

## KLASIFIKASI FAKTOR PENYEBAB RENDAHNYA MINAT MAHASISWA DALAM MENULIS ARTIKEL ILMIAH

Friskila Parhusip<sup>1</sup>, Agus Perdana Windarto<sup>2</sup>, Irfan Sudahri Damanik<sup>3</sup>, Eka Irawan<sup>4</sup>, Ilham Syahputra Saragih<sup>5</sup>

<sup>1,2,3,4,5</sup>Program Studi Sistem Informasi, STIKOM Tunas Bangsa,  
Pematangsiantar, Indonesia

e-mail: [friskilaparhusip98@gmail.com](mailto:friskilaparhusip98@gmail.com)<sup>1</sup>, [agus.perdana@amiktunasbangsa.ac.id](mailto:agus.perdana@amiktunasbangsa.ac.id)<sup>2</sup>,  
[irfansudahridamanik@stikomtb.ac.id](mailto:irfansudahridamanik@stikomtb.ac.id)<sup>3</sup>, [eka.irawan@amiktunasbangsa.ac.id](mailto:eka.irawan@amiktunasbangsa.ac.id)<sup>4</sup>  
[ilham@amiktunasbangsa.ac.id](mailto:ilham@amiktunasbangsa.ac.id)<sup>5</sup>

Received : September, 2020

Accepted : November, 2020

Published : October, 2021

### Abstract

*Scientific articles are scientific writings that can be in the form of review articles or research articles from research reports that are rewritten by the authors to be published in reputable scientific journals. This is one of the main things that becomes a reference for students in higher education during their studies and at the end of their studies. then the problem that must be faced is how to classify the factor of low student interest and being able to increase student interest in writing scientific articles. The factors of low student interest in writing scientific articles are the level of understanding, motivation, difficulty level, time, and infrastructure. The data used comes from the data of students of the 2016 class of information systems study program. It has been obtained the classification of the factor of low student interest in writing scientific articles. The factor that has top priority on the factor of low student interest is the infrastructure factor. The experimental and evaluation results show that the Decision Tree C4.5 Algorithm is accurately applied to classify the factors of low student interest in writing scientific articles with an accuracy rate of 100%.*

**Keywords:** *Scientific Articles, Colleges, Students, Data Mining, C4.5 Algorithm.*

### Abstrak

Artikel ilmiah merupakan tulisan yang ilmiah dapat berbentuk artikel ulasan (*review article*) maupun artikel penelitian (*research article*) dari laporan hasil penelitian yang ditulis kembali oleh para penulisnya untuk dipublikasikan dalam jurnal ilmiah bereputasi. Hal ini adalah salah satu hal utama yang menjadi acuan bagi mahasiswa di perguruan tinggi selama menjalani study maupun berakhirnya study. maka masalah yang harus dihadapi adalah bagaimana mengklasifikasi faktor rendahnya minat mahasiswa dan mampu meningkatkan minat mahasiswa dalam menulis artikel ilmiah. Faktor rendahnya minat mahasiswa dalam menulis artikel ilmiah yaitu Tingkat pemahaman, Motivasi, Tingkat Kesulitan, Waktu, dan Sarana Prasarana. Data yang digunakan berasal dari data mahasiswa prodi sistem informasi angkatan 2016. Telah didapat klasifikasi faktor rendahnya minat mahasiswa dalam menulis artikel ilmiah faktor yang memiliki prioritas utama terhadap faktor rendahnya minat mahasiswa adalah faktor sarana prasarana. Hasil eksperimen dan evaluasi menunjukkan bahwa Algoritma Decision Tree C4.5 akurat diterapkan untuk mengklasifikasi faktor rendahnya minat mahasiswa dalam menulis artikel ilmiah dengan tingkat akurasi 100%.

**Kata Kunci:** *Artikel Ilmiah, Perguruan Tinggi, Mahasiswa, Data Mining, Algoritma C4.5.*

## 1. PENDAHULUAN

Artikel ilmiah merupakan tulisan yang ilmiah dapat berbentuk artikel ulasan (*review article*) maupun artikel penelitian (*research article*) dari laporan hasil penelitian yang ditulis kembali oleh para penulisnya untuk dipublikasikan dalam jurnal ilmiah bereputasi [1]. Dalam peningkatan kualitas institusi perguruan tinggi, institusi mengharapkan dosen dan mahasiswa memberikan kontribusi dalam menulis artikel ilmiah. Pada kenyataannya, seperti halnya kegiatan membaca, kegiatan menulis dikalangan mahasiswa STIKOM Tunas Bangsa belum membudaya. Hal tersebut tampak pada kecenderungan mahasiswa lebih menyukai berbicara daripada menulis. Sehingga perlu adanya metode untuk menyelesaikan masalah penentuan faktor rendahnya minat mahasiswa dalam menulis artikel ilmiah.

