

# KOMPARASI DATA MINING KLASIFIKASI PADA PEMILIHAN DOSEN TERBAIK DENGAN MENGGUNAKAN METODE ALGORITMA C 4.5 DAN NAIVE BAYES

Zelvi Gustiana<sup>1</sup>, Ega Evinda Putri<sup>2</sup>

*Program Studi Teknologi Informasi, Universitas Dharmawangsa<sup>1</sup>*  
*Program Studi Rekayasa Perangkat Lunak, Universitas Dharmawangsa<sup>2</sup>*  
Jl. KL. Yos Sudarso No. 224, Glugur Kota, Medan  
zelvi@dharmawangsa.ac.id , ega@dharmawangsa.ac.id

**Abstract** - Lecturers are one of the important elements in tertiary institutions that carry out the task of tri dharma, namely teaching, education and service. In order for the performance of the lecturers to continue to develop, the lecturers are given awards for their work by holding the selection of the best lecturers. Many previous studies conducted the selection of the best lecturers, one of them by using data mining. Classification data mining is one method that is very useful in this case. In this paper, a comparison of data mining classification algorithm C 4.5 and Naïve Bayes is carried out to get which one is better. High accuracy results obtained with a percentage rate of 96.67% compared to data processing with the C 4.5 algorithm, which is 83.33%

**Keywords** - Algorithm C 4.5, Data Mining, Lecturer, Classification, Naïve Bayes.

**Abstrak** – Dosen merupakan salah satu elemen penting di perguruan tinggi yang mengemban tugas tri dharma yaitu pengajaran, Pendidikan dan pengabdian. Agar kinerja dosen menjadi terus berkembang maka para dosen diberikan penghargaan atas kerja mereka dengan diadakannya pemilihan dosen terbaik. Banyak penelitian sebelumnya yang melakukan pemilihan dosen terbaik, salah satunya dengan cara menggunakan data mining. Data mining klasifikasi merupakan salah satu metode yang sangat berguna untuk kasus ini. Pada tulisan ini dilakukan komparasi data mining klasifikasi Algoritma C 4.5 dan Naïve Bayes untuk mendapatkan mana yang lebih baik. Didapatkan hasil akurasi tinggi dengan tingkat persentase 96,67% dibandingkan pengolahan data dengan Algoritma C 4.5 yaitu sebesar 83,33%.

**Kata Kunci** - Algoritma C 4.5, Data Mining, Dosen, Klasifikasi, Naïve Bayes .

## I. PENDAHULUAN

Dosen merupakan elemen yang sangat penting dalam melakukan penyelenggaraan Pendidikan serta subjek yang berkaitan dengan mewujudkan cita-cita institusi Pendidikan [1]. Dosen adalah tenaga pendidik profesional dan merupakan ilmuwan yang mengemban tugas utama dalam melakukan transformasi, pengembangan dan penyebaran ilmu pengetahuan, teknologi serta seni yang diterapkan melalui tri dharma perguruan tinggi. Dalam mendukung aktivitas proses belajar mengajar pada perguruan tinggi agar terciptanya mahasiswa yang memiliki kompetensi pada bidang konsentrasi yang diambil, maka sangat diperlukan seorang tenaga pengajar atau dosen yang berkompentensi dibidangnya [2].

Untuk pemberian penghargaan terhadap apa yang dosen kerjakan untuk akademis, sudah sepatutnya dosen diberikan penghargaan atas tridharma yang telah dilakukan. Penelitian tentang pemilihan dosen terbaik

pernah dilakukan tahun 2020 dengan menggunakan metode SAW berbasis WEB dengan hasil dari seluruh perhitungan didapatkan dosen terbaik dengan nilai tertinggi dan terendah dengan menggunakan Matrix [3].

Selain itu juga dilakukan penelitian dengan membuat perancangan sistem untuk pemilihan dosen terbaik dengan menggunakan metode SAW yang dilakukan guna mendapatkan perhitungan lebih cepat dibandingkan perhitungan manual [4]. Selain itu juga dilakukan penelitian tentang penentuan dosen berprestasi dengan menggunakan metode AHP (Analytical Hierarchy Process) yang mendapatkan kesesuaian dengan sistem sebanyak 66,3% [5]

Dengan berbagai sistem pendukung keputusan yang telah dilakukan untuk pemilihan dosen terbaik, maka dilakukan penelitian dengan membandingkan dua algoritma klasifikasi yaitu Algoritma C 4.5 dan Naïve Bayes untuk pemilihan dosen terbaik.

