

**KLASIFIKASI TINGKAT PEMAHAMAN MAHASISWA TERHADAP PEMBELAJARAN
DARING SELAMA COVID-19 MENGGUNAKAN ALGORITMA C 4.5**

**CLASSIFICATION OF STUDENT' S UNDERSTANDING OF E-LEARNING DURING
COVID-19 USING ALGORITHM C 4.5**

**Zelvi Gustiana^{1*}, Arif Rahman², Ananda Hadi
Elyas³ Universitas Dharmawangsa^{1,2,3}**

zelvi@dharmawangsa.ac.id^{1}, arif@dharmawangsa.ac.id², nanda@dharmawangsa.ac.id³*

ABSTRACT

Data mining is data processing that is carried out to obtain unknown information by extracting and managing data patterns so that new valuable data can be obtained from data sets in the database. This article raises a topic related to data mining, namely the Algorithm C 4.5 which is part of the classification group. The final result of the Algorithm C 4.5 is a decision tree and rules that become a new knowledge that will be useful in the future.

Keywords: Algorithm C 4.5, Data Mining, Decision Tree, Classification

ABSTRAK

Data mining adalah pengolahan data yang dilakukan untuk menggali informasi yang secara ini tidak diketahui dengan cara melakukan penggalian data-data dengan cara mengekstraksi dan mengenali pola-pola dari data sehingga diperoleh data baru yang berharga dari kumpulan data yang berada dalam basis data. Artikel ini mengangkat topik yang berhubungan dengan Data mining yaitu Algoritma C 4.5 yang merupakan bagian dari kelompok klasifikasi. Hasil akhir dari Algoritma C 4.5 merupakan sebuah pohon keputusan dan aturan-aturan yang menjadi sebuah pengetahuan baru yang akan berguna dimasa mendatang.

Kata kunci: Algoritma C 4.5, Data Mining, Decision Tree, Klasifikasi

1. Pendahuluan

Pandemi Covid-19 yang terjadi merupakan krisis kesehatan yang mengancam masyarakat di dunia saat ini. Virus ini sangat cepat menyebar dan meluas yang mengakibatkan kerugian besar di berbagai sektor, salah satunya sektor Pendidikan. Banyak negara yang memutuskan untuk melakukan menutup sekolah, perguruan tinggi sebagai salah satu upaya untuk meminimalisir dan memutuskan penyebaran virus ini. Proses belajar mengajar tetap dilaksanakan tapi secara daring sampai keadaan pulih. Pembelajaran daring sangat bermanfaat pada masa pandemi saat ini, ini bertujuan agar siswa ataupun mahasiswa tetap bisa belajar dirumah atau dimanapun mereka berada selama mereka terhubung dengan jaringan internet. Salah satu perguruan tinggi yang melakukan pembelajaran daring ini adalah Universitas Dharmawangsa. Universitas Dharmawangsa menerapkan sistem pembelajaran daring untuk kegiatan proses belajar mengajar. Banyak penelitian yang dilakukan untuk mengukur kephahaman mahasiswa yang melakukan pembelajaran daring, salah satunya adalah Algoritma C 4.5

Algoritma C 4.5 merupakan salah satu metode data mining yang berada pada kelompok

Klasifikasi yang bertujuan untuk mencari pola yang menghasilkan aturan-aturan baru untuk mencapai sebuah tujuan. Algoritma C 4.5 merupakan algoritma yang banyak digunakan untuk melakukan pengolahan data secara besar dalam melakukan pengelompokan, klasifikasi atau segmentasi yang memiliki sifat prediktif. Sementara klasifikasi merupakan sebuah proses yang ada dalam data mining yang memiliki tujuan untuk menemukan sebuah pola yang dihasilkan dari sebuah kumpulan data dan memiliki ukuran data yang relatif besar. Algoritma C 4.5 merupakan pengembangan dari Algoritma ID3, yang memiliki beberapa kelebihan yaitu :

- a. Mampu menangani atribut yang bertipe kontinu atau diskrit
- b. Mampu menangani atribut yang datanya hilang (missing value)
- c. Mampu melakukan pemangkasan terhadap pohon keputusan
- d. Pemilihan atribut dilakukan dengan melakukan pencarian Gain ratio

