



# Klasifikasi Kepuasan Siswa Terhadap Kinerja Guru Smk Satrya Budi 2 Perdagangan Menggunakan Algoritma C4.5

Wili Darma Kusuma<sup>1,\*</sup>, Suhada<sup>2</sup>, Saifullah<sup>1</sup>

<sup>1</sup> STIKOM Tunas Bangsa, Pematangsiantar, Sumatera Utara, Indonesia

<sup>2</sup> AMIK Tunas Bangsa, Pematangsiantar, Sumatera Utara, Indonesia

Email: <sup>1</sup>wilidarmakusuma@gmail.com, <sup>3</sup>saifullah@amiktunasbangsa.ac.id

Email Penulis Korespondensi: wilidarmakusuma@gmail.com

**Abstrak**—Tujuan dari penelitian ini adalah untuk Mengukur kepuasan siswa terhadap kinerja guru Smk Satrya Budi 2 Perdagangan dengan menggunakan teknik data mining klasifikasi Algoritma C4.5. Sumber data diperoleh dari kuisioner yang telah diberikan kepada siswa Smk Satrya Budi 2 Perdagangan dan pemberian kuisioner kepada siswa sebanyak 97 siswa. Atribut yang digunakan sebagai parameter penilaian kepuasan siswa antara lain : Kehadiran Guru (C1), Penguasaan Materi (C2), Penjelasan Materi (C3), Tugas (C4), Sikap Guru (C5), dan Penampilan Guru (C6). Hasil dari pengolahan Algoritma C4.5 dengan menggunakan bantuan software Weka adalah atribut kehadiran guru (C1) dan penjelasan materi (C3) menjadi atribut yang berpengaruh terhadap kepuasan siswa, terhadap kinerja guru. Dan nilai akurasi yang ditunjukkan terhadap kesesuaian Algoritma C4.5 dengan kepuasan siswa terhadap kinerja guru adalah 94%. Manfaat dari penelitian ini adalah untuk meningkatkan kualitas kinerja guru dan meningkatkan kualitas mengajar guru.

**Kata Kunci:** Data Mining; Klasifikasi; Algoritma C4.5; Siswa; Kinerja Guru

**Abstract**—The purpose of this study was to measure student satisfaction with teacher performance Smk Satrya Budi 2 Perdagangan using classification data mining techniques C4.5 Algorithm. Data Sources obtained from questionnaires that have been given to students Smk Satrya Budi 2 Perdagangan and giving questionnaires to students as many as 97 students. attributes used as a parameter for assessing student satisfaction include: Teacher attendance (C1), Material Mastery (C2), Material Explanation (C3), Tasks (C4), Attitude Teacher (C5), and Teacher appearance (C6). The results of processing the C4.5 algorithm using the help of Weka software are the attributes of teacher attendance (C1) and material explanation (C3) which are attributes that affect student satisfaction on teacher performance. And the accuracy value shown against the suitability of the C4.5 algorithm with student satisfaction with teacher performance is 94%. The benefit of this research is to improve the quality of teacher performance and improve the quality of teacher teaching.