## A. Dosen

Dosen merupakan salah satu pekerjaan dibidang Pendidikan yang bertugas memdidik dan mentransformasikan ilmu pengetahuan, teknologi dan seni melalui tri dharma yaitu Pendidikan atau pengajaran, Penelitian dan Pengabdian [6].

## B. Algoritma C 4.5

Algoritma decision tree digunakan untuk membangun sebuah pohon keputusan yang mudah dimengerti, fleksibel, dan menarik karena dapat divisualisasikan dalam bentuk gambar Pohon keputusan adalah salah satu metode klasifikasi yang paling populer karena mudah untuk diinterpretasi oleh manusia. Pohon keputusan adalah model prediksi menggunakan struktur pohon atau struktur berhirarki [7]. Konsep dari pohon keputusan adalah mengubah data menjadi pohon keputusan dan aturan-aturan keputusan. Adapun langkah-langkah menjalankan penelitian menggunakan algoritma C 4.5 untuk membangun sebuah pohon keputusan adalah sebagai berikut:

1. Hitung jumlah data, jumlah data berdasarkan anggota atribut hasil dengan syarat tertentu. Untuk proses pertama syaratnya masih kosong.
2. Pilih atribut sebagai Node.
3. Bentuk cabang pada setiap anggota dari Node.
4. Memeriksa nilai entropy dari anggota Node ada yang bernilai nol. Jika ada, tentukan daun yang terbentuk. Jika seluruh nilai entropy anggota Node adalah nol, maka proses pun berhenti.
5. Jika ada anggota Node yang memiliki nilai entropy lebih besar dari nol, ulangi lagi proses dari awal dengan Node sebagai syarat sampai semua anggota dari Node bernilai nol [8].

## C. Naïve Bayes

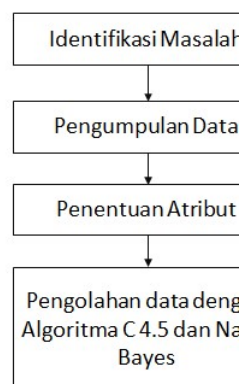
Naïve Bayes adalah salah satu pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu class. Naïve Bayes didasarkan pada teorema Bayes yang memiliki kemampuan klasifikasi serupa dengan decision tree dan neural network. Bayesian classification terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam database dengan data yang besar[9].

Dalam menggunakan metode Naïve Bayes, keuntungan yang bisa didapat adalah metode ini hanya membutuhkan jumlah data training yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. Karena diasumsikan sebagai variabel independen, maka hanya varian dari suatu variabel independen dalam sebuah kelas yang dibutuhkan untuk menentukan klasifikasi, bukan keseluruhan dari matriks kovarians. Naive Bayes merupakan teknik

prediksi berbasis probabilitas sederhana yang berdasarkan pada penerapan teorema Bayes dengan asumsi independensi yang kuat. Naïve Bayes didasarkan pada asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai output. Pengklasifikasian Naïve Bayes dilakukan dengan memilih probabilitas akhir (posterior) tertinggi dari masing-masing kelas.[10]

## II. METODE PENELITIAN

Penelitian ini disusun dengan kerangka penelitian sebagai berikut ini :



Gambar 1. Tahapan Penelitian

1. Identifikasi Masalah  
Pada tahapan ini semua permasalahan akan dilakukan penganalisaan oleh peneliti yang berkaitan dengan hal-hal yang mempengaruhi kinerja dosen
2. Pengumpulan Data  
Data yang dikumpulkan merupakan data yang berasal dari LPM Universitas Dharmawangsa dimana terdapat evaluasi dosen dengan mengisi form yang berkaitan dengan Indeks kerja dosen yang berisikan tentang pengajaran, penelitian dan pengabdian
3. Penentuan Atribut  
Setelah didapatkan data, maka dianalisa kemudian untuk mendapatkan atribut-atribut apa saja yang akan diolah dengan menggunakan Algoritma C 4.5 dan Naïve bayes
4. Pengolahan data dengan Algoritma C 4.5 dan Naïve Bayes.  
Pada tahapan ini dilakukan pemilihan dosen terbaik dengan menggunakan Algoritma C4.5 dan Naïve Bayes yang nantinya akan dibandingkan
5. Penentuan Hasil  
Penentuan hasil tentang komparasi Algoritma C 4.5 dan Naïve Bayes dilakukan dengan pengujian Confusion Matrix pada Rapid Miner untuk

mendapatkan Algoritma mana yang memiliki akurasi paling tinggi dan tepat.