Data Mining[2] merupakan suatu proses penggalian data atau penyaringan data dengan memanfaatkan kumpulan data dengan ukuran yang cukup besar melalui serangkaian proses untuk mendapatkan informasi yang berharga dari data tersebut. Data Mining dapat diterapkan pada berbagai bidang yang mempunyai sejumlah data. Menurut Daryl Pregibon bahwa "Data mining adalah campuran dari statistic, kecerdasan buatan, dan riset basis data" yang masih berkembang.[3] Metode yang digunakan dalam skripsi ini adalah Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Didalam

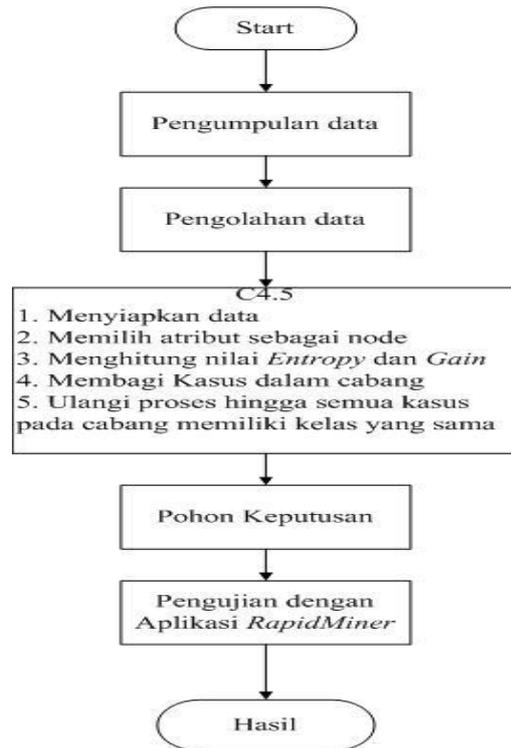
algoritma C4.5 ini, pohon-pohon keputusan yang dibentuk berdasarkan kriteria-kriteria pembentuk keputusan. Didalam penggunaan algoritma C4.5 ada 2 (dua) elemen yang harus dipahami yaitu Entropy dan Gain[4]. *Decision Tree* adalah salah satu pendekatan yang paling populer dalam klasifikasi. *Decision Tree* dipakai diberbagai disiplin ilmu seperti bidang medis, statistik, *machine learning*, *pattern recognition*, dan data mining. Para peneliti telah mengembangkan berbagai metode *Decision Tree* dalam menyelesaikan dataset yang ada.[5]

Dari paparan diatas, maka pada penelitian ini bertujuan mengetahui Klasifikasi faktor rendahnya minat mahasiswa dalam menulis artikel ilmiah melalui proses algoritma *Decision Tree* C4.5 dalam mengklasifikasi faktor setiap kelas faktor serta mengetahui tingkat akurasi yang dihasilkan setiap algoritma menggunakan software rapidminer.

## 2. METODE PENELITIAN

### 2.1. Rancangan Penelitian

Penelitian ini dilakukan dengan menggunakan metode penelitian Studi Literatur dengan pokok penelitian adalah rendahnya minat mahasiswa dalam menulis artikel ilmiah. Data diambil dari kusioner yang dibagi kepada mahasiswa STIKOM Tunas Bangsa. Adapun langkah-langkah kerja dari penelitian ini sebagai berikut:



Gambar 1. Alur Kerja Penelitian

Keterangan Kerangka Kerja :

1. Pengumpulan data  
Pengumpulan data adalah mengumpulkan data-data yang akan digunakan dalam proses algoritma klasifikasi C4.5. Untuk data keseluruhan dapat dilihat pada data yang telah terlampir.
2. Pengolahan data  
Proses pengolahan data dilakukan secara manual dan kemudian dilakukan pengujian menggunakan *RapidMiner* versi 5.3.
3. Algoritma C4.5  
Algoritma C4.5 merupakan algoritma yang digunakan dalam pembuatan pola *Decision Tree* berdasarkan perhitungan dari data yang diperoleh.
4. Pohon Keputusan  
Pohon keputusan adalah hasil dari proses perhitungan *entropy* dan *information gain*, setelah perhitungan berulang-ulang sampai semua atribut pohon memiliki kelas dan tidak bisa lagi dilakukan proses perhitungan.