**Keywords:** Data Mining; Classification; C4.5 Algorithm; Students; Teacher Performance

## 1. PENDAHULUAN

Guru sebagai tenaga pendidikan dituntut untuk tidak hanya sebagai pengajar tapi juga dituntut untuk menjadi pendidik, pelatih dan pembimbing yang memberikan contoh positif kepada siswa [1]. Tidak semua kinerja guru di dunia pendidikan berkinerja baik, Terkadang ada beberapa sekolah yang memiliki kualitas belajar mengajar yang buruk. Proses belajar mengajar yang buruk akan berdampak buruk juga terhadap siswa, sebaliknya jika belajar mengajarnya baik maka akan berdampak baik terhadap siswa tersebut. Sebagaimana dinyatakan oleh [2], Kemampuan dan perkembangan para siswa/siswi bukan hanya dilihat dari belajar dan fasilitas yang memadai, namun guru juga menjadi faktor yang sangat penting dalam meningkatkan kemampuan siswa/siswi. Guru yang berkompeten dan memiliki kemampuan dibidangnya tentunya akan dapat dengan mudah menyampaikan materi kepada para siswa dan siswi pun dapat dengan mudah menerima materi yang disampaikan oleh para guru. Oleh karena itu untuk mendapatkan guru yang berkualitas tentunya pihak sekolah akan melakukan penilaian bagi setiap guru yang mengajar pada sekolah tersebut dengan mengukur tingkat kepuasan siswa ataupun siswi guna memperbaiki serta meningkatkan kualitas sekolah. Setiap siswa maupun siswi tentunya ingin memperoleh pelayanan pendidikan yang baik. Itulah sebabnya perlu diketahui puas atau tidaknya siswa terhadap kinerja guru yang diberikan oleh pihak sekolah. Dari permasalahan diatas, penulis menggunakan data mining klasifikasi dengan algoritma C4.5 untuk mendapatkan solusi yang lebih efisien pada kepuasan siswa terhadap kinerja guru SMK Satrya Budi 2 Perdagangan. Algoritma C4.5 merupakan kelompok algoritma Decision Tree. Algoritma ini mempunyai input berupa training samples dan samples. Training samples berupa data contoh yang akan digunakan untuk membangun sebuah tree yang telah diuji kebenarannya. Sedangkan samples merupakan field-field data yang nantinya akan digunakan sebagai parameter dalam melakukan klasifikasi data [3]. Algoritma C4.5 dapat digunakan untuk memiliki berbagai macam hal, diantaranya penelitian terkait untuk memprediksi kepuasan mahasiswa terhadap kinerja dosen di kota Batam. Pengujian yang dilakukan dengan *software weka* menghasilkan tingkat akurasi 94.12% [4]. [5] Melakukan prediksi prestasi siswa smp nurul jadid menggunakan algoritma c4.5 dan menghasilkan akurasi sebesar 72.4652% namun ketika attribute yang ku- rang berpengaruh diabaikan tingkat akurasi menurun sebesar 2%. Dengan Menggunakan metode Algoritma C4.5 untuk mengetahui kualitas dari seorang guru dalam mengajar di lingkungan SMK Satrya Budi 2 Perdagangan, penelitian ini diharapkan mampu mengetahui kepuasan siswa terhadap kinerja guru dan mengetahui kinerja guru mana yang perlu ditingkatkan untuk memenuhi kepuasan dari siswa.



## 2. METODOLOGI PENELITIAN

### 2.1 Data Mining

Data mining adalah proses yang menggunakan statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai basis data besar. Dapat disimpulkan data mining merupakan suatu proses otomatis untuk menganalisa data. Data yang akan diproses merupakan data yang sangat besar, sehingga sulit untuk memproses dengan manual [6]. Data mining juga dapat diartikan sebagai proses yang mempekerjakan satu atau lebih teknik pembelajaran computer (*machine learning*) untuk menganalisis dan mengekstraksi pengetahuan (*knowledge*) secara otomatis. *Knowledge Discovery in Databases* (KDD) adalah penerapan metode saintifik pada data mining. Dalam konteks ini data mining merupakan satu langkah dari proses KDD [7].

### 2.2 Algoritma C4.5

Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Pohon keputusan dapat membagi kumpulan data yang besar menjadi himpunan-himpunan *record* yang lebih kecil dengan menerapkan serangkaian aturan keputusan [8]. Secara umum algoritma C4.5 yang digunakan untuk membuat pohon keputusan adalah sebagai berikut :

- a) Pilih atribut sebagai akar (*root node*)

Atribut akar dapat dilakukan dengan menghitung jumlah kasus dan jumlah target atribut. Nilai *entropy* dihitung untuk menghasilkan nilai atribut.