### III. HASIL DAN PEMBAHASAN

Dari penelitian yang telah dilakukan didapatkan hasil seperti berikut ini :

#### 1. Data

Data yang digunakan adalah data IKD Dosen yang diadakan tiap semester di Universitas Dharmawangsa. Sebanyak 30 data diproses dengan metode Algoritma C 4.5 dan Naïve Bayes untuk mendapatkan mana Algoritma Klasifikasi yang paling akurat untuk Pemilihan Dosen Terbaik di Universitas Dharmawangsa.

Data yang akan dijadikan diolah adalah data sampel sebanyak 30 data dari keseluruhan data yang didapatkan.

Tabel 1. Data Sampel

No	NIDN	Nama	Pengajaran	Penelitian	Pengabdian	Pe
1	0127049401	Afni Nia Sari, S.Kom, M.Kom	BAIK	BAIK	CUKUP	S/ B.
2	0124118601	Al Firah, M.Si	BAIK	CUKUP	BAIK	S/ B.
3	0107127901	Amru Yasir, S.Kom, M.Kom	BAIK	BAIK	CUKUP	B.
4	0124078701	Ananda Hadi Elyas, S.Kom, M.Kom	SANGAT BAIK	CUKUP	BAIK	CI
...	.....	.....	.....	.....	.....	.....
28	0110069002	J. Prayoga, S.Kom, M.Kom	SANGAT BAIK	CUKUP	BAIK	CI
29	0120069401	Jovi Antares,				

#### 2. Hasil Pengolahan Data Menggunakan Algoritma C 4.5

Algoritma C 4.5 melakukan penyelesaian masalah dengan sistematis dengan cara membentuk decision tree dengan langkah-langkah berikut :

- Melakukan pemilihan atribut yang akan dijadikan akar
- Membuat cabang dari atribut untuk masing-masing record.
- Membagi kasus kedalam cabang.
- Melakukan perulangan untuk masing-masing cabang sampai semua kasus di setiap cabang menghasilkan keputusan yang sudah sesuai

Untuk melakukan perhitungan nilai dari atribut yang akan dijadikan akar, dicari nilai gain yang paling tinggi. Nilai ini akan dijadikan akar (root) pohon keputusan. Rumus untuk mencari nilai gain adalah :

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

Keterangan :

A : Atribut didalam S

n : Jumlah partisi A

|Si| : Jumlah kasus pada partisi ke-i

|S| : Jumlah kasus

Sedangkan untuk mencari nilai entropy rumus yang digunakan adalah sebagai berikut :

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

Keterangan :

S : Himpunan Kasus

A : Atribut dalam S

n : Jumlah partisi S

Pi : Proporsi dari Si terhadap S

Tabel 2. Atribut dan Nilai

Atribut	Nilai
Pengajaran	Sangat Baik, Baik, Cukup, Kurang
Penelitian	Baik, Cukup, Kurang
Pengabdian	Baik, Cukup, Kurang
Penunjang	Sangat Baik, Baik, Cukup, Kurang
Hasil	Sangat Baik, Baik, Cukup, Kurang

- Penentuan sampel kasus data dari hasil penentuan jenis kelulusan mahasiswa diambil sebanyak 30 data yang sudah disiapkan, untuk sampel jumlah kasus pada penelitian ini dapat lihat pada tabel berikut :

Tabel 3. Data Sampel Yang Akan Diproses

Pengajaran	Penelitian	Pengabdian	Penunjang
BAIK	BAIK	CUKUP	SANGAT BAIK
BAIK	CUKUP	BAIK	SANGAT BAIK
BAIK	BAIK	CUKUP	BAIK
SANGAT BAIK	CUKUP	BAIK	CUKUP
.....	.....	.....	.....
.....	.....	.....	.....
.....	.....	.....	.....
.....	.....	.....	.....
SANGAT			

- Perhitungan Entropy

Untuk perhitungan entropy total dapat dilihat dibawah ini.

