

# Penerapan Algoritma C4.5 Dalam Memprediksi Kebutuhan Pembibitan Pohon

Helmida Br Manik

Fakultas Ilmu Komputer dan Teknologi Informasi, Prodi Teknik Informatika, Universitas Budi Darma, Medan, Indonesia

Email: [1manikhelmida1@gmail.com](mailto:1manikhelmida1@gmail.com)

**Abstrak**—Pembibitan pohon merupakan tempat yang di kelola dan di rancang untuk memproduksi bibit pohon yang di besarkan di dalam kondisi baik sampai bibit- bibit ini siap untuk di tanam. Pembibitan pohon ini dapat berupa pembibitan tidak resmi yang berskala kecil atau badan usaha komersial besar. Pembibitan memiliki keragaman dalam hal ukuran, fasilitas (suplai, peralatan, perlengkapan, dan lain-lain.), Tipe bibit yang di produksi, dan operasional. Pembibitan- pembibitan juga memiliki perbedaan signifikan dalam hal kualitas dan kuantitas stok bahan tanam yang di produksi. Namun tujuan utama semua pembibitan adalah memproduksi jumlah bibit berkualitas tinggi yang cukup untuk memenuhi kebutuhan pengguna bibit. Para pengguna bibit mencakup operator pembibitan itu sendiri, perseorangan, organisasi, masyarakat, kelompok petani, badan pemerintahan, organisasi non pemerintah, perusahaan, atau konsumen swasta. Masalah yang di alami selama ini pada dinas kehutanan sumatera utara untuk menentukan perbenihan masih kurang efektif di karenakan menentukan pembibitan pohon pertahunnya masih sangat susah, jadi untuk mengatasi masalah tersebut di terapkan algoritma C4.5 di dalam memprediksi pembibitan pada dinas kehutanan sumatera utara di karenakan Prediksi merupakan suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa lalu dan sekarang yang di miliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat di perkecil. Prediksi tidak harus memberikan jawaban secara pasti kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin yang akan terjadi. Algoritma C4.5 suatu algoritma yang di gunakan unuk menghasilkan suatu pohon keputusan yang mampu mengklasifikasi suatu obyek. C4.5 merepresentasi konsep-konsep dalam bentuk pohon keputusan. Aturan-aturan yang di hasilkan oleh C4.5 mempunyai relasi yang hirarkis seperti suatu pohon (mempunyai akar, titik, cabang, dan daun). Beberapa menyebut struktur model yang di hasilkan C4.5 sebagai pohon keputusan (decision tree) sementara yang lain menyebutnya pohon aturan (rule tree). Di mana di dalam proses pengerjaan pada algoritma C4.5 menghitung nilai gain dan entropy terhadap atribut – atribut data yang telah di sajikan oleh basis data sebelumnya. Dari proses klasifikasi dan hasil yang di dapatkan di gambarkan dalam bentuk pohon keputusan dan berdasarkan pohon keputusan tersebut di dapatkanlah sebuah informasi baru yang berasal dari basis data sebelumnya dalam bentuk aturan – aturan atau rule konsumen swasta.

**Kata Kunci:** Data Mining; Prediksi; Kebutuhan; Pembibitan Pohon; Algoritma C4.5

**Abstract**—A tree nursery is a place that is managed and designed to produce tree seedlings that are raised in good conditions until these seedlings are ready for planting. These tree nurseries can be small-scale informal nurseries or large commercial enterprises. Nurseries vary in size, facilities (supply, equipment, supplies, etc.), type of seed produced, and operations. Nurseries also have significant differences in the quality and quantity of stock of planting material produced. However, the primary goal of all nurseries is to produce sufficient quantities of high quality seed to meet the needs of seedling users. Seed users include the nursery operators themselves, individuals, organizations, communities, farmer groups, government agencies, non-governmental organizations, companies, or private consumers. The problems that have been experienced so far at the North Sumatran Forestry Service to determine seedling are still ineffective because determining the annual tree nursery is still very difficult, so to overcome this problem, the C4.5 algorithm is applied in predicting nurseries at the North Sumatra Forestry Service because of predictions. is a process of systematically estimating something that is most likely to happen in the future based on past and present information that is owned, so that the error (difference between something that happened and the predicted result) can be minimized. Prediction does not have to provide a definite answer to events that will occur, but seeks to find answers as close as possible to what will happen. The C4.5 algorithm is an algorithm that is used to produce a decision tree that is able to classify an object. C4.5 represents concepts in the form of a decision tree. The rules generated by C4.5 have a hierarchical relationship like a tree (having roots, points, branches, and leaves). Some call the structure of the model generated by C4.5 a decision tree while others call it a rule tree. Where in the process of working on the C4.5 algorithm, it calculates the gain and entropy values for the data attributes that have been presented by the previous database. From the classification process and the results obtained are described in the form of a decision tree and based on the decision tree, new information comes from the previous database in the form of rules or rules private consumers.

