

Penerapan Algoritma ID3 dalam Prediksi Kebutuhan Pupuk

Milyani Aritonang

Program Studi Teknik Informatika, Universitas BUDI DARMA, Medan, Indonesia

Email: milyani209@gmail.com

Abstrak—Kebutuhan pupuk pada Unit Pembinaan Perlindungan Tanaman (UPPT) tidak menentu tergantung permintaan petani, oleh sebab itu di butuhkan prediksi kebutuhan pupuk. Pupuk yang diprediksi pada Unit Pembinaan Perlindungan Tanaman (UPPT) ada lima jenis antara lain pupuk Urea, pupuk ZA, pupuk SP-36, pupuk NPK, dan pupuk Organik maka dapat di lakukan prediksi kebutuhan pupuk. Dalam prediksi data mining pada kebutuhan pupuk menggunakan algoritma ID3. Dimana cara kerjanya menghitung nilai entropy dan gain untuk mendapatkan hasil akhir berupa pohon ke putusan dan rule. Pengujian dilakukan dengan menggunakan software tanagra. Hasil pengujian yang dilakukan pada aplikasi tanagra dengan menggunakan algoritma ID3 berbentuk decision tree, sedangkan pada perhitungan hasil yang didapatkan berupa pohon keputusan.

Kata Kunci: Kebutuhan Pupuk; Data Mining; Algoritma ID3

Abstract—The need for fertilizer at the Plant Protection Development Unit (UPPT) is uncertain depending on the demand of farmers, therefore it is necessary to predict fertilizer needs. There are five types of fertilizers predicted by the Plant Protection Development Unit (UPPT), including Urea fertilizer, ZA fertilizer, SP-36 fertilizer, NPK fertilizer, and Organic fertilizer, so fertilizer needs can be predicted. In predicting data mining on fertilizer needs using the ID3 algorithm. Where it works is calculating the value of entropy and gain to get the final result in the form of a tree to the decision and rule. Testing is done using the tanagra software. The results of the tests carried out on the tanagra application using the ID3 algorithm are in the form of a decision tree, while in the calculation the results obtained are in the form of a decision tree.

Keywords: Fertilizer Needs; Data Mining; ID3 Algoritma Algorithm

1. PENDAHULUAN

Pupuk merupakan material yang ditambahkan pada tanah atau tanaman yang berfungsi untuk melengkapi unsur hara sehingga pertumbuhan tanaman menjadi lebih baik. Pupuk merupakan bahan kimia atau organisme yang berperan dalam penyediaan unsur hara bagi keperluan secara langsung atau tidak langsung. Penggunaan pupuk pada tanaman harus di perhatikan oleh petani yaitu teknik dan cara pemupukannya. Jika pemupukan di lakukan secara asal dapat megakibatkan tanaman menjadi rusak. Dalam pemberian pupuk perlu diperhatikan kebutuhan tumbuhan agar tumbuhan tidak mendapat terlalu banyak zat makanan dan tidak berbahaya pada tumbuhan. Unit Pembinaan Perlindungan Tanaman (UPPT) merupakan kantor pertanian yang menyediakan Pupuk untuk di terima masyarakat. Pupuk yang di berikan sesuai permintaan petani. Maka untuk memprediksi ke butuhan pupuk yang di perlukan selanjutnya kita perlu menghitung kebutuhan pupuk tersebut dengan menggunakan algoritma ID3.

Berdasarkan penelitian terdahulu jurnal Edik Informatika penelitian bidang komputer sains dan pendidikan informatika V2.i2(213-219). Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstrasi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar [1]. Berdasarkan penelitian selanjutnya penerapan algoritma klasifikasi data mining ID3 untuk menentukan penjurusan siswa SMAN 6 Semarang yang di buat oleh Obbie Kristanto Metode yang di gunakan adalah algoritma ID3. ID3 merupakan algoritma yang di gunakan untuk membangun sebuah decision tree atau pohon keputusan.

Dalam penerapan algoritma ID3 (*Iterative Dichotomiser 3*) merupakan sebuah metode yang digunakan untuk membangkitkan pohon keputusan. Metode ini menggunakan konsep dari entropy informasi. Pemilihan atribut dengan menggunakan information gain. Pemilihan atribut pada ID3 dilakukan dengan propertis statistik, yang disebut dengan information gain. Gain mengukur seberapa baik suatu atribut memisahkan training example kedalam kelas target. Atribut dengan informasi tertinggi akan di pilih. Dengan tujuan untuk mendefinisikan gain, pertama-tama digunakan ide dari teori informasi yang disebut entropy [2], [3].