5. Pengujian dengan Aplikasi *RapidMiner*  
Pengujian aplikasi dilakukan untuk memastikan bahwa tidak ada kekeliruan dalam penerapan algoritma ke dalam aplikasi.
6. Hasil  
Penelitian ini akan menghasilkan *rules* atau aturan dalam penentuan faktor rendahnya minat mahasiswa dalam menulis artikel ilmiah

## 2.2. Sumber Data

Data yang digunakan dalam penelitian ini adalah data kuisisioner yang dibagi secara acak kepada mahasiswa STIKOM Tunas Bangsa Pematangsiantar. Data yang digunakan adalah kuisisioner yang dibagi kepada Mahasiswa STIKOM Tunas Bangsa Prodi sistem informasi angkatan 2016. Berikut adalah data yang dipakai dalam penelitian ini dapat dilihat pada tabel

Tabel 1. Data hasil kuisioner

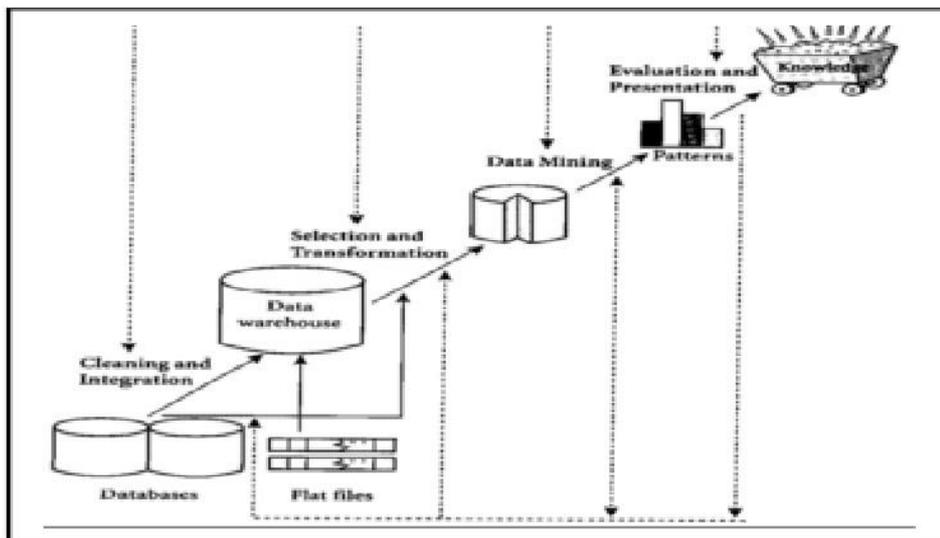
Responden	Tingkat Pemahaman	Motivasi	Tingkat Kesulitan	Waktu	Sarana dan Prasarana	HASIL
Responden1	Tidak Paham	Rendah	Sangat Sulit	Tidak Bekerja	Ada	Tidak Minat
Responden2	Tidak Paham	Rendah	Sangat Sulit	Bekerja	Tidak Ada	Tidak Minat
Responden3	Paham	Tinggi	Mudah	Bekerja	Ada	Minat
Responden4	Tidak Paham	Rendah	Sangat Sulit	Bekerja	Tidak Ada	Tidak Minat
Responden5	Tidak Paham	Rendah	Sangat Sulit	Bekerja	Tidak Ada	Tidak Minat
Responden6	Paham	Tinggi	Mudah	Bekerja	Tidak Ada	Minat
Responden7	Sangat Paham	Sangat Tinggi	Mudah	Bekerja	Ada	Minat
Responden8	Sangat Paham	Tinggi	Mudah	Tidak Bekerja	Ada	Minat
Responden9	Tidak Paham	Rendah	Sangat Sulit	Bekerja	Tidak Ada	Tidak Minat
...	...	....	...	...	...	...
Responden100	Tidak Paham	Rendah	Sangat Sulit	Bekerja	Tidak Ada	Tidak Minat

Sumber : Mahasiswa STIKOM Tunas Bangsa

### 2.3. Data Mining

Data Mining[6][3] merupakan serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data. Salah satu metode data mining adalah klasifikasi yaitu proses pencarian model klasifikasi yang dapat membedakan objek label kelasnya[7]

Karena data mining [8][9] adalah suatu rangkaian proses, data mining dapat dibagi menjadi beberapa tahap. Tahap-tahap tersebut bersifat interaktif di mana pemakai terlibat langsung atau dengan perantaraan knowledge base. Tahap-tahap ini diilustrasikan pada Gambar 1.



