

# IMPLEMENTASI DATA MINING DALAM PREDIKSI KEPUASAN BELAJAR SAAT PANDEMIC COVID MENGGUNAKAN ALGORITMA C 4.5

Asnija Elisabeth Siahaan<sup>1</sup>, Rahmat Fauzi<sup>2</sup>

<sup>1</sup>Mahasiswa Program Studi Teknik Informatika, Universitas Putera Batam

<sup>2</sup>Dosen Program Studi Teknik Informatika, Universitas Putera Batam

email: [pb180210082@upbatam.ac.id](mailto:pb180210082@upbatam.ac.id)

## ABSTRACT

*In 2020, the Covid-virus pandemic has spread throughout the world, including in Indonesia, of course affecting all aspects, especially the lives of Indonesian people, including education during the online learning process and the transition from face-to-face processes, of course, various obstacles arise including the lack of students understanding the teaching and learning process, and teachers who are unable to convey material optimally, that's why this online learning system requires the implementation of Data Mining to maximize learning and help related parties. The method used in this research is analysis and data mining. Analysis is an attempt to observe something in detail. things or things by outlining its constituent components or constituents for further study. Data mining, also known as knowledge, is one of the fields whose rapid growth is due to the high demand for added value from large-scale database batteries. From the research that has been done, it indicates that the decision tree can be a reference for schools in order to increase student and female learning satisfaction in the learning process, and based on the results of system testing it can be concluded that this application was successful in increasing student satisfaction in online learning during the covid pandemic. Data mining is used by the C4.5 algorithm to interest students in online learning during the corona virus.*

**Keywords:** Covid, Data Mining, Algoritma C4.5, Analysis, Study

## PENDAHULUAN

Di tahun 2020 adanya pandemi virus covid yang menyebar ke seluruh dunia termasuk di indonesia, tentunya mempengaruhi seluruh aspek terutama kehidupan masyarakat Indonesia termasuk pendidikan, karena penyebarannya seperti virus flu yang membuat kita mudah terjangkit virus covid tersebut, dan karna keadaan tersebut, tentunya memaksa kita untuk menjaga jarak yang aman dan menghindari keramaian, selama proses

pembelajaran daring dan peralihan dari proses tatap muka tentunya muncul berbagai kendala termasuk kurangnya murid dalam memahami proses belajar mengajar, dan para guru yang kurang bisa menyampaikan materi dengan maksimal karena itulah sistem pembelajaran daring ini diperlukan implementasi *Data Mining* untuk memaksimalkan pembelajaran dan membantuk pihak-pihak yang terkait terutama pihak sekolah sehingga orang tua atau wali murid dan juga siswa dapat

belajar dengan maksimal dan mencapai kepuasan saat pembelajaran daring sedang berlangsung.

memprediksi keputusan dengan menerapkan serangkaian aturan keputusan (Febriani & Sulistiani, 2021).

## KAJIAN TEORI

### 2.1 Knowledge Discovery in Database

*Data mining*, sering juga disebut 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 dipakai untuk memperbaiki pengambilan keputusan dimasa depan (Panggabean et al., 2020).

### 2.2 Data Mining

*Data Mining* adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. *Data Mining* dapat diterapkan pada berbagai bidang yang mempunyai sejumlah data, tetapi karena wilayah penelitian dengan sejarah yang belum lama, dan belum melewati masa 'remaja', maka *data mining* masih diperdebatkan posisi bidang pengetahuan yang memilikinya. Maka Daryl Pregibon menyatakan bahwa " *Data Mining* adalah campuran dari statistik, kecerdasan buatan, dan riset basis data " yang masih berkembang *Data Mining* mempunyai empat akar bidang ilmu yaitu Statistik, Kecerdasan Buatan ( Artificial Intelligence ), Pengenalan Pola, dan Sistem Basis Data (Panggabean et al., 2020).