2. METODOLOGI PENELITIAN

2.1 Algoritma Iterative Dichotomiser 3 (ID3)

Iterative Dichotomiser 3 (ID3) adalah algoritma decision tree learning (algoritma pembelajaran pohon keputusan) yang paling dasar. Algoritma ini melakukan pencarian secara rakus /menyeluruh (greedy) pada semua kemungkinan pohon keputusan [4]–[6].

Salah satu algoritma induksi pohon keputusan yaitu ID3 (Iterative Dichotomiser 3). ID3 dikembangkan oleh J. Ross Quinlan. Algoritma ID3 dapat diimplementasikan menggunakan fungsi rekursif (fungsi yang memanggil dirinya sendiri). Algoritma ID3 berusaha membangun decision tree (pohon keputusan) secara top-down (dari atas ke bawah), mulai dengan pertanyaan : “atribut mana yang pertama kali harus dicek dan diletakkan pada root ?” pertanyaan ini dijawab dengan mengevaluasi semua atribut yang ada dengan menggunakan suatu ukuran statistik (yang banyak digunakan adalah information gain) untuk mengukur efektivitas suatu atribut dalam mengklasifikasikan kumpulan

sampel data. Langkah-langkah pada algoritma ID3 diuraikan sebagai berikut:

1. Entropy

Entropy adalah ukuran dari teori informasi yang dapat mengetahui karakteristik dari impurity, dan homogeneity dari kumpulan data. Dari nilai entropy tersebut kemudian dihitung nilai information gain (IG) masing-masing atribut.

Dimana :

- a. S adalah ruang (data) sample yang digunakan untuk training.
- b. P+ adalah jumlah yang bersolusi positif (mendukung) pada data sample untuk kriteria tertentu
- c. P- adalah jumlah yang bersolusi negatif (tidak mendukung) pada data sample untuk kriteria tertentu. Dari rumus entropy diatas dapat disimpulkan bahwa definisi entropy (S) adalah jumlah bit yang diperkirakan dibutuhkan untuk dapat mengekstrak suatu kelas (+ atau -) dari sejumlah data acak pada suatu ruang sampel S. Entropy bisa dikatakan sebagai kebutuhan bit untuk menyatakan suatu kelas. Semakin kecil nilai entropy maka semakin baik digunakan dalam mengekstraksi suatu kelas.

Panjang kode untuk menyatakan informasi secara optimal adalah $-\log_2 p$ bits untuk messages yang mempunyai probabilitas p. Sehingga jumlah bit yang diperkirakan untuk mengekstraksi S ke dalam kelas adalah $-\sum p \log_2 p$

2. Informasi Gain

Setelah mendapat nilai entropy untuk suatu kumpulan data, maka kita dapat mengukur efektivitas suatu atribut dalam mengklasifikasikan data. Ukuran efektifitas ini disebut information gain. Secara matematis, information gain dari suatu atribut A, dituliskan sebagai berikut :

dimana : A : atribut

V : suatu nilai yang mungkin untuk atribut A

Values (A) : himpunan yang mungkin untuk atribut A

|Sv| : Jumlah sampel Untuk nilai v

|S| : Jumlah seluruh sampel data

Entropy(Sv): entropy untuk sampel-sampel yang memiliki nilai v [2].

3. HASIL DAN PEMBAHASAN

Dalam kebutuhan pupuk dapat melakukan pengumpulan data. Oleh sebab itu, untuk menghasilkan kesimpulan berdasarkan pohon keputusan pada analisa data kebutuhan pupuk diperlukan data terdahulu. Analisa tersebut dilakukan berdasarkan teknik algoritma ID3 dengan beberapa langkah yang telah ditentukan. Data yang diambil merupakan data tahun 2017-2019 pada Unit Pembinaan Perlindungan Tanaman (UPPT) dan Cara kerjanya dapat dilihat pada tabel berikut ini. Pengumpulan data yang ingin di prediksi. Seperti tabel di bawah ini :