**Keywords:** Data Mining; Prediction; Needs; Tree Nurseries; C4.5 Algorithm

## 1. PENDAHULUAN

Pohon merupakan paru-paru dunia yang harus di jaga, karena keberlangsungan kehidupan manusia juga tergantung dari pohon. Pohon merupakan tumbuhan berkayu yang mempunyai sebuah batang utama dengan dahan dan ranting yang jauh dari permukaan tanah. Pohon merupakan tumbuhan berkayu yang memiliki diameter batang setinggi dada (*breast height*). Menurut *Dengler*, pohon merupakan suatu tumbuhan yang mempunyai akar, batang dan tajuk yang jelas dan tinggi minimal 5 meter.

Pembibitan pohon merupakan tempat yang di kelola dan di rancang untuk memproduksi bibit pohon yang di besarkan di dalam kondisi baik sampai bibit- bibit ini siap untuk di tanam. Pembibitan pohon ini dapat berupa pembibitan tidak resmi yang berskala kecil atau badan usaha komersial besar. Pembibitan memiliki keragaman dalam hal ukuran, fasilitas (suplai, peralatan, perlengkapan, dan lain-lain.), Tipe bibit yang di produksi, dan operasional. Pembibitan- pembibitan juga memiliki perbedaan signifikan dalam hal kualitas dan kuantitas stok bahan tanam yang di produksi. Namun tujuan utama semua pembibitan adalah memproduksi jumlah bibit berkualitas tinggi yang cukup untuk memenuhi

kebutuhan pengguna bibit. Para pengguna bibit mencakup operator pembibitan itu sendiri, perseorangan, organisasi, kemasyarakatan, kelompok petani, badan pemerintahan, organisasi non pemerintah, perusahaan, atau konsumen swasta.

Masalah yang di alami selama ini pada dinas kehutanan sumatera utara untuk menentukan perbenihan masih kurang efektif di karenakan menentukan pembibitan pohon pertahunnya masih sangat susah, jadi untuk mengatasi masalah tersebut di terapkan algoritma C4.5 di dalam memprediksi pembibitan pada dinas kehutanan sumatera utara di karenakan Prediksi merupakan suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa lalu dan sekarang yang di miliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat di perkecil. Prediksi tidak harus memberikan jawaban secara pasti kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin yang akan terjadi.

Algoritma C4.5 suatu algoritma yang di gunakan unuk menghasilkan suatu pohon keputusan yang mampu mengklasifikasi suatu obyek. C4.5 merepresentasi konsep-konsep dalam bentuk pohon keputusan. Aturan-aturan yang di hasilkan oleh C4.5 mempunyai relasi yang hirarkis seperti suatu pohon (mempunyai akar, titik, cabang, dan daun). Beberapa menyebut struktur model yang di hasilkan C4.5 sebagai pohon keputusan (*decision tree*) sementara yang lain menyebutnya pohon aturan (*rule tree*). Di mana di dalam proses pengerjaan pada algoritma C4.5 menghitung nilai *gain* dan *entropy* terhadap atribut – atribut data yang telah di sajikan oleh basis data sebelumnya[1], [2]. Dari proses klasifikasi dan hasil yang di dapatkan di gambarkan dalam bentuk pohon keputusan dan berdasarkan pohon keputusan tersebut di dapatkanlah sebuah informasi baru yang berasal dari basis data sebelumnya dalam bentuk aturan – aturan atau *reale*.

## 2. METODOLOGI PENELITIAN

### 2.1 Data Mining

Data mining, sering juga di sebut *knowledge discovery in database* (KDD), adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam *set* data berukuran besar. Keluaran dari data mining bisa di pakai untuk memperbaiki pengambilan keputusan di masa depan[3]–[5].

