

## PENERAPAN ALGORITMA C4.5 UNTUK PENGUKURAN TINGKAT KEPUASAN SISWA TERHADAP KINERJA GURU

**Muhammad Basyier Ardima<sup>1\*</sup>**

<sup>1</sup>Prodi Sistem Informasi, Universitas Semarang  
<sup>1</sup>ardima@usm.ac.id

**Shona Chay Bilqis<sup>2</sup>**

<sup>2</sup>Prodi Sistem Informasi, Universitas Semarang  
<sup>2</sup>shona@mail.unnes.ac.id

**Mochamad Idris<sup>3</sup>**

<sup>2</sup>Prodi Sistem Informasi, Universitas Semarang  
<sup>3</sup>idrez.mochamad@gmail.com

### ABSTRACT

*This research aims to determine the level of student satisfaction with teacher performance at MA Miftahussalam Demak by applying the C4.5 algorithm. By knowing the level of student satisfaction, the school can improve services in terms of teaching if they are not good and further improve services if the level of satisfaction is good. The method used in this research uses data mining techniques, classification methods with the C.45 algorithm, where the data source used is a questionnaire involving 3 sample classes, namely 1 class X, 1 class XI, and 1 class XII from the total number of MA students. Miftahussalam Demak. The attributes used as parameters for assessing student satisfaction include: mastery of the material, explanation of the material, time discipline, concern, and communication. The benefit of this research is to improve the quality of teacher performance and improve the quality of teacher teaching. Of the 118 data, 104 students could be classified as satisfied and 14 students as dissatisfied with the teacher's performance with an accuracy of 96.6%.*

**Keywords:** Classification, satisfaction, E-learning, C4.5.

### 1. PENDAHULUAN

Dalam upaya untuk meningkatkan mutu pendidikan, peran guru menjadi faktor penentu dalam keberhasilan proses pembelajaran. Oleh karena itu, penting untuk memahami dan mengukur kepuasan siswa terhadap kinerja guru agar dapat mengidentifikasi area yang perlu ditingkatkan dalam pengajaran dan pembelajaran di lingkungan sekolah. 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. Berikut beberapa masalah yang akan digunakan sebagai bahan penelitian dari beberapa pertanyaan yang ada di koesioner: Bagaimana faktor disiplin waktu kedatangan guru, Strategi pembelajaran

yang diberikan oleh guru, materi yang diberikan oleh guru, evaluasi yang diberikan oleh guru. Dari permasalahan diatas, peneliti menggunakan teknik data mining pengukuran dengan algoritma C4.5 untuk mengetahui kepuasan siswa terhadap kinerja guru MA Miftahussalam Demak.

Berdasarkan uraian diatas, salah satu pendekatan yang digunakan untuk mengatasi permasalahan klasifikasi, termasuk klasifikasi kepuasan siswa terhadap kinerja guru, adalah dengan menerapkan algoritma C4.5. Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Sedang pohon keputusan dapat diartikan suatu cara untuk memprediksi atau mengklasifikasi yang sangat kuat. Pohon keputusan dapat membagi kumpulan data yang besar menjadi himpunan-himpunan record yang lebih kecil dengan menerapkan serangkaian aturan keputusan (Irnanda, 2020). Algoritma ini telah terbukti efektif dalam mengklasifikasikan data dengan aturan-aturan yang mudah dimengerti dan dapat diterapkan dalam berbagai bidang termasuk bidang pendidikan. Dalam konteks

penelitian ini, kami akan mengidentifikasi kontribusi utama dari algoritma C4.5 dalam klasifikasi kepuasan siswa terhadap kinerja guru untuk memberikan model aturan dalam menganalisa apakah siswa puas atau tidak puas terhadap kinerja pengajaran guru adapun atribut yang digunakan sebagai parameter penilaian kepuasan siswa antara lain: penguasaan materi, penjelasan materi, disiplin waktu, kepedulian, dan komunikasi.

## 2. METODE PENELITIAN

### 2.1 Data Mining

Data mining adalah sebuah *knowledge discovery in database* (KDD). Data Mining merupakan salah satu cara yang digunakan untuk mendapatkan pengetahuan baru dengan memanfaatkan jumlah data yang sangat besar. Beberapa teknik telah dikembangkan dan dimplementasikan untuk mengumpulkan pengetahuan yang mungkin berguna untuk pengambilan keputusan. Teknik-teknik yang digunakan untuk mengekstrakan pengetahuan dalam data mining adalah pengenalan pola, clustering, asosiasi, prediksi dan klasifikasi (Defiyanti, 2015).