### 2.3 Algoritma C4.5

Dalam *data mining* terdapat beberapa metode, salah satunya adalah *algoritma C4.5*. *Algoritma C4.5* merupakan algoritma yang digunakan untuk

### 2.4 Rapidminer

*RapidMiner* adalah sebuah aplikasi atau perangkat lunak yang berperan sebagai sarana pembelajaran ilmu data mining. Platform ini dikembangkan oleh perusahaan yang didedikasikan untuk semua tahapan yang melibatkan data dalam jumlah besar dalam masalah bisnis, penelitian, pendidikan, pelatihan, dan pembelajaran. *RapidMiner* memiliki sekitar 100 solusi pembelajaran untuk pengelompokan, klasifikasi, dan analisis regresi, *RapidMiner* juga mendukung sekitar 22 format file, seperti .xls, .csv, dan lainnya.

*RapidMiner* membawa kecerdasan buatan ke bisnis melalui platform ilmu data yang terbuka dan dapat diskalakan. *RapidMiner* dibangun untuk tim analitik, mengintegrasikan seluruh siklus ilmu data, mulai dari persiapan data hingga pembelajaran mesin hingga penerapan model prediktif. Lebih dari 625.000 analis menggunakan produk *RapidMiner* untuk meningkatkan pendapatan, mengurangi biaya, dan menghindari risiko (Prasetyo et al., 2021)

## METODE PENELITIAN

### 3.1 Tahapan Penelitian

Tahapan penelitian akan dijelaskan secara singkat dalam pengertian ini antara lain adalah tahapan proses penelitian lokasi atau variabel yang harus diteliti pada tahapan ini.



Gambar 3.1 Data Mining

Metode yang digunakan dalam penelitian ini adalah Analisa dan *data mining*, Analisa atau *analysis* adalah suatu usaha untuk mengamati secara detail sesuatu hal atau benda dengan cara menguraikan komponen-komponen pembentuknya atau penyusunya untuk dikaji lebih lanjut.

Penambangan data atau *data mining*, juga dikenal sebagai pengetahuan, adalah salah satu bidang tersebut pertumbuhan yang cepat karena permintaan yang tinggi untuk nilai tambah dari baterai database berskala besar terakumulasi seiring perkembangan teknologi informasi dengan sangat cepat penambangan data juga dikenal sebagai penemuan pengetahuan (Fauziah et al., 2020).

### 3.2 Operasional Variabel

Dalam melakukan analisis data diperlukan variable yang harus diolah dalam proses pembelajaran khususnya variable yang menentukan kepuasan belajar siswa, menggunakan beberapa aspek tentang algoritma dan mengukur kepuasan siswa-siswi dalam 5 aspek yaitu *tangible*, *reability*, *empathy*, *assurance* dan *responsiveness*.

### 3.3 Lokasi Penelitian

Penelitian ini berlokasi di SMAN5

Batam beralamat di JL. KAVLING LAMA, Sagulung Kota, Kec. Sagulung, Kota Batam, Kepulauan Riau, dengan kode pos 29439.



Gambar 3.2 Lokasi SMAN 5 Batam

## HASIL DAN PEMBAHASAN

### 4.1 Hasil

Dalam bab ini akan diberikan hasil dari aplikasi yang telah dirancang dan telah diuji dalam penelitian ini.

### 4.2 Pembahasan

Untuk mengetahui apakah aplikasi ini telah berjalan dengan baik maka diperlukan sebuah pengujian dan akan ditunjukkan beberapa hasil pengujian aplikasi ini.

#### 4.2.1 Pengujian Siswa

Pengujian dilakukan pada siswa dan siswi di 1 Kelas IPA dan 1 Kelas IPS di SMAN 5 Batam, Penelitian ini mengukur tingkat seberapa puas siswa dan siswi terhadap 5 aspek penting yaitu *tangible*, *reability*, *empathy*, *assurance*, dan *responsiveness*.

##### a. aspek *tangible*

Kualitas pembelajaran yang diberikan dalam bentuk daring

### b. aspek reability

Kemampuan guru dalam memberikan bahan ajar yang berkualitas kepada dan siwi

### c. aspek empathy

Kesediaan guru dalam memberikan solusi dalam permasalahan siwa dalam pembelajaran.

### d. aspek assurance

Kecepatan guru dalam menjawab dan merespon pertanyaan atau masalah yang sedang dihadapi siswa dalam pembelajaran.

### e. aspek responsiveness

Bagaimana kualitas interaksi antara guru dan juga siswa dalam menggunakan sistem Daring?