Berikut ini adalah rumus dasar perhitungan *entropy*:

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

Dimana :

S : Himpunan kasus

n : Jumlah partisi S

pi : Proporsi dari Si terhadap S

- b) Setelah menghitung entropy setiap kasus, selanjutnya menghitung nilai *gain* untuk pemisah objek dengan rumus:

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

Dimana :

S : Himpunan kasus

A : Atribut

n : Jumlah partisi atribut A

|Si| : Jumlah kasus pada partisi ke-i

|S| : Jumlah kasus dalam S

Setelah semua atribut dihitung menggunakan persamaan diatas, maka atribut yang memiliki informasi tertinggi dibanding atribut yang lain dijadikan *node* (akar).

- c) Membuat cabang untuk tiap-tiap nilai

Setelah diperoleh atribut *gain* tertinggi maka atribut tersebut digunakan sebagai *node*. *Node* ini memiliki *instance* sehingga *instance* dijadikan sebagai cabang dari *node*.

- d) Membagi kasus dalam cabang

Nilai *instance* diklasifikasikan berdasarkan makna dari nilai *instance* tersebut agar menjadi lebih sederhana. Tetapi jika nilai *instance* tidak dapat disederhanakan lagi, maka perlu melakukan perhitungan lebih lanjut.

- e) Ulangi proses untuk setiap cabang sampai semua kasus pada semua cabang memiliki kelas yang sama.

## 3. HASIL DAN PEMBAHASAN

Dalam penelitian ini data yang digunakan adalah hasil kuesioner yang diberikan kepada Siswa Smk Satyra Budi 2 Perdagangan. Dari hasil kuesioner yang dilakukan penulis mengambil sampel data sebanyak 97 responden. Data hasil dari kuesioner selanjutnya diolah menggunakan *Weka* dengan operator *Performance* yang berfungsi sebagai validasi dan reabilitas data untuk mencari keakuratan data. Data yang akurat dilakukan pengolahan data untuk mencari hasil dari masalah penelitian dengan menggunakan *Weka* dan mengambil keputusan dari hasil klasifikasi dengan algoritma C4.5. Berikut ini sampel data hasil rekapitulasi kuesioner yang diolah di *Microsoft Excel* pada Tabel 1.

**Tabel 1.** Data Hasil Rekapitulasi Kuesioner Penelitian

Responden	Kehadiran	Penguasaan	Penjelasan	Tugas	Sikap	Penampilan	Hasil
Responden 1	TW	TM	SJ	SS	ST	RA	Baik
Responden 2	TW	SM	SJ	SS	ST	RA	Baik
Responden 3	TW	SM	TJ	SS	ST	RA	Baik
Responden 4	TW	TM	TJ	SS	ST	RA	Baik
Responden 5	TW	SM	SJ	SS	ST	RA	Baik
Responden 6	TW	TM	SJ	SS	ST	RA	Baik



Responden	Kehadiran	Penguasaan	Penjelasan	Tugas	Sikap	Penampilan	Hasil
Responden 7	TW	TM	TJ	SS	ST	TR	Baik
Responden 8	TW	TM	SJ	SS	ST	RA	Baik
Responden 9	TW	SM	SJ	SS	ST	RA	Baik
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
Responden 97	TW	SM	SJ	TP	ST	RA	Baik

Perhitungan Algoritma C4.5 untuk memperoleh model aturan pohon keputusan dapat diuraikan sebagai berikut :

- Menghitung jumlah kasus, jumlah kasus untuk keputusan Baik, jumlah kasus untuk keputusan Tidak Baik.
- Menghitung *Entropy* dari semua kasus dan kasus yang dibagi berdasarkan kelas atribut dengan persamaan (1). Selanjutnya dilakukan penghitungan *Gain* untuk masing-masing atribut dengan persamaan (2). Berikut ini adalah perhitungan nilai *entropy* dan *gain*.