### 2.2 Konsep Decision Tree

*Decision Tree* atau pohon keputusan adalah pohon yang digunakan sebagai prosedur penalaran untuk mendapatkan jawaban dari masalah yang dimasukkan. Pohon yang dibentuk tidak selalu berupa pohon biner. Jika semua fitur dalam data set menggunakan 2 macam nilai kategorikal maka bentuk pohon yang didapatkan berupa pohon biner. Jika dalam fitur berisi lebih dari 2 macam nilai kategorikal atau menggunakan tipe numerik maka bentuk pohon yang didapatkan biasanya tidak berupa pohon biner [9].

*Decision Tree* terdapat 3 jenis Node, yaitu [12] :

1. *Root Node*, merupakan *node* paling atas, pada *node* ini tidak ada *input* dan bisa tidak mempunyai *output* atau mempunyai *output* lebih dari satu.
2. *Internal Node*, merupakan *node* percabangan, pada *node* ini hanya terdapat satu *input* dan mempunyai *output* minimal dua. *Lead Node* atau *Terminal Node*, merupakan *node* akhir, pada *node* ini hanya terdapat satu *input* dan tidak mempunyai *output*.

Algoritma C4.5 merupakan metode yang dapat digunakan untuk melakukan pembentukan pohon keputusan. Decision tree tersebut mampu menghasilkan keputusan yang kompleks menjadi lebih sederhana, sehingga pengambil keputusan akan lebih menginterpretasikan solusi dari permasalahan. Secara umum metode C4.5 untuk membangun decision tree adalah sebagai berikut (Ardiansyah Sembiring et al., 2018):

#### a. Pilih atribut sebagai akar.

Langkah-langkah untuk memperoleh atribut sebagai akar adalah dengan menghitung jumlah kasus

dan jumlah target atribut. Setelah itu menghitung nilai entropy yang digunakan untuk menentukan seberapa informatif sebuah input atribut untuk menghasilkan output atribut. Rumus dasar dari entropy tersebut

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

Keterangan :

- S : himpunan kasus  
N : jumlah partisi S  
pi : proporsi dari Si terhadap S

Persamaan di atas nantinya digunakan untuk menghitung entropy total pada kasus ini. Adapun jumlah kasus (S) yang nantinya akan dihitung dengan total kasus awal, sedangkan untuk n (jumlah partisi S) adalah banyaknya respon tidak puas dan respon puas dalam kasus ini. Setelah menghitung entropy setiap kasus, maka digunakan informasi gain untuk pemisahan obyek. Dengan menggunakan rumus:

$$Gain(S,A) = Entropy(S) - \sum_{i=1}^n \left[ \left( \frac{|S_i|}{|S|} \right) \times Entropy(S_i) \right] \quad (2).$$

Keterangan:

- S : himpunan kasus  
A : atribut  
n : jumlah partisi atribut A  
|Si| : jumlah kasus pada partisi ke-i  
|S| : jumlah kasus dalam S

Untuk S (himpunan kasus), sedangkan atribut (A) yang digunakan dalam kasus ini meliputi atribut penguasaan materi, penjelasan materi, disiplin waktu, kepedulian, dan komunikasi. Sedangkan n adalah jumlah partisi dari masing-masing atribut. Setelah semua atribut dihitung menggunakan persamaan diatas, maka atribut yang memiliki nilai informasi tertinggi dibanding atribut yang lain dijadikan sebagai node (akar).

#### b. Buat cabang.

Setelah diperoleh atribut yang mempunyai nilai gain tertinggi, maka atribut tersebut digunakan sebagai node. Node ini memiliki instance sehingga instance dijadikan sebagai cabang dari node.

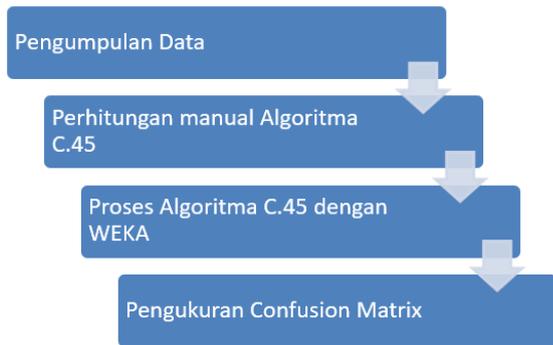
#### c. Bagi kasus dalam cabang.