Setelah mengetahui aspek apa saja yang akan di uji maka akan diberikan data siapa saja yang akan menjadi penguji aplikasi yang terkait, Tabel dibawah ini

**Tabel 4.1** Data Responden

Siswa-siswi	11 IPS	12 IPA
Perempuan	22	21
Laki-laki	13	14

Tabel dibawah ini adalah hasil yang diberikan setelah melakukan pengujian kepada para responden dan

hasil dari pengujian tersebut adalah seperti berikut :

**Tabel 4.2** Hasil Rekapitulasi Kuisisioner

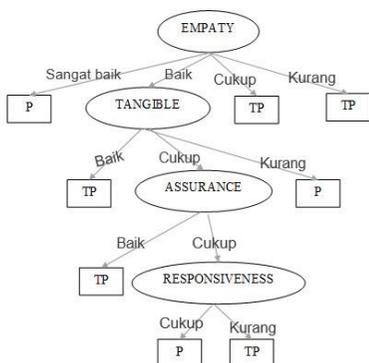
No	TANGIBL E	REABILIT Y	ASSURAN CE	EMPATHY	RESPONSIVENE SS	HASI L
1	Kurang	Baik	Cukup	Baik	Kurang	TP
2	Baik	Cukup	Kurang	Kurang	Kurang	TP
3	Kurang	Cukup	Baik	Cukup	Kurang	TP
4	Cukup	Cukup	Kurang	Cukup	Kurang	TP

5	Cukup	Baik	Baik	Cukup	Kurang	TP
6	Baik	Baik	Baik	Baik	Baik	P
7	Kurang	Cukup	Cukup	Cukup	Kurang	TP
8	Kurang	Kurang	Kurang	Kurang	Kurang	P
9	Baik	Cukup	Baik	Cukup	Kurang	TP
10	Kurang	Kurang	Kurang	Kurang	Kurang	TP
11	Cukup	Cukup	Cukup	Cukup	Cukup	TP
12	Kurang	Baik	Kurang	Cukup	Kurang	TP
13	Kurang	Cukup	Kurang	Kurang	Kurang	TP
14	Kurang	Kurang	Kurang	Kurang	Kurang	TP
15	Kurang	Kurang	Kurang	Kurang	Kurang	TP
16	Kurang	Kurang	Cukup	Kurang	Cukup	TP
17	Kurang	Kurang	Cukup	Kurang	Cukup	TP
18	Kurang	Cukup	Cukup	Kurang	Kurang	TP
19	Baik	Baik	Baik	Baik	Cukup	TP
20	Cukup	Baik	Cukup	Baik	Cukup	TP
21	Kurang	Kurang	Kurang	Kurang	Kurang	TP
22	Cukup	Baik	Cukup	Baik	Kurang	TP
23	Cukup	Cukup	Cukup	Cukup	Kurang	TP
24	Baik	Baik	Baik	Baik	Kurang	TP
25	Cukup	Baik	Baik	Baik	Kurang	TP
26	Kurang	Cukup	Kurang	Kurang	Kurang	TP
27	Cukup	Kurang	Kurang	Kurang	Cukup	TP
28	Cukup	Cukup	Cukup	Cukup	Cukup	TP
29	Cukup	Baik	Baik	Baik	Kurang	TP
30	Cukup	Baik	Baik	Baik	Kurang	TP
31	Cukup	Cukup	Cukup	Cukup	Kurang	TP
32	Kurang	Cukup	Cukup	Kurang	Kurang	TP
33	Kurang	Baik	Cukup	Baik	Kurang	TP
34	Kurang	Baik	Cukup	Baik	Kurang	P
35	Cukup	Baik	Baik	Sangat baik	Kurang	P