**Menghitung entropy total :**

$$Entropy [Total] = \left(-\frac{66}{97} \times \log_2 \left(\frac{66}{97}\right)\right) + \left(-\frac{31}{97} \times \log_2 \left(\frac{31}{97}\right)\right)$$

$$Entropy [Total] = 51,3951943$$

**Menghitung entropy dan gain Kehadiran Guru :**

*Entropy* [Kehadiran Guru-Tepat Waktu]

$$= \left(-\frac{66}{77} \times \log_2 \left(\frac{66}{77}\right)\right) + \left(-\frac{11}{84} \times \log_2 \left(\frac{11}{77}\right)\right)$$

$$= 0,591672779$$

*Entropy* [Kehadiran Guru-Tidak Tepat Waktu]

$$= \left(-\frac{0}{20} \times \log_2 \left(\frac{0}{20}\right)\right) + \left(-\frac{20}{20} \times \log_2 \left(\frac{20}{20}\right)\right)$$

$$= 0$$

*Gain* [Total, Kehadiran Guru]

$$= 51,3951943 - \left(\left(\frac{77}{97} \times 0,591672779\right) - \left(\frac{20}{97} \times 0\right)\right)$$

$$= 50,92551591$$

Begitu seterusnya hingga diperoleh nilai *entropy* dan *gain* untuk seluruh kriteria. Berikut ini hasil perhitungan nilai *entropy* dan *gain* yang diuraikan pada tabel 2

**Tabel 2.** Hasil Perhitungan *Node 1*

Node 1	Jlh Kasus	Baik	Tidak Baik	Entropy	Information Gain
TOTAL	97	66	31	51,3951943	
Kehadiran Guru					50,92551591
Tepat Waktu	77	66	11	0,591672779	
Tidak Tepat Waktu	20	0	20	0	
Penguasaan Materi					50,82566981
Sangat Menguasai	48	47	1	0,14609425	
Tidak Menguasai	49	19	30	0,984313285	
Penjelasan Materi					50,91599794
Sangat Jelas	48	48	0	0	
Tidak Jelas	49	18	31	0,948613198	
Tugas					50,66141316
Sering Sekali	31	35	2	0,303374836	
Tidak Pernah	60	31	29	0,999198354	
Sikap					50,80333517
Sangat Tegas	70	61	9	0,553507265	
Tidak Tegas	27	5	22	0,691289869	
Penampilan Guru					50,56514895
Rapi	78	59	19	0,800959108	
Tidak Rapi	19	7	12	0,949452015	

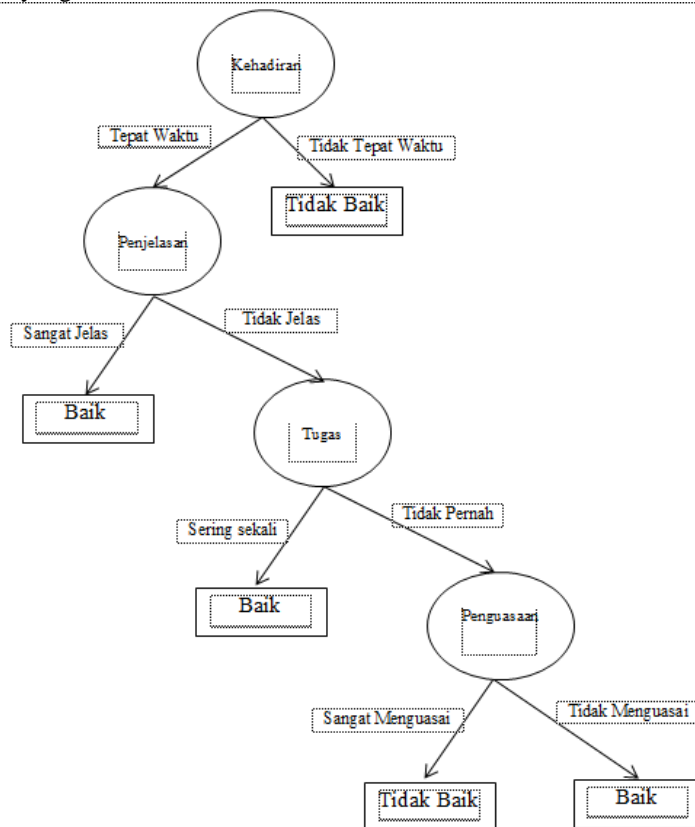
- Dari hasil perhitungan pada Tabel 2. diperoleh nilai atribut tertinggi adalah *Kehadiran* dengan *gain* sebesar 50,92551591. Maka atribut *Kehadiran* dipilih sebagai *node* akar. Nilai kelas atribut Tidak Tepat Waktu mengklasifikasikan kasus menjadi satu keputusan yaitu Tidak Baik. Sedangkan Nilai kelas atribut Tepat Waktu, belum diperoleh hasil antara keputusan Baik dan Tidak Baik, maka perlu dilakukan perhitungan lebih lanjut. Pada

