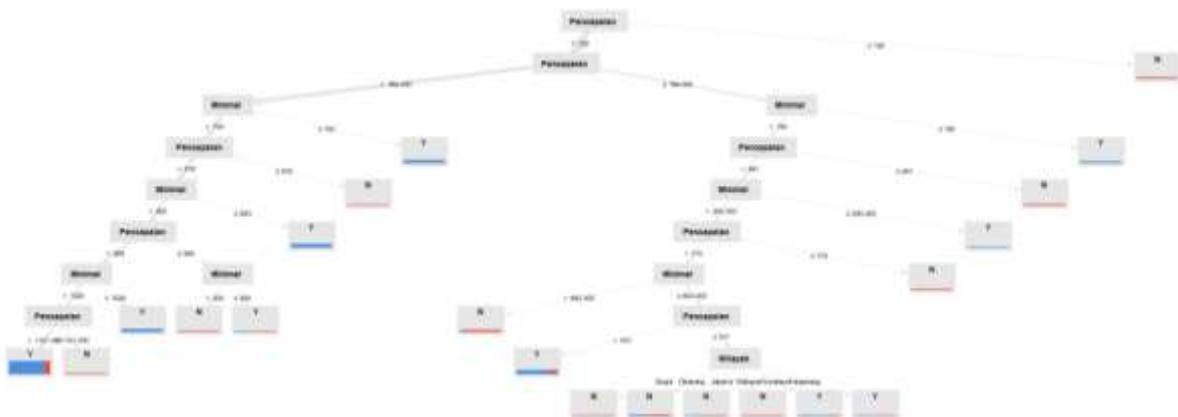


Gambar 2. Konektivitas antara data dan model pohon keputusan



Gambar 3. Mengukur Performance

Pada gambar 2 dan 3 dapat dilihat proses pengujian dari algoritma C5.5 dengan menggunakan tools rapidminer studio. Dari gambar tersebut nantinya didapatkan hasil yang digunakan pada penelitian.



Gambar 4. Pohon Keputusan (*Decision Tree*)

Table View | Plot View

accuracy: 77.73% +/- 9.11% (micro average) 77.73%

	True Y	True N	class precision
pred Y	157	28	80.51%
pred N	10	21	67.50%
class recall	82.50%	85.50%	

Gambar 5. Accuracy

Table View | Plot View

recall: 85.67% +/- 24.19% (macro average) 75.50% (positive class: N)

	True Y	True N	class precision
pred Y	157	28	80.51%
pred N	10	21	67.50%
class recall	82.50%	85.50%	

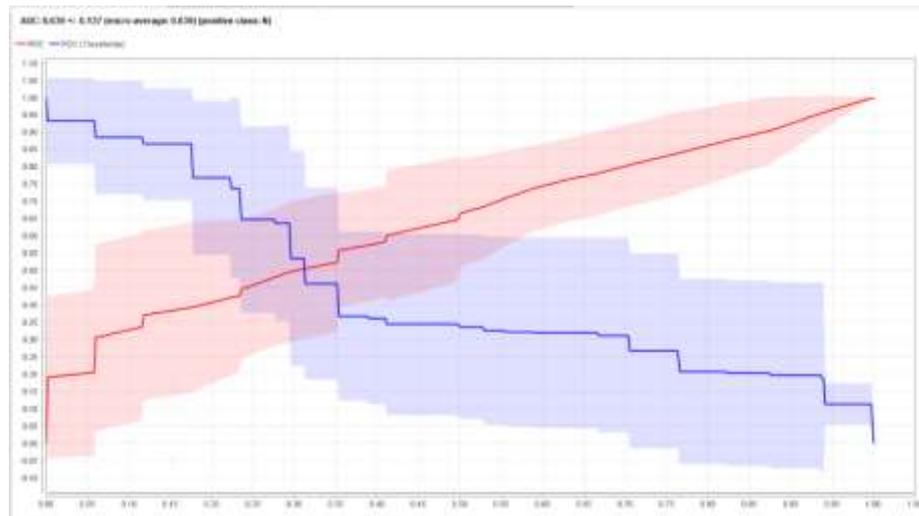
Gambar 6. Recall

Table View | Plot View

precision: 61.76% (positive class: R)

	True P	True N	class precision
pred P	107	38	85.07%
pred N	12	21	61.76%
class recall	92.30%	33.33%	

Gambar 7. Precision



Gambar 8. AUC

Pada Gambar 4, Gambar 5, Gambar 6, Gambar 7 dan Gambar 8 dapat dilihat hasil dari prediksi dari proses algoritma C4.5. Pada gambar tersebut dapat dilihat hasil pohon keputusan dan juga nilai accuracy, recall, precision dan AUC.