Tabel 1. Data Sampel

Bulan	UREA	ZA	SP-36	NPK	ORGANIK
2017 Januari	57000	28500	13500	13500	0
2017 Februari	57000	13500	13500	1200	0
2017 Maret	0	0	0	0	0
2017 April	12600	9635	3145	13545	0
2017 Mei	2800	2800	2800	0	0
2017 Juni	9550	6600	0	13200	0
2017 Juli	7880	5415	1900	6970	0
2017 Agustus	4000	3475	1900	8950	10500
2017 September	3880	1940	0	5820	0
2017 Oktober	7150	6450	3650	7850	910
2017 November	3650	3650	3650	3650	910
2017 Desember	3500	2800	0	4200	0
2018 Januari	0	0	0	0	0
2018 Februari	2650	0	100	3950	300
2018 Maret	1300	500	950	4200	400
2018 April	0	750	150	3800	0
2018 Mei	0	750	1150	4350	200
2018 Juni	0	550	1200	4050	0
2018 Juli	0	50	300	1600	0
2018 Agustus	3100	400	700	3350	450
2018 September	2950	800	950	4800	200
2018 Oktober	2200	725	1500	5600	500
2018 November	0	0	0	0	0
2018 Desember	0	0	0	0	0

Bulan	UREA	ZA	SP-36	NPK	ORGANIK
2019 Januari	70930	51005	15280	81260	1700
2019 Februari	57630	41030	11580	61310	29000
2019 Maret	38925	28800	13950	50580	25000
2019 April	113851	92957	7766	98657	0
2019 Mei	7140	48762	7766	51522	13958
2019 Juni	42450	48195	0	45335	0

Setelah itu ubah bentuk data seperti yang ada pada tabel di bawah:

Tabel 2. Kebutuhan Pupuk

UREA	ZA	SP-36	NPK	ORGANIK	Kebutuhan Pupuk
Naik	Naik	Naik	Naik	Turun	Naik
Tetap	Turun	Tetap	Turun	Tetap	Tetap
Turun	Turun	Turun	Turun	Tetap	Turun
Naik	Naik	Naik	Naik	Tetap	Naik
Turun	Turun	Turun	Turun	Tetap	Turun
Naik	Naik	Turun	Naik	Tetap	Naik
Turun	Turun	Naik	Turun	Tetap	Turun
Turun	Turun	Tetap	Naik	Naik	Naik
Turun	Turun	Turun	Turun	Turun	Turun
Naik	Naik	Naik	Naik	Naik	Naik
Turun	Turun	Tetap	Turun	Tetap	Tetap
Turun	Turun	Turun	Naik	Turun	Turun
Turun	Turun	Tetap	Turun	Tetap	Tetap
Naik	Tetap	Naik	Naik	Naik	Naik
Turun	Naik	Naik	Naik	Naik	Naik
Turun	Naik	Turun	Turun	Turun	Turun
Tetap	Tetap	Naik	Naik	Naik	Naik
Tetap	Turun	Naik	Turun	Turun	Turun
Tetap	Turun	Turun	Turun	Tetap	Turun
Naik	Naik	Naik	Naik	Naik	Naik
Turun	Naik	Naik	Naik	Turun	Naik
Turun	Turun	Naik	Naik	Naik	Naik
Turun	Turun	Turun	Turun	Turun	Turun
Tetap	Tetap	Tetap	Tetap	Tetap	Tetap
Naik	Naik	Naik	Naik	Naik	Naik
Turun	Turun	Turun	Turun	Naik	Turun
Turun	Turun	Naik	Turun	Turun	Turun
Naik	Naik	Turun	Naik	Turun	Naik
Turun	Turun	Tetap	Turun	Naik	Turun
Naik	Turun	Turun	Turun	Turun	Turun

Kemudian Hitunlah *Entropy* untuk node akar (semua data)

$$\begin{aligned}
 Entropy(\text{Semua}) &= -\sum p_i \log_2 p_i \\
 &= -[(P(\text{Naik}|\text{Semua}) \times \log_2 (P(\text{Naik}|\text{Semua}))) + (P(\text{Tetap}|\text{Semua}) \times \log_2 (P(\text{Tetap}|\text{Semua}))) + (P(\text{Turun}|\text{Semua}) \times \log_2 (P(\text{Turun}|\text{Semua}))) \\
 &= -[(13/30) \times \log_2 (13/30)] + [(4/30) \times \log_2 (4/30)] + [(13/30) \times \log_2 (13/30)] \\
 &= 1,4331
 \end{aligned}$$

Setelah itu lakukan Perhitungan *entropy* dan gain pada setiap atribut dengan cara perhitungan seperti dibawah ini.