Terdapat dua buah metode yang dilakukan dalam pemangkasan pohon keputusan, yakni:

- a. Prepruning, yaitu dengan melakukan penghentian pembangunan terhadap subtree lebih awal, yakni dengan melakukan pemutusan agar partisi data training tidak terjadi lebih jauh. Pada pendekatan ini, sebuah pohon dipangkas dengan menggunakan cara menghentikan pembangunan tidak signifikan pada partisi yang akan dibuat.
- b. Postpruning, yaitu dengan cara menyederhanakan pohon keputusan dengan cara membuang beberapa cabang pada subtree setelah pembangunan selesai. Metode ini merupakan metode standar yang digunakan pada Algoritma C 4.5

Reduced Error Pruning merupakan salah satu bagian dari algoritma ini. Algoritma ini membagi data menjadi dua bagian, yakni training data dan test data. Training data merupakan data yang akan digunakan data untuk pembentukan pohon keputusan sedangkan test data merupakan digunakan untuk penghitungan nilai error rate pada pemangkasan pohon keputusan

Pohon keputusan sangat berguna dalam menciptakan pengetahuan baru untuk masa depan. Algoritma C 4.5 saat ini berperan penting dalam menciptakan aturan-aturan pada pohon keputusan. Jadi peneliti memutuskan untuk menggunakan Algoritma dalam penyelesaian masalah terhadap klasifikasi pemahaman mahasiswa terhadap pembelajaran daring selama Covid-19. Sehingga tercipta aturan-aturan baru yang berguna untuk Universitas sehingga dapat berguna untuk meningkatkan atribut-atribut yang menjadi faktor utama dalam pemahaman mahasiswa selama daring.

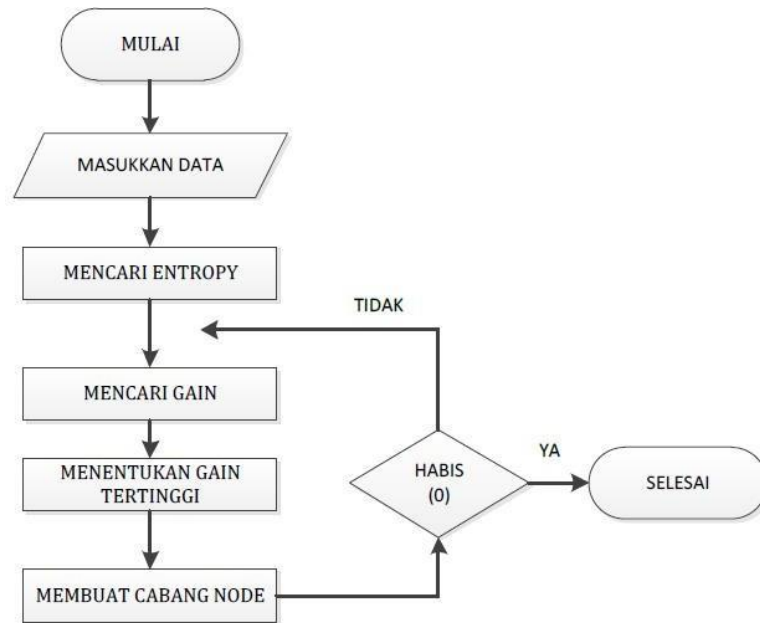
2. Tinjauan pustaka dan pengembangan hipotesis

Tujuan melakukan pendekatan algoritma ini adalah untuk memudahkan dalam melihat dan menentukan sebuah pola yang dibutuhkan. Data mining decision tree bertujuan untuk melihat pola aturan-aturan yang mempengaruhi sebuah tujuan yang berasal dari atribut-atribut yang ada pada data. Decision tree merupakan salah satu Data Mining yang banyak digunakan dalam pengolahan data yang berbentuk pohon keputusan dan hasil akhir berupa aturan-aturan baru. Dalam studi pustaka, dilakukan melalui membedah dan mereview beberapa jurnal yang berkaitan dengan Algoritma C 4.5 dan pemahaman mahasiswa dalam pembelajaran. Hasil akhir dari decision tree Algoritma C 4.5 ini merupakan aturan-aturan yang bisa digunakan atau dimanfaatkan untuk penggunaannya dalam meningkatkan efektifitas yang berkaitan dengan hal yang diteliti dan menjadi sebuah pengetahuan baru pada masa mendatang.