Gambar 2. Tahap-tahap data mining

### 2.3. Algoritma C4.5

Algoritma C4.5 [10] merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Sedang pohon keputusan dapat diartikan suatu cara untuk memprediksi atau mengklarifikasi yang sangat kuat. Pohon

keputusan merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal.[11] Adapun tahapan dalam membuat sebuah pohon keputusan dalam algoritma C4.5[5] yaitu:

1. Mempersiapkan data *training*. Data *training* biasanya diambil dari data histori yang pernah terjadi sebelumnya atau disebut data masa lalu dan sudah dikelompokkan dalam kelas-kelas tertentu.

2. Menghitung akar dari pohon. Akar akan diambil dari atribut yang akan terpilih, dengan cara menghitung nilai *gain* dari masing-masing atribut, nilai *gain* yang paling tinggi yang akan menjadi akar pertama. Sebelum menghitung nilai *gain* dari atribut, hitung dahulu nilai *entropy*. Untuk menghitung nilai *entropy* digunakan rumus :

$$Entr(S) = \sum - p_i \log_2 p_i \dots \dots \dots (1)$$

Keterangan:

- $S$  = Himpunan kasus
- $n$  = Jumlah partisi  $S$
- $p_i$  = Proporsi  $S_i$  terhadap  $S$

Kemudian menghitung nilai *gain* menggunakan rumus :

$$G(S,A) = Entropy(S) - \sum |S_i| / |S| * n_i = 1 Entropy(S_i) \dots \dots (2)$$

Keterangan:

- $S$  = Himpunan kasus
- $A$  = Atribut
- $n$  = Jumlah partisi atribut  $A$
- $|S_i|$  = Proporsi  $S_i$  terhadap  $S$
- $|S|$  = Proporsi kasus dalam  $S$

3. Ulangi langkah ke-2 dan langkah ke-3 hingga semua *record* terpartisi.

4. Proses partisi pohon keputusan akan berhenti saat:

1. Semua *record* dalam simpul  $N$  mendapat kelas yang sama.
2. Tidak ada atribut di dalam *record* yang dipartisi lagi.
3. Tidak ada *record* di dalam cabang yang kosong.

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Jenis Kelas Data

(judul subbab Calibri, 11, Bold)

Tabel 2. Jenis kelas data

Atribut	Nama Field	Kelas Data yang Digunakan
C1	Tingkat Pemahaman	Sangat Paham. Paham, Tidak Paham
C2	Motivasi	Sangat Tinggi, Tinggi, Rendah.
C3	Tingkat Kesulitan	Sangat Sulit, Sulit, Mudah
C4	Waktu	Bekerja, Tidak Bekerja
C5	Sarana dan Prasarana	Ada, Tidak Ada

Keterangan Jenis Kelas Data

- 1) Tingkat Pemahaman (C1), mahasiswa yang tidak paham dalam menulis artikel ilmiah maka tidak minat dalam menulis artikel ilmiah.
- 2) Motivasi (C2), mahasiswa yang kurang mendapat motivasi dalam menulis artikel ilmiah maka tidak minat dalam menulis artikel ilmiah.
- 3) Tingkat Kesulitan (C3), faktor Tingkat kesulitan akan membuat mahasiswa tidak minat dalam menulis artikel ilmiah, yang menurut nya membuat suatu artikel ilmiah itu sulit.
- 4) Waktu (C4), Mahasiswa yang tidak memiliki waktu dalam menulis artikel ilmiah seperti mahasiswa yang kuliah sambil bekerja , cenderung tidak minat dalam menulis artikel ilmiah.