$$\begin{aligned}
 Entropy(\text{Semua Naik}) &= -\sum p_i \log_2 p_i \\
 &= -[(P(\text{Naik}|\text{Naik}) \times \log_2 P(\text{Naik}|\text{Naik})) + (P(\text{Tetap}|\text{Naik}) \times \log_2 P(\text{Tetap}|\text{Naik})) + (P(\text{Turun}|\text{Naik}) \times \log_2 (P(\text{Turun}|\text{Naik}))) \\
 &= -[(8/9) \times \log_2 (8/9)] + [(0/9) \times \log_2 (0/9)] + [(1/9) \times \log_2 (1/9)] \\
 &= 0
 \end{aligned}$$

$$\begin{aligned}
 Entropy(\text{Semua Tetap}) &= -\sum p_i \log_2 p_i \\
 &= -[(P(\text{Naik}|\text{Tetap}) \times \log_2 P(\text{Naik}|\text{Tetap})) + (P(\text{Tetap}|\text{Tetap}) \times \log_2 P(\text{Tetap}|\text{Tetap})) + (P(\text{Turun}|\text{Tetap}) \times \log_2 (P(\text{Turun}|\text{Tetap}))) \\
 &= -[(1/5) \times \log_2 (1/5)] + [(2/5) \times \log_2 (2/5)] + [(2/5) \times \log_2 (2/5)]
 \end{aligned}$$

$$= 1,5219$$

$$\begin{aligned} \text{Entropy (Semua Turun)} &= -\sum p_i \log_2 p_i \\ &= -[(P(\text{Naik}|\text{Turun}) \times \log_2 P(\text{Naik}|\text{Turun})) + (P(\text{Tetap}|\text{Turun}) \times \log_2 \\ &P(\text{Tetap}|\text{Turun})) + (P(\text{Turun}|\text{Turun}) \times \log_2 P(\text{Turun}|\text{Turun}))] \\ &= -[(4/16) \times \log_2 (4/16) + (2/16) \times \log_2 (2/16) + (10/16) \times \log_2 (10/16)] \\ &= 1,2987 \end{aligned}$$

$$\begin{aligned} \text{Gain (Semua, Urea)} &= \text{Entropy (Semua)} - P(V_i/S) \times \text{Entropy (S}_i) \\ &= E(\text{Semua}) - [(P(\text{Naik}|\text{Semua}) \times E(\text{Semua Naik})) + (P(\text{Tetap}|\text{Semua}) \times E(\text{Semua Tetap})) + (P(\text{Turun}|\text{Semua}) \times E \\ &(\text{Semua Turun}))] \\ &= 1,4331 - ((9/30) \times 0) + ((5/30) \times 1,5219) + ((16/30) \times 1,2987) \\ &= 0,4868 \end{aligned}$$

Dari perhitungan di atas maka akan di dapatkan node akar seperti pada tabel di bawah ini.

Tabel 3. Hasil perhitungan entropy dan gain untuk node akar

Node	Jumlah	Naik	Tetap	Turun	Entropy	Gain	
1 Total	30	13	4	13	1,4331		
Urea	Naik	9	8	0	1	0	0,4868
	Tetap	5	1	2	2	1,5219	
	Turun	16	4	2	10	1,2987	
ZA	Naik	10	9	0	1	0	1,1625
	Tetap	3	2	1	0	0	
	Turun	17	2	3	12	1,1595	
SP-36	Naik	13	10	0	3	0	1,1827
	Tetap	6	1	4	1	1,2516	
	Turun	11	2	0	9	0	
NPK	Naik		14	13	0	11	1,4331
	Tetap		1	0	1	0	
	Turun		15	0	3	12	
Organik	Naik		10	8	0	2	0
	Tetap		10	2	4	4	
	Turun		10	3	0	7	

Maka langkah selanjutnya melakukan pemisahan menjadi 3 kelompok seperti pada tabel dibawah.