36	Cukup	Baik	Cukup	Kurang	Cukup	TP
37	Cukup	Baik	Cukup	Kurang	Cukup	TP
38	Cukup	Baik	Cukup	Kurang	Cukup	TP
39	Cukup	Baik	Cukup	Kurang	Cukup	TP
40	Cukup	Cukup	Cukup	Baik	Kurang	TP
41	Cukup	Cukup	Cukup	Cukup	Cukup	P
42	Kurang	Cukup	Kurang	Kurang	Kurang	TP
43	Cukup	Cukup	Kurang	Kurang	Cukup	P
44	Cukup	Cukup	Cukup	Cukup	Cukup	TP
45	Kurang	Baik	Cukup	Cukup	Baik	TP
46	Kurang	Cukup	Cukup	Cukup	Kurang	TP
47	Cukup	Kurang	Cukup	Cukup	Kurang	P
48	Cukup	Baik	Kurang	Cukup	Cukup	TP
49	Kurang	Baik	Baik	Baik	Kurang	P
50	Kurang	Baik	Baik	Baik	Kurang	P
51	Cukup	Cukup	Kurang	Cukup	Baik	TP
52	Cukup	Baik	Cukup	Kurang	Cukup	TP
53	Cukup	Cukup	Cukup	Cukup	Cukup	P
54	Cukup	Kurang	Cukup	Cukup	Kurang	TP
55	Cukup	Cukup	Cukup	Baik	Kurang	TP
56	Cukup	Baik	Cukup	Baik	Kurang	TP
57	Kurang	Kurang	Kurang	Kurang	Kurang	P
58	Kurang	Cukup	Kurang	Baik	Kurang	P
59	Cukup	Baik	Cukup	Baik	Cukup	P
60	Kurang	Cukup	Cukup	Kurang	Kurang	TP
61	Cukup	Cukup	Kurang	Cukup	Cukup	TP
62	Cukup	Baik	Baik	Baik	Baik	TP
63	Cukup	Baik	Cukup	Baik	Cukup	P
64	Cukup	Cukup	Cukup	Cukup	Cukup	TP
65	Cukup	Baik	Cukup	Kurang	Cukup	TP
66	Cukup	Baik	Cukup	Baik	Cukup	P

67	Sangat baik	P				
68	Cukup	Baik	Baik	Cukup	Baik	TP
69	Cukup	Cukup	Cukup	Cukup	Baik	TP
70	Cukup	Baik	Baik	Cukup	Baik	TP

### 4.2.2 Perhitungan Hasil Algoritma C4.5



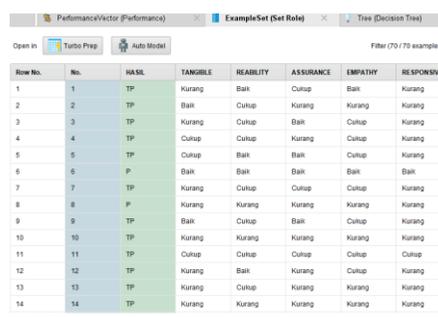
**Gambar 4.1** Pohon Keputusan Algoritma C4.5

Terdapat 8 aturan untuk menentukan tingkat belajar dan kepuasan siswa dan juga siswi dalam pembelajaran selama pandemi covid.

### 4.2.3 Implementasi Data Mining menggunakan RapidMiner

Pengujian dilakukan pada 70 sample data dan menggunakan aplikasi yang disebut *rapidminer* adalah sebagai berikut:

#### a. Hasil Penginputan

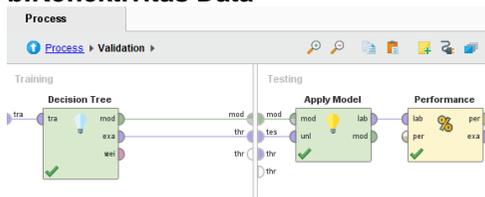


Row No.	No.	KASL	TANGIBLE	REABILITY	ASSURANCE	EMPATHY	RESPONSIV...
1	1	TP	Kurang	Baik	Cukup	Baik	Kurang
2	2	TP	Baik	Cukup	Kurang	Kurang	Kurang
3	3	TP	Kurang	Cukup	Baik	Cukup	Kurang
4	4	TP	Cukup	Cukup	Kurang	Cukup	Kurang
5	5	TP	Cukup	Baik	Baik	Cukup	Kurang
6	6	P	Baik	Baik	Baik	Baik	Baik
7	7	TP	Kurang	Cukup	Cukup	Cukup	Kurang
8	8	P	Kurang	Kurang	Kurang	Kurang	Kurang
9	9	TP	Baik	Cukup	Baik	Cukup	Kurang
10	10	TP	Kurang	Kurang	Kurang	Kurang	Kurang
11	11	TP	Cukup	Cukup	Cukup	Cukup	Cukup
12	12	TP	Kurang	Baik	Kurang	Cukup	Kurang
13	13	TP	Kurang	Cukup	Cukup	Kurang	Kurang
14	14	TP	Kurang	Kurang	Kurang	Kurang	Kurang