penelitian ini hasil klasifikasi diperoleh hingga *Node* 1.1.1.1. berikut adalah hasil klasifikasi akhir dari pengolahan data menggunakan algoritma C4.5:

**Tabel 3.** Hasil Perhitungan *Node* 1.1.1.1

Node 1.1.1.1	Jlh Kasus	Baik	Tidak Baik	Entropy	Information Gain
TOTAL	97	66	31	51.3951943	
Penguasaan Materi					51,28700529
Sangat Menguasai	3	3	0	0	
Tidak Menguasai	14	3	11	0,749595257	
Sikap					51,23676063
Sangat Tegass	12	5	7	0,979868757	
Tidak Tegass	5	1	4	0,721928095	
Penampilan					51,23632925
Rapi	9	4	5	0,99107606	
Tidak Rapi	8	2	6	0,811278124	

Dari hasil perhitungan pada Tabel 3. diatas, atribut yang menjadi *node* cabang dari Kehadiran Guru = Tepat Waktu dan Penjelasan Materi = Tidak Jelas dan *Tugas* = Tidak Pernah adalah *Penguasaan materi* dengan nilai *gain* tertinggi sebesar 51,28700529. Nilai kelas atribut Sangat Menguasai, telah mengklasifikasikan satu keputusan Baik. Kemudian Kelas atribut Tidak Menguasai adalah Tidak Baik. Maka dari itu perhitungan selesai pada *Node* 1.1.1.1 sebagai *node* cabang terakhir yang terbentuk.



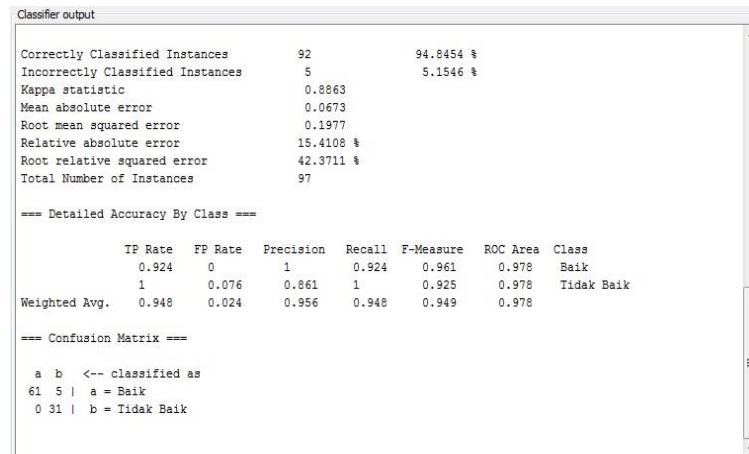
**Gambar 1.** Pohon Keputusan Hasil Perhitungan Algoritma C4.5

Dengan memperhatikan pohon keputusan pada gambar 1 diketahui bahwa semua kasus sudah masuk dalam kelas dengan demikian pohon keputusan pada gambar merupakan pohon keputusan terakhir yang terbentuk. Setelah pohon terbentuk, dihasilkan sejumlah aturan dalam pohon tersebut. Contoh aturan yang dapat terbentuk dari pohon pada gambar 4.5 adalah sebagai berikut:

- 1) Jika *Kehadiran Guru* = Tidak Tepat Waktu, maka hasilnya “Tidak Baik”.
- 2) Jika *Kehadiran Guru* = Tepat Waktu dan *Penjelasan Materi* = Sangat Jelas, maka hasilnya “Baik”.
- 3) Jika *Kehadiran Guru* = Tepat Waktu dan *Penjelasan Materi* = Tidak Jelas dan *Tugas* = Sering Sekali maka hasilnya “Baik”.
- 4) Jika *Kehadiran Guru* = Tepat Waktu dan *Penjelasan Materi* = Tidak Jelas dan *Tugas* = Tidak Pernah dan *Penguasaan Materi* = Sangat menguasai maka hasilnya “Baik”.

- 5) Jika *Kehadiran Guru* = *Tepat Waktu* dan *Penjelasan Materi* = Tidak Jelas dan *Tugas* = Tidak Pernah dan *Penguasaan Materi* = Tidak Menguasai, dan maka hasilnya “Tidak Baik”.
- 6) Jika *Kehadiran Guru* = *Tepat Waktu* dan *Penjelasan Materi* = Tidak Jelas dan *Tugas* = Tidak Pernah dan *Penguasaan Materi* = Sangat Menguasai, maka hasilnya “Baik”.

Pada tahap akhir penerapan Algoritma C4.5 dilakukan penyesuaian hasil perhitungan manual melalui pengujian menggunakan *software Weka* 3.6.9. Berikut ini beberapa tahap proses pengujian menggunakan *software weka*. Hasil penerapan Algoritma C4.5 menggunakan *software RapidMiner* dengan operator *Cross Validation* diperoleh nilai akurasi yaitu sebesar 94,00% dan *classification error* yaitu sebesar 5,00%. Hasil akurasi tersebut diperoleh dengan pengaturan pada operator *Cross Validation* dengan nilai *number of validation* = 10 dan *sampling type* = *shuffled sampling*. Berikut ini adalah hasil akurasi yang diperoleh dapat dilihat pada Gambar 2.



```
Classifier output
Correctly Classified Instances      92          94.8454 %
Incorrectly Classified Instances    5           5.1546 %
Kappa statistic                    0.8863
Mean absolute error                0.0673
Root mean squared error            0.1977
Relative absolute error            15.4108 %
Root relative squared error        42.3711 %
Total Number of Instances         97

=== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
      -----  -
      0.924    0        1          0.924  0.961     0.978    Baik
      1        0.076    0.861     1       0.925    0.978    Tidak Baik
Weighted Avg.   0.948    0.024    0.956    0.948  0.949     0.978

=== Confusion Matrix ===
 a b  <-- classified as
 61 5 | a = Baik
 0 31 | b = Tidak Baik
```

**Gambar 2.** Nilai Akurasi Algoritma C4.5

Dari gambar diatas nilai *Accuracy* sebesar 94,00%, artinya aturan atau rule yang dihasilkan mendekati 100%, dan *classification error* yaitu sebesar 5,00%. Artinya hasil proses yang dilakukan peneliti pada perhitungan Algoritma C4.5 dan *Weka* diperoleh hasil yang sama dan sesuai. Sehingga pengujian dengan *Weka* dapat dikatakan berhasil dan dapat menemukan pohon keputusan pada kasus kepuasan siswa terhadap kinerja guru smk satria budi 2 perdagangan.

## 4. KESIMPULAN

Berdasarkan hasil dan pembahasan sebelumnya dapat disimpulkan bahwa data mining dengan algoritma C4.5 dapat diterapkan untuk mengklasifikasi tingkat kepuasan siswa terhadap kinerja guru smk satria budi 2 perdagangan. Hasil yang diperoleh dari pengujian metode algoritma C4.5 kedalam *Weka* memiliki nilai akurasi yang sama. Hasil Pengujian Algoritma C4.5 dengan menggunakan *Software Weka*, Dalam menentukan kepuasan siswa terhadap kinerja guru smk satria budi 2 perdagangan telah didapatkan pohon keputusan dan rule yang menunjukkan bahwa atribut *Kehadiran Guru* dan *penjelasan materi* sebagai atribut yang paling berpengaruh terhadap kinerja guru. Algoritma C4.5 sangat cocok diterapkan dalam mengkasifikasi tingkat kepuasan kepuasan siswa terhadap kinerja guru smk satria budi 2 perdagangan yang mana diperoleh tingkat akurasi yaitu 94,00%.