Entropy (Sangat Baik, Baik, Cukup dan Kurang Baik)

$$\begin{aligned}
 &= \left( - \left( \frac{\text{Jumlah Sangat Baik}}{\text{Jumlah Kasus}} * \log_2 \frac{\text{Jumlah Sangat Baik}}{\text{Jumlah Kasus}} \right) \right) \\
 &+ \left( - \left( \frac{\text{Jumlah Baik}}{\text{Jumlah Kasus}} * \log_2 \frac{\text{Jumlah Baik}}{\text{Jumlah Kasus}} \right) \right) \\
 &+ \left( - \left( \frac{\text{Jumlah Cukup}}{\text{Jumlah Kasus}} * \log_2 \frac{\text{Jumlah Cukup}}{\text{Jumlah Kasus}} \right) \right) \\
 &+ \left( - \left( \frac{\text{Jumlah Kurang Baik}}{\text{Jumlah Kasus}} * \log_2 \frac{\text{Jumlah Kurang Baik}}{\text{Jumlah Kasus}} \right) \right) \\
 \\
 \text{Total Entropy} &= \left( - \left( \frac{2}{30} * \log_2 \frac{2}{30} \right) + \left( - \left( \frac{14}{30} * \log_2 \frac{14}{30} \right) \right) \right) \\
 &+ \left( - \left( \frac{6}{30} * \log_2 \frac{6}{30} \right) \right) + \left( - \left( \frac{8}{30} * \log_2 \frac{8}{30} \right) \right) \\
 &= 0,260459 + 0,513117 + 0,464386 \\
 &+ 0,508504 \\
 &= 1,746466
 \end{aligned}$$

Dilakukan juga untuk perhitungan entropy Pengajaran (Sangat Baik, Baik, Cukup dan Kurang), entropy Penelitian (Baik, Cukup, Kurang), Entropy Pengabdian (Baik, Cukup, Kurang), dan Penunjang (Sangat Baik, Baik, Cukup, Kurang) dengan menggunakan rumus yang sama.

c. Perhitungan Gain

Untuk perhitungan ini, penulis mengambil sampel untuk perhitungan gain pada kriteria pengajaran, yang mana atributnya adalah Sangat Baik, Baik, Cukup, dan Kurang Baik.

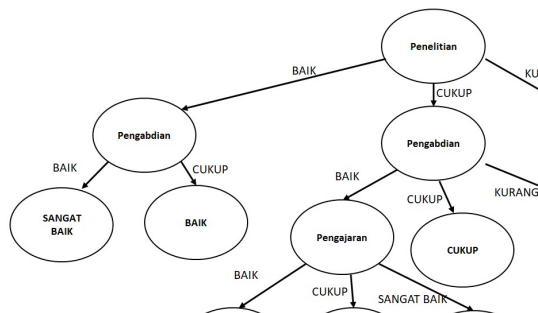
$$\begin{aligned}
 \text{Gain (Pengajaran)} &= \text{Total Entropy} - \left( \frac{\text{Sangat Baik}}{\text{Jumlah Kasus}} * \text{Entropy Sangat Baik} \right) \\
 &+ \left( \frac{\text{Sangat Baik}}{\text{Jumlah Kasus}} * \text{Entropy Baik} \right) \\
 &+ \left( \frac{\text{Cukup}}{\text{Jumlah Kasus}} * \text{Entropy Cukup} \right) \\
 &+ \left( \frac{\text{Kurang}}{\text{Jumlah Kasus}} * \text{Entropy Kurang} \right) \\
 \\
 \text{Gain (Pengajaran)} &= 1,746466 - \left( \left( \frac{6}{30} * 1,459148 \right) + \left( \frac{14}{30} * 0,946373 \right) \right) \\
 &+ \left( \frac{10}{30} * 0,970951 \right) + \left( \frac{0}{30} * 0 \right) + \left( \frac{3}{30} * 0 \right) \\
 &= 0,689345
 \end{aligned}$$

Gain juga dilakukan untuk atribut lainnya, sehingga nanti didapatkan nilai paling besar yang berfungsi sebagai root dari pohon keputusan yang akan dibuat

Tabel 4. Perhitungan Nilai dan Entropy

TOTAL		JUMLAH	SANGAT BAIK	BAIK	CUKUP	KURANG BAIK	E
		30	2	14	6	8	
PENGAJARAN	SANGAT BAIK	6	2	3	1	0	
	BAIK	14	0	11	1	2	
	CUKUP	10	0	0	4	6	
	KURANG	0	0	0	0	0	
PENELITIAN	BAIK	10	2	7	1	0	
	CUKUP	14	0	7	5	2	
	KURANG	6	0	0	0	6	
PENGABDIAN	BAIK	12	2	8	2	0	
	CUKUP	11	0	6	3	2	
	KURANG	7	0	0	1	6	

Setelah dilakukan pencarian Gain tertinggi maka didapatkan pohon keputusan seperti yang ada pada gambar dibawah ini.