Setiap nilai pada instance memiliki nilai yang berbeda. Nilai instance ini 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.

### 2.3 Kerangka Pemikiran

Kerangka pemikiran merupakan garis besar dari langkah – langkah penelitian yang sedang dilakukan,

kerangka pemikiran dijadikan acuan untuk melakukan tahap – tahap yang sedang dilakukan dalam penelitian.



Gambar 1. Kerangka berfikir

2.4 Jenis Data

Jenis data yang digunakan dalam penelitian ini yaitu:

- a. Data primer yaitu data yang diperoleh secara langsung dari objek penelitian, sebagai berikut : Kegiatan yang ada di fakultas TIK, Jumlah peserta, kemudian ranking yang akan di dapatkan.
- b. Data sekunder yaitu data yang diperoleh dari literature, buku referensi, maupun browsing internet.

2.5 Metode Pengumpulan Data

Data yang digunakan untuk kepuasan penggunaan dalam penelitian berupa penyebaran kuesioner. Dengan menyebarkan kuesioner kepada siswa MA Miftahussalam. Setelah mendapatkan data, penulis kemudian akan mengolah dengan metode Algoritma C4.5 dan mendapatkan hasil yang akurat saat melakukan analisis dalam kepuasan pengguna terhadap kinerja guru.

2.6 Tahapan Penelitian

Pengumpulan kriteria atau atribut untuk digunakan pada saat pengelolaan data.

- a. Pembuatan kuesioner sesuai dengan atribut yang sudah dikumpulkan.
- b. Proses data pertama yaitu dengan memisahkan data kuesioner yang sudah terkumpul ke dalam suatu file , lalu disimpan di Microsoft Excel 2016.
- c. Kemudian data yang sudah tersimpan diubah ke dalam bentuk format .csv.
- d. Implementasi dengan Weka 3.8
- e. Buka aplikasi Weka 3.8, pilih tool Classify, dan pilih penggunaan metode yang akan digunakan. Penulis menggunakan metode c.45 dalam penelitian ini.
- f. Lakukan pengujian dengan berbagai variasi untuk mendapatkan hasil akurasi yang tinggi
- g. Hasil akan didapatkan setelah melakukan proses dengan menggunakan metode c.45.

3. HASIL DAN PEMBAHASAN

3.1 Penentuan Kriteria

Klasifikasi berdasarkan atribut-atribut memungkinkan analisis yang lebih mendalam terhadap faktor-faktor yang mempengaruhi kepuasan siswa. Pada penelitian ini mengidentifikasi 5 kriteria penilaian, yaitu : penguasaan materi, penjelasan materi, disiplin waktu, kepedulian, dan komunikasi. Alternatif yang digunakan pada penelitian adalah nama- nama siswa yang menjawab pertanyaan dari kuesioner. Berikut ini adalah kriteria dalam menentukan hasil berdasarkan ranking menggunakan metode klasifikasi: C.45 berikut ini :

Tabel 3.1 Tabel Kriteria

No	Kriteria/Atribut	Keterangan
1	Penguasaan Materi	Sangat Menguasai (SM), Menguasai (M), Kurang Menguasai (KM).
2	Penjelasan Materi	Sangat Jelas (SJ), Jelas (J), Kurang Jelas (KJ).
3	Disiplin Waktu	Sangat Tepat (ST), Tepat (T), Kurang Tepat (KT).
4	Kepedulian	Sangat Peduli (SP), Peduli (P), Kurang Peduli (KP).
5	Komunikasi	Sangat Baik (SB), Baik (B), Kurang Baik (KB).

3.2 Data

Dalam penelitian ini data yang digunakan adalah hasil dari kuesioner yang diberikan kepada siswa Madrasah Aliyah Miftahussalam Demak. Dari hasil kuesioner yang dilakukan penulis mengambil sampel sebanyak 118 responden dari 3 kelas siswa MA Miftahussalam berikut ini:

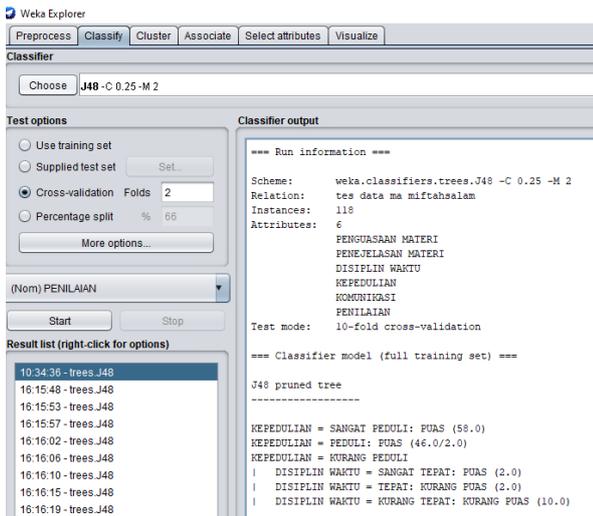
Tabel 3.2 Tabel Data Penelitian

Nama Siswa	Kelas	Pngsaan Materi	Pnjlasan Materi	Disiplin Waktu	Kepedu lian	Komu nkasi	Penil aian
ssw1	x	Sangat Menguasai	Jelas	Sangat Tepat	Sangat Peduli	Sangat Baik	Puas
ssw2	X	Menguasai	Jelas	Tepat	Sangat Peduli	Baik	Puas
Ssw3	X	Sangat Menguasai	Sangat Jelas	Sangat Tepat	Sangat Peduli	Sangat Baik	Puas
ssw4	X	Sangat Menguasai	Sangat Jelas	Sangat Tepat	Sangat Peduli	Sangat Baik	Puas
ssw5	X	Kurang Menguasai	Sangat Jelas	Sangat Tepat	Peduli	Sangat Baik	Puas
ssw6	X	Menguasai	Sangat Jelas	Tepat	Peduli	Baik	Kurang Puas
ssw7	X	Menguasai	Jelas	Sangat Tepat	Peduli	Baik	Puas
ssw 118	X	Menguasai	Kurang Jelas	Tepat	Kurang Peduli	Baik	Kurang Puas

Data yang telah didapatkan kemudian di ubah kedalam format .csv agar dapat di implementasikan kedalam aplikasi Weka 3.8.

### 3.3 Implementasi Klasifikasi Naive Bayesian Pada Weka

Implementasi merupakan sebuah proses menentukan informasi dari data yang digunakan yaitu metode prediksi dengan algoritma naive bayes dengan software Weka 3.8. Prosesnya yaitu klik explorer ->open file ->pilih file .csv yang akan kita proses ->open->classify-> pilih algoritmanya. maka tampilannya seperti dibawah ini :



Gambar 2. Tampilan Weka

Dari gambar 2 menunjukkan terdapat 6 atribut yaitu penguasaan materi, penjelasan materi, disiplin waktu, kepedulian, dan komunikasi, serta atribut penilaian sebagai atribut klasifikasi puas atau kurang puas. Untuk hasil klasifikasi lebih jelas ditunjukkan pada gambar 3 sebagai berikut:

```

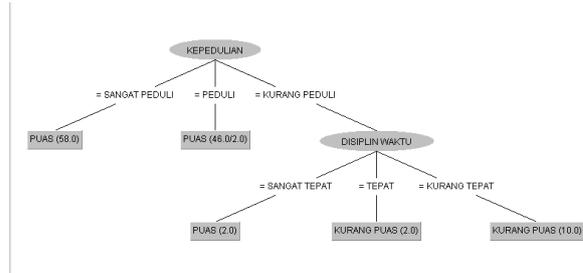
Correctly Classified Instances 114 96.6102 %
Incorrectly Classified Instances 4 3.3898 %
Kappa statistic 0.8379
Mean absolute error 0.0503
Root mean squared error 0.1769
Relative absolute error 23.437 %
Root relative squared error 54.599 %
Total Number of Instances 118

==== Detailed Accuracy By Class ====
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area FRC Area Class
0.981 0.143 0.981 0.981 0.981 0.838 0.940 0.990 PUAS
0.857 0.019 0.857 0.857 0.857 0.838 0.940 0.870 KURANG PUAS
Weighted Avg. 0.966 0.128 0.966 0.966 0.966 0.838 0.940 0.975

==== Confusion Matrix ====
a b <-- classified as
102 2 | a = PUAS
2 12 | b = KURANG PUAS
    
```

Gambar 3. Hasil Klasifikasi

Berdasarkan gambar diatas menunjukkan hasil klasifikasi dari 118 record data, yaitu 104 siswa menyatakan puas dan 14 siswa menyatakan kurang puas terhadap kinerja guru. Adapun tree yang terbentuk dapat dilihat pada gambar 5.7 pohon keputusan berikut ini:



Gambar 3. Pohon Keputusan

Cabang pertama muncul pada kriteria Kepedulian yaitu sangat peduli puas 58, peduli puas 46, dan kurang peduli menghasilkan cabang disiplin waktu yaitu sangat tepat puas 2, tepat kurang puas 2, dan kurang tepat kurang puas 10.