## REFERENCES

- [1] M. Kristiawan And N. Rahmat, “Peningkatan Profesionalisme Guru Melalui Inovasi Pembelajaran,” *J. Iqra’ Kaji. Ilmu Pendidik.*, Vol. 3, No. 2, Pp. 373–390, 2018, Doi: <https://doi.org/10.25217/Ji.V3i2.348>.
- [2] B. G. Gerardo, S. Saifullah, And E. Irawan, “Teknik Data Mining Dalam Penilaian Pengajaran Guru Berdasarkan Indeks Kepuasan Siswa,” *Komik (Konferensi Nas. Teknol. Inf. Dan Komputer)*, Vol. 3, No. 1, Pp. 508–514, 2019, Doi: 10.30865/Komik.V3i1.1634.
- [3] C. U. Memprediksi And P. Diabetes, “Analisa Metode Classification-Decission Tree Dan Algoritma,” Vol. 2, No. 1, Pp. 10–17, 2019.
- [4] A. D. P. Yulia, “Computer Based Information System Journal Data Mining Menggunakan Algoritma C4.5 Untuk Memprediksi Kepuasan Mahasiswa Terhadap Terhadap Kinerja Dosen Di Kota Batam,” *Comput. Based Informaon Syst. J.*, Vol. 07, No. 02, Pp. 56–66, 2019.
- [5] K. Malik And M. Faid, “Prediksi Prestasi Siswa Smp Nurul Jadid Menggunakan Algoritma C4.5,” *Nusant. J. Comput. Its Appl.*, Vol. 2, No. 4, Pp. 2–5, 2016.
- [6] M. G. Sadewo, A. P. Windarto, And D. Hartama, “Penerapan Datamining Pada Populasi Daging Ayam Ras Pedaging Di Indonesia Berdasarkan Provinsi Menggunakan K-Means Clustering,” *Infotekjar (Jurnal Nas. Inform. Dan Teknol. Jaringan)*, Vol. 2, No. 1, Pp. 60–67, 2017.
- [7] F. Riandari And A. Simangunsong, “Penerapan Algoritma C4.5 Untuk Mengukur Tingkat Kepuasan Mahasiswa,” *Terakreditasi Dikti*, Vol. 3, No. 2, Pp. 1–7, 2019.



- [8] E. Elisa, “Analisa Dan Penerapan Algoritma C4.5 Dalam Data Mining Untuk Mengidentifikasi Faktor-Faktor Penyebab Kecelakaan Kerja Kontruksi Pt.Arupadhatu Adisesanti,” *J. Online Inform.*, Vol. 2, No. 1, P. 36, 2017, Doi: 10.15575/Join.V2i1.71.
- [9] D. N. Batubara And A. P. Windarto, “Analisa Klasifikasi Data Mining Pada Tingkat Kepuasan Pengunjung Taman Hewan Pematang Siantar Dengan Algoritma,” *Komik (Konferensi Nas. Teknol. Inf. Dan Komputer)*, Vol. 3, No. 1, Pp. 588–592, 2019, Doi: 10.30865/Komik.V3i1.1664.
- [10] A. Shiddiq, R. K. Niswatin, And I. N. Farida, “Ahmad Shiddiq Analisa Kepuasan Konsumen Menggunakan Klasifikasi Decision Tree Di Restoran Dapur Solo (Cabang Kediri),” *Gener. J.*, Vol. 2, No. 1, P. 9, 2018, Doi: 10.29407/Gj.V2i1.12051.