### 2.2 Pembibitan Pohon

Pembibitan pohon adalah faktor penting untuk meraih kesuksesan di dalam berbagai interfensi pembangunan kehutanan dan pertanian selama 2 dekade terakhir World Agroforestry Centry (ICRAF), Winrock Internasional, dan University of Philippines Los Banos (UPLB) telah bekerja sama dengan ratusan perbenihan pohon skala kecil dan besar di kawasan Asia Tenggara. Sebagian besar pembibitan ini berlokasi di Indonesia dan Filipina. Pembibitan ini memiliki beragam tujuan mulai dari produksi biomassa secara komersial, rehabilitasi lahan dan konserfasi hutan, hingga pengembangan kapasitas dan peningkatan matapencaharian. Para mitra yang terlibat dalam pengoperasian pembibitan ini meliputi para petani, pengusaha, perusahaan komersial, lembaga swadaya masyarakat (LSM), masyarakat, proyek dan badan pemerintahan.

### 2.3 Algoritma C4.5

Algoritma C4.5 merupakan salah satu algoritma yang di gunakan untuk membentuk *decision tree* berdasarkan *training data*. Algoritma ini merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal [6]. *Decision tree* adalah model prediksi menggunakan struktur berhirarki. Konsep dari *decision tree* adalah mengubah data menjadi *decision tree* dan aturan-aturan keputusan (*decision rules*). Algoritma C4.5 merupakan pengembangan dari ID3. Beberapa pengembangan yang di lakukan C4.5 adalah bisa mengatasi *missing value*, *continue data* dan *pruning*. Algoritma C4.5 mempunyai input berupa *training samples* dan *samples*, *training samples* berupa data contoh yang akan di gunakan untuk membangun sebuah *tree* yang telah di uji kebenarannya, sedangkan untuk *semples* merupakan *field-field* data yang nantinya akan kita gunakan sebagai parameter dalam melakukan klasifikasi data[6]–[8]. Adapun langkah-langkah untuk penyelesaian Algoritma C4.5 yaitu sebagai berikut:

1. Mencari Nilai *Entropy* dari kriteria-kriteria
2. Mencari Nilai *Gain* dari setiap atribut
3. Pembentukan Atribut sebagai akar berdasarkan *gain* tertinggi
4. Pembentukan Cabang berdasarkan masing-masing nilai
5. Ulangi Proses untuk masing-masing cabang

Untuk mencari Rumus pada Algoritma C4.5 pada pohon keputusan yaitu sebagai berikut:

$$Gain(S,A) = Entrophy(S) - \sum_{i=1}^n \frac{S_i}{S} * Entrophy(S_i) \quad (1)$$

Di mana:

S : himpunan kasus

A : atribut

n : jumlah partisi atribut A

|Si| : jumlah kasus pada partisi ke-i

|S| : jumlah kasus dalam S

$$Entrophy(S) = Entrophy(S) - \sum_{i=1}^n \frac{S_i}{S} * pi * \log_2 pi \quad (2)$$

Di mana :

S : himpunan kasus

A : atribut

n : jumlah partisi S

$p_i$  : proporsi dari  $S_i$  terhadap S

### 3. HASIL DAN PEMBAHASAN

Masalah yang di alami selama ini pada dinas kehutanan sumatera utara untuk menentukan pembibitan masih kurang efektif di karena menentukan pembibitan pohon pertahunnya masih sangat susah, jadi untuk mengatasi masalah tersebut di terapkan algoritma C4.5 di dalam memprediksi pembibitan pada dinas kehutanan sumatera utara di karenakan Prediksi merupakan suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa lalu dan sekarang yang di miliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat di perkecil. Prediksi tidak harus memberikan jawaban secara pasti kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin yang akan terjadi.