Tabel 4. Hasil pemisahan data oleh fiktur ‘NPK’

UREA	ZA	SP-36	NPK	ORGANIK	Kebutuhan Pupuk
Naik	Naik	Naik	Naik	Turun	Naik
Naik	Naik	Naik	Naik	Tetap	Naik
Naik	Naik	Turun	Naik	Tetap	Naik
Turun	Turun	Tetap	Naik	Naik	Naik
Naik	Naik	Naik	Naik	Naik	Naik
Turun	Turun	Turun	Naik	Turun	Turun
Naik	Tetap	Naik	Naik	Naik	Naik
Turun	Naik	Naik	Naik	Naik	Naik
Tetap	Tetap	Naik	Naik	Naik	Naik
Naik	Naik	Naik	Naik	Naik	Naik
Turun	Naik	Naik	Naik	Turun	Naik
Turun	Turun	Naik	Naik	Naik	Naik
Naik	Naik	Naik	Naik	Naik	Naik
Naik	Naik	Turun	Naik	Turun	Naik
Tetap	Tetap	Tetap	Tetap	Tetap	Tetap
Tetap	Turun	Tetap	Turun	Tetap	Tetap
Turun	Turun	Turun	Turun	Tetap	Turun
Turun	Turun	Turun	Turun	Tetap	Turun

UREA	ZA	SP-36	NPK	ORGANIK	Kebutuhan Pupuk
Turun	Turun	Naik	Turun	Tetap	Turun
Turun	Turun	Turun	Turun	Turun	Turun
Turun	Turun	Tetap	Turun	Tetap	Tetap
Turun	Turun	Tetap	Turun	Tetap	Tetap
Turun	Naik	Turun	Turun	Turun	Turun
Tetap	Turun	Naik	Turun	Turun	Turun
Tetap	Turun	Turun	Turun	Tetap	Turun
Turun	Turun	Turun	Turun	Turun	Turun
Turun	Turun	Turun	Turun	Naik	Turun
Turun	Turun	Naik	Turun	Turun	Turun
Turun	Turun	Tetap	Turun	Naik	Turun
Naik	Turun	Turun	Turun	Turun	Turun

Setelah itu maka kita akan mencari node akar ke dua dengan cara perhitungan seperti mencari node akar pertama. Dan akan mendapat hasil perhitungan untuk node akar ke dua seperti tabel di bawah in

Table 4. Hasil perhitungan entropy dan gain untuk node 2

Node	Jumlah	Naik	Turun	Entropy	Gain
2 Total	14	13	1	0,3712	

Langka selanjutnya menghitung node akar ke 4 dengan cara perhitungan seperti mencari node sebelumnya. Dan akan mendapatkan hasil perhitungan seperti tabel di bawah ini.

Tabel 5. Hasil perhitungan entropy dan gain untuk node 4

Node	Jumlah	Tetap	Turun	Entropy	Gain	
4 Total	15	3	12	0,7219		
UREA	Naik	1	1	0	0	
	Tetap	3	1	2	0,9182	0,0366
	Turun	11	9	2	0,684	
ZA	Naik	1	0	1	0	0,0223
	Naik	3	0	3	0	
	Tetap	4	3	1	0,8112	
SP-36	Turun	8	0	8	0	0,5055
	Naik	2	0	2	0	
	Tetap	7	3	4	0,9852	
Organik	Turun	6	0	6	0	0,2621

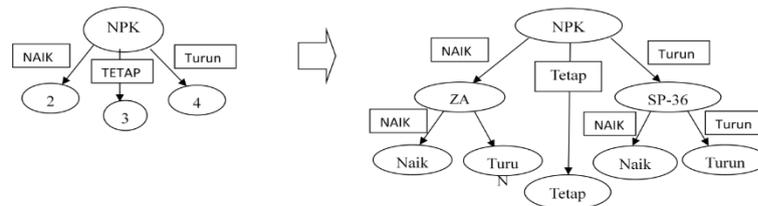
Langka selanjutnya melakukan pemisahan dari setiap hasil node yang di dapat seperti pada tabel di bawah ini.

Tabel 6. Tabel Hasil Pemisahan Data

UREA	ZA	SP-36	NPK	ORGANIK	Kebutuhan Pupuk
Naik	Naik	Naik	Naik	Turun	Naik
Naik	Naik	Naik	Naik	Tetap	Naik
Naik	Naik	Turun	Naik	Tetap	Naik
Naik	Naik	Naik	Naik	Naik	Naik
Turun	Naik	Naik	Naik	Naik	Naik
Naik	Naik	Naik	Naik	Naik	Naik
Turun	Naik	Naik	Naik	Turun	Naik
Naik	Naik	Naik	Naik	Naik	Naik
Naik	Naik	Turun	Naik	Turun	Naik
Naik	Tetap	Naik	Naik	Naik	Naik
Tetap	Tetap	Naik	Naik	Naik	Naik
Turun	Turun	Tetap	Naik	Naik	Naik