#### 4. Kesimpulan

Berdasarkan penelitian yang di lakukan dapat disimpulkan bahwa data mining yang digunakan dengan metode algoritma C45 dapat diterapkan dengan tingkat *accuracy* = 77.73%, dengan nilai entropy 0.829762 dan gain information terbesar 0.038638 sedangkan untuk recall = 35.67% dan precision = 61.76%. Dengan penelitian ini hasil algoritma yang didapatkan dari data yang diperoleh, dapat memberikan rekomendasi untuk pengambil keputusan

#### Referensi

- [1] A. D. Saputri, S. Handayani, dan M. K. Dp, "Pengaruh Disiplin Kerja dan Pemberian Insentif terhadap Kinerja Karyawan PT Putra Karisma Palembang," *Jurnal Nasional Manajemen Pemasaran Dan SDM*, vol. 2, no. 1, hlm. 25–42, Mar. 2021, doi: 10.47747/jnmpsdm.v2i1.211.
- [2] G. Y. A. P. Pratama, "Pengaruh Insentif dan Motivasi Terhadap Kinerja Karyawan pada PT. Aditama Bogor," *Jurnal Pajak Dan Bisnis*, vol. 2, no. 1, hlm. 59–65, Mar. 2021, doi: 10.55336/jpb.v2i1.27.
- Malaeny, C. S., Katuuk, M., & Onibala, F. (2017). Hubungan Riwayat Lama Merokok Dan Kadar Kolesterol Total Dengan Kejadian Penyakit Jantung Koroner Di Poliklinik Jantung Rsu Pancaran Kasih Gmim Manado. *Jurnal Keperawatan*, 5(1).
- [3] Y. Yulisman dan R. Wahyuni, "Sistem Pendukung Keputusan Penentuan Bonus Karyawan Dengan Metode SAW Pada PT. Delima Makmur Aceh Singkil," *JTIM: Jurnal Teknologi Informasi Dan Multimedia*, vol. 3, no. 2, hlm. 78–90, Aug. 2021, doi: 10.35746/jtim.v3i2.154.
- Novitasari, F. (2023). Sistem Klasifikasi Penyakit Jantung Menggunakan Teknik Pendekatan SMOTE Pada Algoritma Modified K-Nearest Neighbor. *Jurnal Building of Informatics, Technology and Science (BITS)*, 5(1), 274-284.
- [4] N. E. Ahmad dan R. P. Bambang, "Pengaruh Gaji, Insentif, Dan Motivasi Terhadap Kinerja Karyawan Pada Pt. Mulia Wisnu Setya Di Surabaya," 2021.

- [5] T. Widiastuti, K. Karsa, dan C. Juliane, "Evaluasi Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Akademik Menggunakan Metode Klasifikasi Algoritma C4.5," *Technomedia Journal*, vol. 7, no. 3, hlm. 364–380, Des. 2022, doi: 10.33050/tmj.v7i3.1932.
- [6] A. Nurzaman dan T. N. Suharsono, "Implementasi Metode Algoritma C4.5 Untuk Prediksi Gejala Autisme Pada Anak," *j-innovative.org*, Okt. 2023, doi: 10.31004/innovative.v3i5.4934.
- [7] F. M. Almufqi dan A. Voutama, "Perbandingan Metode Data Mining Untuk Memprediksi Prestasi Akademik Siswa," *Jurnal Teknika*, vol. 15, no. 1, hlm. 61–66, Apr. 2023, doi: 10.30736/jt.v15i1.929.
- [8] A. Supriyadi, "Perbandingan Algoritma Naive Bayes dan Decision Tree(C4.5) dalam Klasifikasi Dosen Berprestasi," *Generation Journal*, vol. 7, no. 1, hlm. 39–49, Mar. 2023, doi: 10.29407/gj.v7i1.19797.
- [9] Effendi dan R. Noviana, "Perancangan Web Sistem Analisis Sentimen Media Sosial Twitter Dengan Metode Valence Aware Dictionary And Sentimen Reasoner (Vader) Menggunakan PHP & MySQL pada Pemerintah Kota Bekasi," *Jurnal Ilmiah Komputasi*, vol. 20, no. 1, Mar. 2021, doi: 10.32409/jikstik.20.1.369.
- [10] R. N. Sari, N. Novrina, dan S. A. S, "Penerapan Algoritma C4.5 Untuk Memprediksi Bencana Gunung Meletus Di Indonesia," Mei 22, 2024. [Online]. Tersedia di: <http://journal.admi.or.id/index.php/JTS/article/view/1265>.
- [11] Y. F. Wijaya and A. Triayudi, "Penerapan Data Mining Pada Prediksi Harga Emas dengan Menggunakan Algoritma Regresi Linear Berganda dan ARIMA," *Journal of Computer System and Informatics/Journal of Computer System and Informatics (JoSYC)*, vol. 5, no. 1, pp. 73–81, Nov. 2023, doi: 10.47065/josyc.v5i1.4615.