3. Metode penelitian

Penelitian ini dilakukan di Universitas Dharmawangsa dengan menggunakan metode Algoritma C 4.5 yang diproses dengan menggunakan software Rapidminer untuk melakukan mengelompokan atau klasifikasi pada penentuan tingkat pemahaman mahasiswa terhadap pembelajaran daring selama Covid-19. Pengumpulan data dalam penelitian ini menggunakan survey kuisioner. Data yang terkumpul sebanyak 348 data mahasiswa. Data yang digunakan sebagai data training sebanyak 35 data

Untuk memudahkan dalam pembuatan pohon keputusan, digunakan tahapan dasar dari Algoritma C 4.5. Bentuk diagram alir (flowchar) mengenai proses dari Algoritma C 4.5 dapat dilihat pada Gambar 1.



Gambar 1. Flowchart Algoritma C 4.5

Langkah-langkah dalam pengelompokan Algoritma Decision Tree C 4.5 adalah sebagai berikut :

1. Persiapan data yang akan diolah
2. Pemilihan terhadap atribut yang akan menjadi akar
3. Perhitungan untuk mendapatkan Nilai Gain

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

4. Perhitungan untuk mendapatkan Nilai Entropy

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

5. Pembagian kasus terhadap cabang

1. Pengilangan proses untuk masing-masing cabang sampai semua kasus pada cabang mempunyai kelas yang sama ataupun tidak ada atribut di dalam tupel yang berada di dalam partisi lagi atau ada tupel dalam cabang yang kosong.

Agar alur penelitian ini dapat dilakukan secara sistematis, maka dijabarkan dalam beberapa langkah, yaitu :

- a. Melakukan studi literatur yang bertemakan teori-teori Data dan Algoritma C 4.5 yang terbaru untuk mendapatkan metode yang signifikan dengan permasalahan yang ada.
- b. Pemilihan objek yang menjadi atribut yang akan memudahkan peneliti dalam melakukan klasifikasi tingkat pemahaman mahasiswa terhadap pembelajaran selama daring.
- c. Pada penelitian ini digunakan 5 atribut yang akan digunakan untuk klasifikasi terhadap tingkat pemahaman mahasiswa. Atributnya yang akan digunakan adalah sebagai berikut:
 1. Komunikasi (C1), atribut komunikasi merupakan faktor utama antara mahasiswa dan dosen dalam proses pembelajaran baik secara tatap muka ataupun daring.
 2. Media Pembelajaran (C2), atribut ini merupakan alat bantu yang digunakan oleh dosen dalam penyampaian materi. Hal ini dapat mempengaruhi pusat perhatian mahasiswa
 3. Suasana pembelajaran (C3), atribut ini merupakan bagaimana cara dosen untuk membangun suasana kelas yang dapat membantu mahasiswa dalam memahami mata kuliah

4. Cara mengajar (C4), atribut ini memperlihatkan cara dosen dalam mengajar mahasiswa. Apakah kaku atau memiliki cara tersendiri untuk menarik perhatian mahasiswa di dalam kelas
 5. Minat dan Motivasi Mahasiswa (C5), atribut ini merupakan bagaimana cara mahasiswa memotivasi dirinya sendiri dalam memahami pembelajaran dalam kelas
- d. Menentukan nilai class untuk masing-masing atribut. Nilai class untuk atribut (C) bisa dilihat pada tabel dibawah ini :

Tabel 1. Jenis Kelas Data

Atribut	Nama Field	Kelas Data
C1	Komunikasi	Satu arah, Kadang-kadang, Dua arah
C2	Media Pembelajaran	Baik, Cukup, Kurang baik
C3	Suasana Pembelajaran	Sangat baik, Baik, Cukup, Kurang baik
C4	Cara Mengajar	Sangat baik, Baik, Cukup, Kurang baik
C5	Minat dan Motivasi Mahasiswa	Sangat baik, Baik, Cukup, Kurang