**Gambar 4.2** Penginputan RapidMiner

Gambar diatas merupakan proses input dan permodelan pohon keputusan

#### b. Konektivitas Data



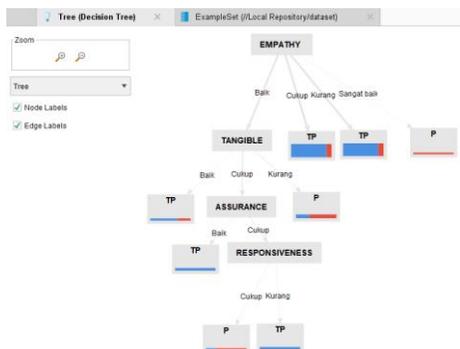
**Gambar 4.3** Konektivitas data dan pohon keputusan

Menampilkan keputusan antara data dan pohon keputusan maupun sebaliknya.

### 4.2.4 Decision Tree

Pohon keputusan merupakan salah satu metode klasifikasi yang digunakan representasi dari struktur pohon di mana setiap node mewakili atribut, cabang mewakili nilai atribut dan daun mewakili kelas ” (Fauziah et al., 2020), Setelah data

diatas diproses maka akan muncul tampilan seperti berikut.



**Gambar 4.4** Decision Tree pada RapidMiner

Setelah diproses hasil dari pengolahan data dengan model dari pohon keputusan adalah seperti gambar diatas.

```

Tree
EMPATHY = Baik
| TANGIBLE = Baik: TP (TP=2, P=1)
| TANGIBLE = Cukup
| | ASSURANCE = Baik: TP (TP=4, P=0)
| | ASSURANCE = Cukup
| | | RESPONSIVENESS = Cukup: P (TP=1, P=3)
| | | RESPONSIVENESS = Kurang: TP (TP=4, P=0)
| TANGIBLE = Kurang: P (TP=2, P=4)
EMPATHY = Cukup: TP (TP=21, P=3)
EMPATHY = Kurang: TP (TP=20, P=3)
EMPATHY = Sangat baik: P (TP=0, P=2)
    
```

**Gambar 4.5** Rule Decision Rapidminer

Hasil implementasi algoritma C4.5 menggunakan aplikasi RapidMiner menggunakan operator Split Validation diperoleh nilai eksak sebesar 65,71%. Hasil nilai presisi diperoleh dengan menetapkan operator validasi split dengan nilai rasio split = 0,5 dan jenis sampling = sampling linier.

ceVector (Performance)			
ExampleSet (Set Role)			
Tree (Decision Tree)			
accuracy: 65.71%			
	true TP	true P	class precision
pred TP	23	12	65.71%
pred P	0	0	0.00%
class recall	100.00%	0.00%	

**Gambar 4.6** Nilai akurasi Algoritma C4.5

Gambar di atas merupakan hasil akurasi yang telah didapatkan dari hasil perhitungan.

## PerformanceVector

```

PerformanceVector:
accuracy: 65.71%
ConfusionMatrix:
True:   TP   P
TP:     23  12
P:       0   0
    
```

**Gambar 4.7** Performa Algoritma C4.5

Gambar di atas menunjukkan nilai eksak sebesar 65,71% yang berarti aturan yang dihasilkan mendekati 100%.

## SIMPULAN

Dari penelitian yang telah dilakukan menandakan bahwa pohon keputusan dapat menjadi acuan untuk sekolah dalam rangka meningkatkan kepuasan belajar siswa dan siswi pada proses pembelajaran, dan berdasarkan hasil pengujian sistem dapat disimpulkan bahwa aplikasi ini berhasil untuk meningkatkan kepuasan siswa dalam pembelajaran daring selama pandemi covid berlangsung.