Pemrosesan data *mining* menggunakan algoritma C4.5 dengan melakukan pengumpulan data dan mengklasifikasikan data menggunakan representasi struktur pohon di mana setiap *node* merepresentasikan nilai dari *atribut*, cabangnya merepresentasikan nilai dari *atribut* dan daun merepresentasikan *node*. Suatu pengklasifikasian kebutuhan pembibitan pohon dari data yang berskala besar dan yang mengasosiasikan data yang satu dengan yang lainnya dengan menggunakan algoritma C4.5 yang bisa melakukan penelusuran pada data pembibitan untuk mengklasifikasi penyebab terjadinya kebutuhan pembibitan pohon sehingga di temukan aturan-aturan yang saling berhubungan. Data yang di miliki akan di susun menjadi sebuah tabel dan jumlah responden sebelum di lakukan perhitungan untuk mencari nilai *entropy* dan *gain*. Sedangkan *output* nya berupa proses klasifikasi dengan membuat pohon keputusan akhir menjadi kebutuhan pembibitan pohon. Hasil dari *output* memiliki dua kategori yaitu meningkat dan menurun, maka di gunakan teknik klasifikasi yang termasuk dalam pengelompokan data *mining* karena dalam klasifikasi terdapat target variabel dan kategori.

Pengujian di lakukan terhadap sampel contoh kasus yang akan di uji. Pengambilan contoh uji di ambil dari beberapa kriteria dari sampel data yang di gunakan. Pengujian ini di lakukan untuk menentukan apakah sampel tersebut mempunyai hasil keputusan yang sama.

#### 3.1 Penerapan Algoritma C4.5

Sistem yang di gunakan dalam prediksi kebutuhan pembibitan pohon yaitu dengan menggunakan data awal *numeric* dan non *numeric* yang akan di bagi melalui atribut-atribut untuk memudahkan analisa berikutnya, kemudian setelah semua data yang di bagi melalui atributnya maka akan di lakukan proses klasifikasi dengan membuat pohon keputusan sebagai *output*. Proses pengambilan keputusan untuk mengkalasifikasi prediksi kebutuhan pembibitan pohon yang menjadi keputusan adalah berdasarkan kebutuhan pembibitan pohon yang di butuhkan yang akan di jadikan sampel untuk analisa dan untuk pengujian di ambil dari tahun 2016 -2018.

**Tabel 1.** Data Hasil kebutuhan pohon Priode 2016 -2018

No	Penerima	Sasaran Tanam (ha)	Hasil kebutuhan	Curah Hujan	Hama	Kategori
1	Jaring Hijau	Sempit	Sedikit	Sedang	Tidak	Meningkat
2	Ditpolair	Luas	Banyak	Sedang	Tidak	Meningkat
3	Kejati DelSer	Luas	Banyak	Sedang	Tidak	Meningkat
4	Gamadiksi	Luas	Banyak	Tinggi	Tidak	Meningkat
5	Sukma	Sempit	Sedang	Tinggi	Ya	Meningkat
6	IMM FISIP UMSU	Sempit	Sedikit	Tinggi	Ya	Meningkat
7	HMTG TOBA ITM	Sedang	Sedikit	Tinggi	Ya	Meningkat
8	KTH Lestari Alam	Sedang	Sedang	Tinggi	Tidak	Meningkat
9	KTH Hijau Mekar	Sedang	Sedang	Tinggi	Tidak	Menurun

Pohon keputusan di buat setelah menghitung *entropy* total, *entropy* masing - masing *atribut* dan menghitung *gain* dan menentukan *gain* tertinggi. *Entropy* merupakan kebutuhan untuk menyatakan suatu kelas yang di gunakan dalam mengekstrak suatu kelas. Sedangkan *Gain* merupakan salah satu *attribute selection measure* yang di gunakan untuk memilih *test attribute* tiap *node* pada *tree*. Seperti yang telah di jelaskan sebelumnya, algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut :

1. Pilih atribut sebagai akar.
  2. Buat cabang untuk tiap-tiap nilai.
  3. Bagi kasus dalam cabang.
  4. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.
- Langkah menghitung Gain dan Entropy yang di mana Gain dan Entropy di gunakan untuk pembentukan Pohon Keputusan dapat di lihat pada tabel 4.3 berikut:

$$Entropy(total(s)) = \left(-\frac{8}{9} * \log_2\left(\frac{8}{9}\right)\right) = 0,503258335$$

*Entropy(total)* adalah menghitung nilai total kategori Sasaran Tanaman (3) dan Hasil kebutuhan (3) Curah hujan (2) dan hama (2), Kategori (2).