- e. Melakukan analisis untuk klasifikasi tingkat pemahaman mahasiswa terhadap pembelajaran selama daring sesuai dengan entropi berdasarkan masing-masing atribut dengan menggunakan Algoritma C 4.5

4. Hasil dan pembahasan

Pada bagian ini merupakan hasil penelitian yang telah dilakukan. Data yang digunakan adalah data kuisioner yang telah dibagikan kepada mahasiswa Universitas Dharmawangsa, yang tercatat dalam format excel yang akan diolah agar mendapatkan informasi yang bermanfaat. Kemudian data diolah menggunakan software rapidminer untuk mendapatkan hasil pohon keputusan dan aturan-aturan baru. Untuk hasil pohon keputusan dari data yang telah diolah dapat dilihat pada Gambar 2 dibawah ini:



Gambar 2. Hasil Pohon Keputusan

Dan untuk deskripsi yang dihasilkan dari pohon keputusan yang telah diolah melalui software Rapidminer dapat dilihat pada Gambar 3 dibawah ini :

Tree

```
MEDIA PEMBELAJARAN = BAIK
| MINAT DAN MOTIVASI = BAIK: PAHAM {PAHAM=3, TIDAK PAHAM=0, PAHAM=0}
| MINAT DAN MOTIVASI = CUKUP: PAHAM {PAHAM=1, TIDAK PAHAM=1, PAHAM=0}
| MINAT DAN MOTIVASI = KURANG
| | KOMUNIKASI = DUA ARAH: PAHAM {PAHAM=1, TIDAK PAHAM=1, PAHAM=0}
| | KOMUNIKASI = KADANG-KADANG: TIDAK PAHAM {PAHAM=0, TIDAK PAHAM=2, PAHAM=0}
| MINAT DAN MOTIVASI = SANGAT BAIK: PAHAM {PAHAM=8, TIDAK PAHAM=0, PAHAM=0}
MEDIA PEMBELAJARAN = CUKUP: TIDAK PAHAM {PAHAM=2, TIDAK PAHAM=7, PAHAM=0}
MEDIA PEMBELAJARAN = KURANG BAIK
| KOMUNIKASI = DUA ARAH: PAHAM {PAHAM=0, TIDAK PAHAM=0, PAHAM=2}
| KOMUNIKASI = KADANG-KADANG: TIDAK PAHAM {PAHAM=0, TIDAK PAHAM=3, PAHAM=1}
| KOMUNIKASI = SATU ARAH: PAHAM {PAHAM=2, TIDAK PAHAM=1, PAHAM=0}
```

Gambar 3. Deskripsi Pohon keputusan

Jika dilihat dari pohon keputusan, atribut yang dijadikan akar adalah media pembelajaran dari dosen. Hal ini yang menjadi faktor utama dalam penentuan tingkat pemahaman mahasiswa Universitas Dharmawangsa dalam pembelajaran daring selama Covid-19.

Dari pohon keputusan dapat dijabarkan aturan-aturan yang menjadi pengetahuan baru yang berguna, yaitu :

1. Jika Media Pembelajaran = Baik, Minat dan Motivasi = Baik, maka Paham
 2. Jika Media Pembelajaran = Baik, Minat dan Motivasi = Cukup, maka Paham
 3. Jika Media Pembelajaran = Baik, Minat dan Motivasi = Kurang, Komukasi = Dua arah, maka Paham
 4. Jika Media Pembelajaran = Baik, Minat dan Motivasi = Kurang, Komukasi = Kadang-kadang, maka Tidak Paham
 5. Jika Media Pembelajaran = Cukup, maka Tidak Paham
 6. Jika Media Pembelajaran = Kurang Baik, Komunkasi = Dua arah, maka Paham
 7. Jika Media Pembelajaran = Kurang Baik, Komunkasi = Kadang-kadang, maka Tidak Paham
 8. Jika Media Pembelajaran = Kurang Baik, Komunkasi = Satu arah, maka Paham.
- (Kurniawan et al., 2020) (Amalia et al., 2018)