- 5) Sarana dan Prasana (C5), Mahasiswa yang tidak memiliki saranana dan prasana

#### 3.2 Pembahasan

##### 3.2.1. Proses Algoritma C4.5

1. Mempersiapkan data *training*. Data *training* biasanya diambil dari data histori yang pernah terjadi sebelumnya atau disebut data masa lalu dan sudah dikelompokkan dalam kelas-kelas tertentu.
2. Menghitung jumlah kasus, jumlah kasus untuk keputusan Minat, jumlah kasus untuk keputusan Tidak Minat, dan *Entropy* dari semua kasus dan kasus yang dibagi berdasarkan atribut. Kolom *Entropy* pada baris total dihitung seperti berikut :

$$Entropy(Total) = \left(-\frac{42}{100} \times \log_2\left(\frac{42}{100}\right)\right) + \left(-\frac{57}{100} \times \log_2\left(\frac{57}{100}\right)\right)$$

$$Entropy(Total) = 0,985815037$$

Entropy Tingkat Pemahaman

$$Entropy(Sangat Paham) = \left(-\frac{19}{19} \times \log_2\left(\frac{19}{19}\right)\right) + \left(-\frac{0}{19} \times \log_2\left(\frac{0}{19}\right)\right) \dots\dots dst$$

sampai Entropy Sarana dan Prasarana

dari semua kasus dan kasus yang dibagi berdasarkan atribut.

Kolom *Gain* pada baris Tingkat Pemahaman dihitung seperti berikut :

$$Gain(Tingkat Pemahaman) = 0,985815037$$

$$\left(\frac{19}{100} \times 0\right) + \left(\frac{57}{100} \times 0,845350937\right) + \left(\frac{59}{100} \times 0,572580466\right)$$

$$Gain = 0,462015356$$

.....dst

Sampai Gain sarana dan Prasarana

3. Menghitung jumlah kasus, jumlah kasus untuk keputusan Minat, jumlah kasus untuk keputusan Tidak Minat, dan *Gain*

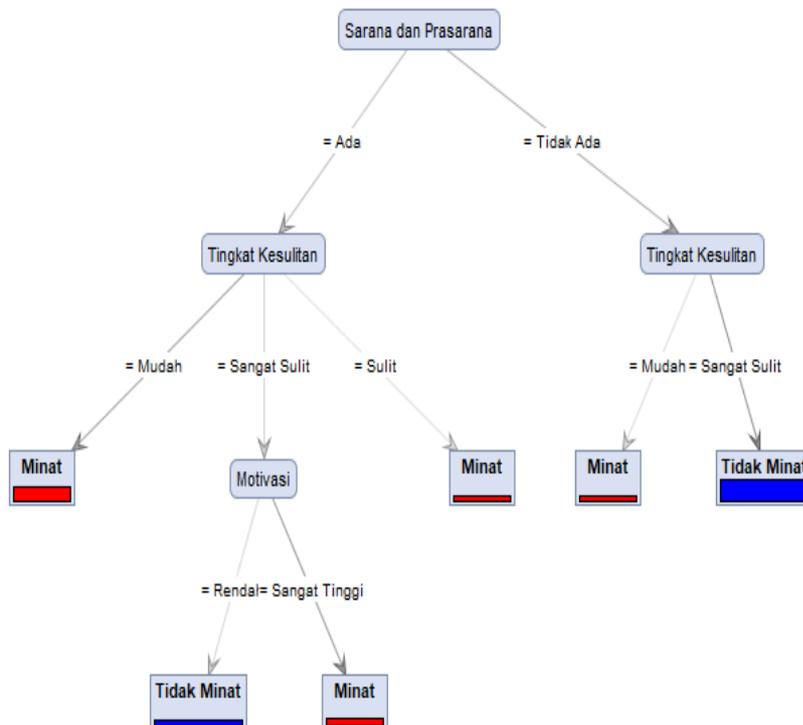
Adapun hasil *preprocessing* data mahasiswa yang di dapat adalah sebagai berikut:

Tabel 3. Node Akhir

Node 1.2.		S	S1	S2	Entropy	Information Gain
Sarana dan Prasarana- Tidak Ada		61	5	56	0,409073139	
Tingkat Pemahaman						0,229821789
	Sangat Paham	0	0	0	0	
	Paham	11	5	6	0,994030211	
	Tidak Paham	50	0	50	0	
Motivasi						0,283901467
	Sangat Tinggi	0	0	0	0	
	Tinggi	8	5	3	0,954434003	
	Rendah	53	0	53	0	
Tingkat Kesulitan						0,409073139
	Sangat Sulit	56	0	56	0	
	Sulit	0	0	0	0	
	Mudah	5	5	0	0	
Waktu						0,012805364
	Bekerja	55	5	50	0,439496987	
	Tidak Bekerja	6	0	6	0	