Turun	Turun	Turun	Naik	Turun	Turun
Turun	Turun	Naik	Naik	Naik	Naik
Tetap	Tetap	Tetap	Tetap	Tetap	Tetap
Turun	Turun	Naik	Turun	Tetap	Turun
Tetap	Turun	Naik	Turun	Turun	Turun
Turun	Turun	Naik	Turun	Turun	Turun
Tetap	Turun	Tetap	Turun	Tetap	Tetap
Turun	Turun	Tetap	Turun	Tetap	Tetap
Turun	Turun	Tetap	Turun	Naik	Turun
Turun	Turun	Turun	Turun	Tetap	Turun
Turun	Turun	Turun	Turun	Turun	Turun
Turun	Turun	Turun	Turun	Turun	Turun
Turun	Naik	Turun	Turun	Turun	Turun
Tetap	Turun	Turun	Turun	Tetap	Turun
Turun	Turun	Turun	Turun	Turun	Turun
Turun	Turun	Turun	Turun	Naik	Turun
Naik	Turun	Turun	Turun	Turun	Turun

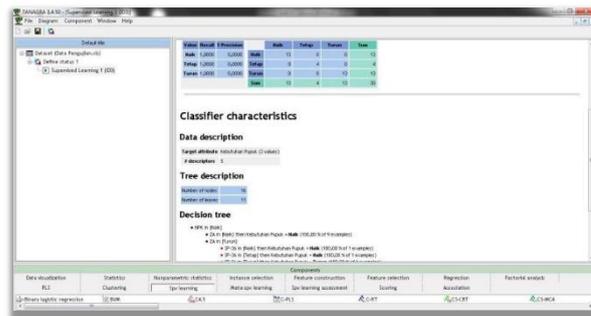
Setelah itu langkah selanjutnya membuat pohon keputusan dari hasil di atas.



Gambar 1. Pohon Keputusan

3.1 Pengujian

Untuk pengujian metode ID3 menggunakan tools tanagra versi 1.4.50 . Dan Ini adalah tampil tanagra sesudah data di masukan lewat microsoft excel. Ini adalah tampilan akhir dari decision tree.



Gambar 2. Hasil Perhitungan Tanagra

4. KESIMPULAN

Berdasarkan hasil-hasil analisis dan percobaan yang dilakukan pada bab sebelumnya, maka kesimpulan yang dapat diambil cara kerja algoritma ID3 dengan hasil akhir adalah pohon keputusan dan juga rule, dimana pohon keputusan dan rule dibentuk berdasarkan nilai dari entropy dan juga gain. Data mining dapat diterapkan untuk memprediksi kebutuhan pupuk pada Unit Pembinaan Perlindungan Tanaman (UPPPT). *Tools* tanagra dapat digunakan untuk menguji hasil menggunakan algoritma ID3.

REFERENCES

- [1] Y. Mardi, "Data Mining : Klasifikasi Menggunakan Algoritma C4.5," *J. Edik Inform.*, vol. 2, no. 2, pp. 213–219, 2017.
- [2] G. G. Reader, "Organization and Development of a Comprehensive Care Program," *Am. J. Public Heal. Nations Heal.*, vol. 44, no. 6, pp. 760–765, 2008.
- [3] A. E. Pramadhani and T. Setiadi, "PENERAPAN DATA MINING UNTUK KLASIFIKASI PREDIKSI PENYAKIT ISPA (Infeksi Saluran Pernapasan Akut) DENGAN ALGORITMA DECISION TREE (ID3)," *J. Sarj. Tek. Inform. e-ISSN 2338-5197*, vol. 2, no. 1, pp. 831–839, 2014.
- [4] D. M. Agustina and Wijanarto, "Analisis Perbandingan Algoritma ID3 Dan C4.5 Untuk Klasifikasi Penerima Hibah Pemasangan Air Minum Pada PDAM Kabupaten Kendal," *J. Appl. Intell. Syst.*, vol. 1, no. 3, pp. 234–244, 2016.
- [5] E. Buulolo, *Data Mining Untuk Perguruan Tinggi*. Deepublish, 2020.
- [6] W. W. Ariestya, Y. E. Praptiningsih, and W. Supriatin, "Decision Tree Learning Untuk Penentuan Jalur Kelulusan Mahasiswa," *J. Ilm. FIFO*, vol. 8, no. 1, p. 97, 2016.