Hasil penelitian dengan menggunakan algoritma C4.5 dijadikan sebagai prediktor tingkat kepuasan pendidikan pelajar atau mahasiswa di masa pandemi COVID-19. Dari hasil prediksi diharapkan lembaga pendidikan atau instansi terkait dapat lebih memahami faktor-faktor yang mempengaruhi tingkat kepuasan pendidikan. Ini dapat membantu mereka mengambil tindakan yang lebih efektif untuk meningkatkan pengalaman belajar siswa selama keadaan darurat seperti pandemic, nilai dari penelitian ini adalah untuk memberikan panduan kepada pembuat kebijakan untuk merancang kebijakan atau program yang lebih baik dan lebih fleksibel untuk mengatasi tantangan pendidikan selama pandemi. Dengan menerapkan hasil prediksi, institusi pendidikan dapat membuat rencana pembelajaran jarak jauh yang lebih efektif, memberikan siswa dukungan psikologis yang lebih baik, dan mengatur program untuk merespons situasi darurat secara efektif, namun penelitian ini juga memiliki keterbatasan seperti kualitas data yang digunakan, sampel yang dipilih atau parameter yang digunakan pada algoritma C4.5.

Validasi dan perbandingan tambahan dengan metode penambangan data lainnya mungkin diperlukan untuk mengukur keakuratan hasil prediksi. Oleh karena itu, perlu dilakukan penelitian lebih lanjut untuk memperkuat dan melengkapi temuan penelitian ini, oleh karena itu, penerapan Data Mining dan memprediksi kepuasan pendidikan selama pandemi COVID-19 menggunakan algoritma C4.5 berpotensi memberikan manfaat yang besar bagi dunia akademik. Dengan memiliki pemahaman yang lebih baik tentang tingkat kepuasan pendidikan, lembaga pendidikan dapat mengambil langkah-langkah yang tepat dan cepat

untuk meningkatkan kualitas pendidikan di tengah keadaan darurat. Rekomendasi lain termasuk penelitian lebih lanjut di bidang prediksi dan evaluasi kepuasan belajar dengan mempertimbangkan metode pengumpulan data lain atau metode gabungan, dan memasukkan lebih banyak variabel yang dapat mempengaruhi kepuasan belajar siswa.

### DAFTAR PUSTAKA

- Fauziah, F., Hartama, D., & Damanik, I. S. (2020). Analisa Kepuasan Pelanggan Menggunakan Klasifikasi Data Mining. *Brahmana : Jurnal Penerapan Kecerdasan Buatan*, 2(1), 41–48. <https://doi.org/10.30645/brahmana.v2i1.47>
- Prasetyo, V. R., Lazuardi, H., Mulyono, A. A., & Lauw, C. (2021). Penerapan Aplikasi RapidMiner Untuk Prediksi Nilai Tukar Rupiah Terhadap US Dollar Dengan Metode Linear Regression. *Jurnal Nasional Teknologi Dan Sistem Informasi*, 7(1), 8–17. <https://doi.org/10.25077/teknosi.v7i1.2021.8-17>
- Andayani, S. (2007). Formation of clusters in Knowledge Discovery in Databases by Algorithm K-Means. *SEMNAS Matematika Dan Pendidikan Matematika 2007*.
- Febriani, S., & Sulistiani, H. (2021). Analisis Data Hasil Diagnosa Untuk Klasifikasi Gangguan Kepribadian Menggunakan Algoritma C4.5. *89Jurnal Teknologi Dan Sistem Informasi (JTSI)*, 2(4), 89–95.
- Panggabean, D. S. O., Buulolo, E., & Silalahi, N. (2020). Penerapan Data Mining Untuk Memprediksi Pemesanan Bibit Pohon Dengan Regresi Linear Berganda.

	<p><b>Biodata</b></p> <p>Penulis pertama, Asnija Elisabeth merupakan mahasiswa Prodi Teknik Informatika Universitas Putera Batam.</p>
	<p><b>Biodata</b></p> <p>Penulis kedua, Rahmat Fauzi, merupakan Dosen Prodi Teknik Informatika Universitas Putera Batam. Penulis banyak berkecimpung di bidang Teknik dan Komputer</p>

*JURIKOM (Jurnal Riset Komputer)*,  
7(1), 56.  
<https://doi.org/10.30865/jurikom.v7i1.1947>