Hasil *preprocessing* yang didapat kemudian diolah kedalam *Decision Tree* menggunakan *software Rapidminer* untuk mengetahui faktor

rendahnya minat mahasiswa dalam menuliskan artikel ilmiah, adapun gambar pohon keputusan sebagai berikut :



Gambar 2. Pohon Keputusan Rapidminer

### 3.2.2. Hasil Pengujian

Pada penelitian yang didukung hasil pemrosesan data menggunakan *software RapidMiner* didapat hasil akurasi penerapan

algorithm C4.5 yaitu 100%. Untuk melihat nilai akurasi dapat dilihat pada Gambar 3 dibawah ini:

accuracy: 100.00%			
	true Tidak Minat	true Minat	class precision
pred. Tidak Minat	17	0	100.00%
pred. Minat	0	13	100.00%
class recall	100.00%	100.00%	

Gambar 3. Nilai Akurasi algoritma C4.5

## 4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, didapatkan beberapa kesimpulan yaitu :

1. Telah didapat klasifikasi faktor rendahnya minat mahasiswa dalam menulis artikel ilmiah di STIKOM Tunas Bangsa Pematangsiantar. Variabel yang memiliki prioritas utama terhadap faktor rendahnya minat mahasiswa adalah faktor sarana prasarana.
2. pengolahan data menggunakan *software RapidMiner* didapat nilai akurasi sebesar 100 %. Sesuai dengan ketentuan tersebut maka hasil perhitungan manual dengan

pengujian *RapidMiner* akan mendapat hasil yang sama. Sehingga pengujian dengan *RapidMiner* dapat dikatakan berhasil

## DAFTAR PUSTAKA

- [1] N. Rofiqo, A. P. Windarto, and A. Wanto, "Penerapan Metode VIKOR Pada Faktor Penyebab Rendahnya Minat Mahasiswa Dalam Menulis Artikel Ilmiah," *Semin. Nas. Sains Teknol. Informasi (SENSASI)*, no. Mcdm, pp. 228–237, 2018.
- [2] R. W. Sari and D. Hartama, "Data

- Mining : Algoritma K-Means Pada Pengelompokan Wisata Asing ke Indonesia Menurut Provinsi,” pp. 322–326, 2018.
- [3] H. Sulastri and A. I. Gufroni, “Penerapan Data Mining Dalam Pengelompokan Penderita Thalassaemia,” *J. Teknol. dan Sist. Inf.*, vol. 3, no. 2, p. 299, 2017.
- [4] P. Algoritma and C. Untuk, “Penerapan algoritma c4.5 untuk prediksi penggunaan jenis kontrasepsi berbasis web,” vol. 04, no. 02, pp. 174–187, 2017.
- [5] M. Hasyim and U. Y. Pasuruan, “PENANGANAN DATA MISSING VALUE PADA KUALITAS PRODUKSI JAGUNG,” vol. 2, no. 2, 2019.
- [6] A. P. Windarto, P. Studi, S. Informasi, and D. Mining, “Penerapan Data Mining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering,” vol. 16, no. 4, pp. 348–357, 2017.
- [7] S. Lorena, B. Ginting, and R. P. Trinanda, “TEKNIK DATA MINING MENGGUNAKAN METODE BAYES CLASSIFIER UNTUK OPTIMALISASI PENCARIAN Oleh :,” vol. d.
- [8] Andri, Y. N. Kunang, and S. Murniati, “Implementasi Teknik Data Mining Untuk Memprediksi Tingkat Kelulusan Mahasiswa Pada Universitas Bina Darma Palembang,” *Semin. Nas. Inform. 2013 (semnasIF 2013)*, vol. 2013, no. A, pp. 56–63, 2013.
- [9] S. K. Yadav and S. Pal, “Data Mining : A Prediction for Performance Improvement of Engineering Students using Classification,” *World Comput. Sci. Inf. Technol. J. WCSIT*, vol. 2, no. 2, pp. 51–56, 2012.
- [10] N. W. Wardani *et al.*, “PREDIKSI CUSTOMER CHURN DENGAN ALGORITMA DECISION TREE C4 . 5,” vol. 1, no. 1, pp. 16–24, 2018.
- [11] D. Pernikahan, P. Kantor, and K. U. A. Siantar, “2 , 3 , 4,” vol. 1, pp. 64–68, 2016.