Gambar 1. Pohon Keputusan

Dilihat dari pohon keputusan dan penghitungan Dosen Terbaik menggunakan Algoritma C4.5 dapat dilihat yang menjadi akar adalah Penelitian, didapatkan delapan aturan dari pencarian tersebut.

B. Hasil Pengolahan Data Menggunakan Naïve Bayes

Pencarian selanjutnya dilakukan pencarian hasil Dosen terbaik dengan menggunakan Naïve Bayes, dimana rumusnya adalah :

$$P(A | B) = P(B | A)P(A)P(B)$$

Keterangan:

P(A | B) : Probabilitas A terjadi dengan bukti bahwa B telah terjadi (probabilitas superior)

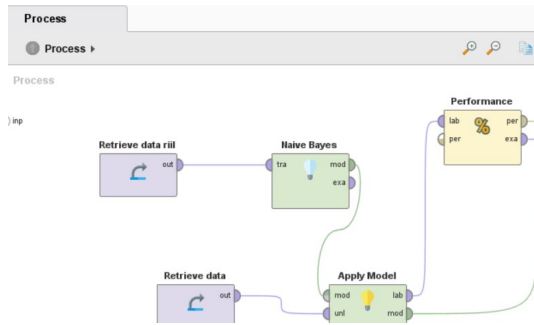
P(B | A) : Probabilitas B terjadi dengan bukti bahwa A telah terjadi

P(A) : Peluang terjadinya A

P(B) : Peluang terjadinya B

Untuk mempermudah proses menganalisa data menggunakan algoritma naïve bayes maka perlu dilakukan data cleaning atau preprocessing, dengan tujuan menghilangkan noise atau missing value. Dari preprocessing dari data sebanyak 113, sebanyak 30 data dijadikan dataset uji. Data transformation dilakukan

dengan memberikan inialisasi terhadap data yang memiliki nilai nominal menjadi bernilai numerik.



Gambar 2. Pengolahan data Naive Bayes pada Rapid Miner

Row No.	HASIL	prediction(H...	confidence_...	confidence_...	confidence_...	confidence_...	PENGAJARAN	PENELITIAN
1	BAIK	BAIK	1.000	0.000	0.000	0.000	BAIK	BAIK
2	CUKUP	CUKUP	0.001	0.999	0.000	0.000	SANGAT BAIK	CUKUP
3	BAIK	BAIK	1.000	0.000	0.000	0.000	BAIK	CUKUP
4	CUKUP	CUKUP	0.001	0.999	0.000	0.000	CUKUP	CUKUP
5	SANGAT BAIK	SANGAT BAIK	0.149	0.000	0.851	0.000	SANGAT BAIK	BAIK
6	KURANG BAIK	KURANG BAIK	0.000	0.000	0.000	1.000	CUKUP	KURANG
7	KURANG BAIK	KURANG BAIK	0.000	0.000	0.000	1.000	CUKUP	KURANG
8	BAIK	BAIK	0.957	0.043	0.000	0.000	BAIK	BAIK
9	BAIK	BAIK	0.639	0.361	0.000	0.000	SANGAT BAIK	CUKUP
10	KURANG BAIK	KURANG BAIK	0.000	0.000	0.000	1.000	BAIK	KURANG
11	BAIK	BAIK	0.904	0.096	0.000	0.000	BAIK	CUKUP
12	KURANG BAIK	KURANG BAIK	0.000	0.001	0.000	0.999	CUKUP	CUKUP
13	BAIK	BAIK	0.639	0.361	0.000	0.000	SANGAT BAIK	CUKUP
14	BAIK	BAIK	0.933	0.067	0.000	0.000	BAIK	BAIK
15	BAIK	BAIK	0.991	0.009	0.000	0.000	SANGAT BAIK	BAIK

Gambar 3. Hasil Penghitungan Confidence di RapidMiner

4. Pengujian Hasil

Untuk pengujian hasil dicari dengan cara menguji kedua Algoritma Klasifikasi ini dengan menggunakan Confusion Matrix untuk menilai mana Algoritma yang lebih akurat diantara Algoritma C4.5 atau Naive Bayes untuk Pemilihan Dosen Terbaik.

- a. Confusion Matrix Algoritma C 4.5  
Setelah dilakukan proses dengan data uji didapatkan akurasi sebesar 83,33% .