**A. Sasaran Tanaman**

$$entropy( sempit ) = \left(-\frac{0}{3} * \log_2\left(\frac{0}{3}\right)\right) = 0$$

$$entropy( sedang ) = \left(-\frac{1}{2} * \log_2\left(\frac{1}{2}\right)\right) = 0,918295834$$

$$entropy( luas ) = \left(-\frac{0}{3} * \log_2\left(\frac{0}{3}\right)\right) = 0$$

**B. Atribut hasil Kebutuhan**

$$entropy( sedikit ) = \left(-\frac{0}{3} * \log_2\left(\frac{0}{3}\right)\right) = 0$$

$$entropy( sedang ) = \left(-\frac{1}{2} * \log_2\left(\frac{1}{2}\right)\right) = 0,918295834$$

$$entropy( banyak ) = \left(-\frac{0}{3} * \log_2\left(\frac{0}{3}\right)\right) = 0$$

**C. Atribut Curah Hujan**

$$entropy( sedang ) = \left(-\frac{0}{3} * \log_2\left(\frac{0}{3}\right)\right) = 0$$

$$entropy( tinggi ) = \left(-\frac{1}{5} * \log_2\left(\frac{1}{5}\right)\right) = 0,650022422$$

**D. Atribut Hama**

$$entropy( tidak ) = \left(-\frac{1}{5} * \log_2\left(\frac{1}{5}\right)\right) = 0,650022422$$

$$entropy( ya ) = \left(-\frac{0}{3} * \log_2\left(\frac{0}{3}\right)\right) = 0$$

Menghitung nilai *Entropy* Kebutuhan Pembibitan Pohon berdasarkan atribut dari tiap-tiap kelas (Sempit, Sedang, Luas, Banyak, Tidak, Ya) pada atribut .Menghitung Nilai *Gain* tiap-tiap atribut menggunakan rumus (1): Menghitung Nilai *Gain* tiap-tiap atribut

**Gain (Total, sasaran tanaman, hasil kebutuhan, curah hujan, hama)**

$$Entropy(S) - \sum_{i=1}^n \frac{|sasaran\ tanaman\ i|}{|total|} * entropy(sasaran\ tanaman)$$

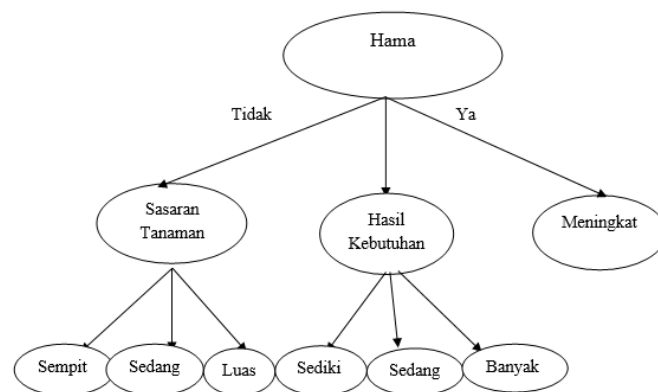
$$= 0,503258335 - \left(\left(\frac{0}{3} * 0\right) + \left(\frac{1}{2} * 0,918295834\right) + \left(\frac{0}{3} * 0\right) + \left(\frac{0}{3} * 0\right) + \left(\frac{1}{2} * 0,918295834\right) + \left(\frac{0}{3} * 0\right) + \left(\frac{0}{3} * 0\right) + \left(\frac{1}{5} * 0,650022422\right) + \left(\frac{1}{5} * 0,650022422\right) + \left(\frac{0}{3} * 0\right)\right) = 1.1783048028$$

Berdasarkan perhitungan di atas, maka dapat di lihat pada tabel perhitungan node berikut :

**Tabel 2.** Perhitungan Node

Node	Atribut	Nilai	Jumlah Kasus / Data Set	Meningkat	Menurun	Entropy	Gain
1.1	Hama --> Tidak		6	5	1	0,650022422	
	Sasaran Tanam						0,31668909
		Sempit	1	1	0	0	
		Sedang	2	1	1	1	
		Luas	3	3	0	0	
	Hasil Kebutuhan						0,31668909