5. Kesimpulan

Dari hasil penelitian dan pembahasan dengan menggunakan algoritma FP-Growth lebih efisien dibandingkan algoritma apriori pada data mining karena FP-Growth memberikan pola yang lebih dinamis dan relevan dipakai pada masa yang mendatang. Hasil penelitian yang dilakukan dan diperoleh dapat ditarik kesimpulan bahwa pemahaman mahasiswa terkait pembelajaran daring selama Covid-19 yaitu delapan (8) aturan rule keputusan dari sasaran yang ingin dicapai yakni lima (5) keputusan Paham dan tiga (3) keputusan Tidak Paham. Dari pohon keputusan terlihat bahwa media pembelajaran merupakan paktor utama yang menjadi akar dari pemahaman mahasiswa berdasarkan perhitungan Algoritma C 4.5 Pengaruh ini dapat dilihat dari bagaimana seorang dosen bisa membuat kelas senyaman mungkin untuk mahasiswa melakukan proses belajar mengajar terutama dalam kelas daring pada masa Covid-19.(Wiguna & Riana, 2020) (Rofiqo et al., 2019)(Novika et al., 2021) (Wiza & Febriadi, 2019) (Oktaria et al., 2019)(Mazhona et al., 2020)(Istiawan & Khikmah, 2019)(Umam et al., 2020)

REFERENSI

- [1] Amalia, H., Puspitasari, A., & Lestari, A. F. (2018). *STUDENT PERFORMANCE ANALYSIS USING C4.5 ALGORITHM TO*. 149–154.
- [2] Istiawan, D., & Khikmah, L. (2019). Implementation of C4.5 Algorithm for Critical Land Prediction in Agricultural Cultivation Areas in Pemali Jratun Watershed. *Indonesian Journal of Artificial Intelligence and Data Mining*, 2(2), 67. <https://doi.org/10.24014/ijaidm.v2i2.7569>
- [3] Kurniawan, D., Anggrawan, A., & Hairani, H. (2020). Graduation Prediction System On Students Using C4.5 Algorithm. *MATRIK : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, 19(2), 358–365. <https://doi.org/10.30812/matrik.v19i2.685>
- [4] Mazhona, I. O., Studi, P., & Komputer, M. (2020). *Satisfaction Level of BPJS Kesehatan Participants Using the C4.5 Algorithm*. 2(3), 132–143.
- [5] Novika, T., Poningsih, P., Okprana, H., Windarto, A. P., & Siahaan, H. (2021). Penerapan Data Mining Klasifikasi Tingkat Pemahaman Siswa Pada Pelajaran Matematika. *Jurnal Media Informatika Budidarma*, 5(1), 9–17. <https://doi.org/10.30865/mib.v5i1.2498>
- [6] Oktaria, R., Komarudin, M., & Muda, M. A. (2019). Analisa Klasifikasi Kualitas Mahasiswa Lulusan Berdasarkan Jalur Penerimaan Menggunakan Algoritma C4.5 (Studi Kasus: Universitas Lampung). *Jurnal Teknik Informatika*, 12(2), 183–192. <https://doi.org/10.15408/jti.v12i2.11171>
- [7] Rofiqo, N., Windarto, A. P., & Irawan, E. (2019). Penerapan Algoritma C4.5 pada Penentuan Tingkat Pemahaman Mahasiswa Terhadap Matakuliah. *Prosiding Seminar Nasional Riset Information Science (SENARIS)*, 1(September), 307. <https://doi.org/10.30645/senaris.v1i0.36>
- [8] Umam, K., Puspitasari, D., & Nurhadi, A. (2020). Penerapan Algoritma C4.5 Untuk Prediksi Loyalitas Nasabah PT Erdika Elit Jakarta. *Jurnal Media Informatika Budidarma*, 4(1), 65. <https://doi.org/10.30865/mib.v4i1.1652>
- [9] Wiguna, W., & Riana, D. (2020). Diagnosis of Coronavirus Disease 2019 (Covid-19) Surveillance Using C4.5 Algorithm. *Jurnal Pilar Nusa Mandiri*, 16(1), 71–80. <https://doi.org/10.33480/pilar.v16i1.1293>
- [10] Wiza, F., & Febriadi, B. (2019). Classification Analysis Using C4.5 Algorithm To Predict The Level of Graduation of Nurul Falah Pekanbaru High School Students. *IJISTECH (International Journal Of Information System & Technology)*, 2(2), 43. <https://doi.org/10.30645/ijistech.v2i2.21>