True \ Description	BAIK	CUKUP	SANGAT BAIK	KURANG BAIK
BAIK	3	0	0	0
CUKUP	0	1	0	1
SANGAT BAIK	0	0	0	0
KURANG BAIK	0	0	0	1

Gambar 4. Confusion Matrix Algoritma C 4.5

- b. Confusion Matrix Naive Bayes  
Pada Algoritma Naive Bayes didapatkan hasil akurasi yang cukup tinggi setelah dilakukan proses

True \ Description	BAIK	CUKUP	SANGAT BAIK	KI
BAIK	14	1	0	0
CUKUP	0	5	0	0

Gambar 4. Confusion Matrix Naive Bayes

Dari pengujian yang telah dilakukan dapat disimpulkan Naive Bayes memiliki akurasi lebih tinggi dengan nilai 96,67% dibandingkan Algoritma C 4.5 yang memiliki persentase 83,33%

IV. KESIMPULAN

Berdasarkan penelitian, implementasi dan pengujian, maka dapat diambil kesimpulan bahwa pemilihan dosen terbaik memiliki akurasi tinggi dengan tingkat persentase 96,67% dibandingkan pengolahan data dengan Algoritma C 4.5 yaitu sebesar 83,33%.

Penulis berharap penelitian ini dapat dikembangkan lagi dengan menggunakan algoritma klasifikasi lainnya.

DAFTAR PUSTAKA

- [1] V. Putratama and R. Andarsyah, "Penentuan Dosen Terbaik Menggunakan Metode Simple Additive Weighting (SAW) Di Politeknik Pos Indonesia," *Competitive*, vol. 15, no. 2, pp. 144-154, 2020, doi: 10.36618/competitive.v15i2.962.
- [2] Y. Y. Moh. Ilbad Dzulfadli, Eka Larasati Amalia, "Sistem Penilaian Dosen Berprestasi Menggunakan Metode Waspa (Studi Kasus Politeknik Negeri Malang)," *Siap 2020*, pp. 389-394, 2020.
- [3] M. Assiddiq, "Peguruang: Conference Series," vol. 2, no. April, 2020.
- [4] S. Amal, "Perancangan Sistem Pendukung Keputusan Pemilihan Dosen Terbaik Menggunakan Metode Saw Berbasis Web (Studi Kasus Politeknik Ganesha): 25-34," *REMIK (Riset dan E-Jurnal Manaj. Inform.*

- ..., vol. 2, no. 1, 2017, [Online]. Available: <https://www.jurnal.polgan.ac.id/index.php/remik/article/view/10859>.
- [5] W. D. Puspitasari and D. K. Ilmi, "Sistem Pendukung Keputusan Pemilihan Dosen Berprestasi Menggunakan Metode Analytical Hierarchy Process (Ahp)," *ANTIVIRUS J. Ilm. Tek. Inform.*, vol. 10, no. 2, pp. 56–68, 2016, doi: 10.30957/antivirus.v10i2.163.
- [6] A. Yuliana and D. B. Pratomo, "Algoritma Decision Tree (C4.5) Untuk Memprediksi Kepuasan Mahasiswa Terhadap Kinerja Dosen Politeknik Tedc Bandung," *Semnasinotek 2017*, pp. 377–384, 2017.
- [7] N. Tahun, P. Algoritma, C. D. Sistem, S. Kasus, and D. I. Kabupaten, "Penerapan algoritma c 4.5 dalam sistem pendukung keputusan evaluasi kinerja fasilitator pamsimas (studi kasus di kabupaten kampar)," vol. 1, no. 1, pp. 20–28, 2020.
- [8] Z. Gustiana, W. Satria, and J. Simon, "Penerapan Algoritma C 4.5 Pada Pengaruh Iklan Online Terhadap Minat Beli Konsumen di Masa Pandemic Covid-19," *J. SAINTIKOM (Jurnal Sains Manaj. Inform. dan Komputer)*, vol. 20, no. 2, p. 91, 2021, doi: 10.53513/jjis.v20i2.3751.
- [9] H. F. Putro, R. T. Vulandari, and W. L. Y. Saptomo, "Penerapan Metode Naive Bayes Untuk Klasifikasi Pelanggan," *J. Teknol. Inf. dan Komun.*, vol. 8, no. 2, 2020, doi: 10.30646/tikomsin.v8i2.500.
- [10] W. P. Nurmayanti, "Penerapan Naive Bayes dalam Mengklasifikasikan Masyarakat Miskin di Desa Lepak," *Geodika J. Kaji. Ilmu dan Pendidik. Geogr.*, vol. 5, no. 1, pp. 123–132, 2021, doi: 10.29408/geodika.v5i1.3430.