Node	Atribut	Nilai	Jumlah Kasus / Data Set	Meningkat	Menurun	Entropy	Gain
		Sedikit	1	1	0	0	
		Sedang	2	1	1	1	
		Banyak	3	3	0	0	
	Hasil Kebutuhan						0,31668909
		Sedikit	1	1	0	0	
		Sedang	2	1	1	1	
		Banyak	3	3	0	0	
	Curah Hujan						0,1908745
		Sedang	3	3	0	0	
		Tinggi	3	2	1	0,918295834	



**Gambar 1.** Pohon Keputusan

1. IF Hama Ya Then Kategori = Meningkatkan
2. IF Hama Tidak Sasaran Tanaman Then Kategori = Sempit
3. IF Hama Tidak Sasaran Tanaman Then Kategori = Sedang
4. IF Hama Tidak Sasaran Tanaman Then Kategori = Luas
5. IF Hama Tidak Hasil Kebutuhan Then Kategori = Sedikit
6. IF Hama Tidak Hasil Kebutuhan Then Kategori = Sedang
7. IF Hama Tidak Hasil Kebutuhan Then Kategori = Banyak
8. Else IF Hama Ya Then Kategori = Meningkatkan.

Berdasarkan pembentukan pohon keputusan di atas, dan jumlah hasil perhitungan *entropy* dan Gainnya sama dan mendapatkan nilai Meningkatkan. Maka proses perhitungannya berhenti. Oleh karena itu maka di dapatkan hasil prediksi kebutuhan pembibitan pohon yaitu jika jumlah kategori hasil Sempit, Sedang, Luas Kategorinya meningkat.

#### 4. KESIMPULAN

Berdasarkan hasil analisa dari penelitian yang dilakukan, maka kesimpulan dari penelitian ini Faktor kebutuhan pohon dapat diketahui dengan menggunakan data kebutuhan pohon sehingga dapat mengetahui kebutuhan pohon untuk tahun berikutnya. Algoritma C4.5 dapat melakukan pengumpulan dan pengklasifikasian data dengan representasi struktur pohon dimana setiap *node* dan cabangnya merepresentasikan nilai dari *atribut*, dan daun merepresentasikan *node* yang bisa melakukan penelusuran data kebutuhan pohon untuk mengetahui kebutuhan pohon sehingga ditemukan aturan-aturan yang saling berhubungan antara satu sama lain.

#### REFERENCES

- [1] R. Rahim *et al.*, "C4.5 Classification Data Mining for Inventory Control," *Int. J. Eng. Technol.*, vol. 7, no. 2.3, pp. 68–72, 2018.
- [2] F. F. Harryanto and S. Hansun, "Penerapan Algoritma C4 . 5 untuk Memprediksi Penerimaan Calon Pegawai Baru di PT WISE," *Jatiji*, vol. 3, no. 2, pp. 95–103, 2017.
- [3] E. Buulolo, *Data Mining Untuk Perguruan Tinggi*. Deepublish, 2020.
- [4] A. P. Windarto, U. Indriani, M. R. Raharjo, and L. S. Dewi, "Bagian 1 : Kombinasi Metode Klastering dan Klasifikasi ( Kasus Pandemi Covid-19 di Indonesia )," *J. Media Inform. Budidarma*, vol. 4, no. 3, pp. 855–862, 2020.

- [5] A. Widiastari, Solikhun, and Irawan, "Analisa Datamining dengan Metode Klasifikasi C4 . 5 Sebagai Faktor Penyebab Tanah Longsor," *J. Comput. Syst. Informatics*, vol. 2, no. 3, pp. 247–255, 2021.
- [6] Y. S. Nugroho, "Penerapan Algoritma C4.5 Untuk Klasifikasi Predikat Kelulusan Mahasiswa Fakultas Komunikasi Dan Informatika Universitas Muhammadiyah Surakarta," p. 7, 2014.
- [7] D. H. Kamagi and S. Hansun, "Implementasi Data Mining dengan Algoritma C4 . 5 untuk Memprediksi Tingkat Kelulusan Mahasiswa," no. April, 2016.
- [8] E. N. Azizah, U. Pujianto, E. Nugraha, and Darusalam, "Comparative performance between C4.5 and Naive Bayes classifiers in predicting student academic performance in a Virtual Learning Environment," *2018 4th Int. Conf. Educ. Technol. ICET 2018*, no. 1, pp. 18–22, 2018.