### 3.4 Pengukuran Confusion Matrix

Pada tahap pengukuran confusion matrix untuk mengevaluasi hasil dilakukan dengan menggunakan cross validation dengan nilai K=2, K=3, K=4, K=5, K=6, K=7, K=8, K=9, K=10. Pengujian evaluasi menggunakan tools weka dan confusion matrix. Berikut merupakan hasil algoritma C.45 dengan menggunakan cross validation mengenai nilai akurasi, precision, dan recall, yang ditampilkan pada tabel 3.3

Tabel 3.3 Tabel Pengujian

K=n	Akurasi	Precision	Recall
K=2	91.5 %	0.907	0.915
K=3	96.6 %	0.966	0.966
K=4	96.6 %	0.966	0.966
K=5	96.6 %	0.966	0.966
K=6	96.6 %	0.966	0.966
K=7	96.6 %	0.966	0.966
K=8	96.6 %	0.966	0.966
K=9	96.6 %	0.966	0.966

Dapat diambil kesimpulan dengan menggunakan cross validation nilai K yang berbeda memberikan hasil evaluasi yang sama, hanya nilai K=2 yang menunjukkan nilai yang sedikit berbeda, dengan hasil akurasi terbaik yaitu 96.6 %, precision sebesar 0.966, dan recall sebesar 0.966. Dari proses pada WEKA data yang diproses menghasilkan 4 data error, dari 118 data. Tingginya nilai akurasi, precision, dan recall menunjukkan konsistensi dan keandalan model Algoritma C4.5 dalam memprediksi kepuasan siswa berdasarkan atribut yang telah ditentukan diatas. Konsistensi hasil ini mendukung validitas model dan sesuai dengan teori yang menunjukkan bahwa factor seperti kepedulian dan disiplin waktu mempengaruhi kepuasan siswa dalam kinerja guru. Analisis terhadap data error menunjukkan bahwa factor individual siswa atau kondisi khusus yang tidak tercakup dalam model dapat mempengaruhi hasil. Hal ini diketahui dari hasil confusion matrix, berikut diberikan tabel pengetahuan dari hasil confusion matrix.

Tabel 3.4 Hasil Confusion Matrix

Jumlah Data	Jumlah Error	True Positive	False Positive	False Negative	True Negative
118	4	102	2	12	2

Dari tabel diatas dengan data sebanyak 118 yang diuji terdapat kasus *error* yaitu 4 data, kasus dengan hasil dan prediksi yang tidak sama. TP (*true positive*) ialah data dengan hasil puas dan prediksinya juga puas 102 data. FP (*false positive*) yaitu data dengan hasil puas namun prediksinya tidak puas 2 data. FN (*false negative*) yaitu data dengan hasil tidak puas namun prediksinya puas 22 data, dan TN (*true negative*) ialah data dengan hasil tidak puas dan prediksi tidak puas 12 data.

Brahmana : Jurnal Penerapan Kecerdasan Buatan, 2(1), pp. 33–40. doi: 10.30645/brahmana.v2i1.46.

#### 4. KESIMPULAN

Dari penelitian yang dilakukan dengan menggunakan 118 data hasil kuisioner siswa MA Miftahussalam Demak dapat diambil kesimpulan bahwa baik perhitungan manual dengan menggunakan *Algoritma C4.5* ataupun perhitungan menggunakan alat bantu *WEKA 3.8.3* mempunyai hasil yang sama untuk model *Pohon keputusan*, yaitu “Kepedulian” yang menjadi akar. Perhitungan *Gain* yang tertinggi yang sama yaitu faktor kepuasan siswa yang mempunyai akar tertinggi ada pada “Kepedulian”.

#### 5. UCAPAN TERIMA KASIH

Peneliti ucapkan terimakasih kepada pihak Dikti dan Universitas Semarang yang telah membiayai penelitian ini.

#### 6. DAFTAR PUSTAKA

- ARDIANSYAH, S. M. (2016). Penerapan Metode Decision Tree Algoritma C4.5 Untuk Memprediksi Hasil Belajar Mahasiswa Berdasarkan Riwayat Akademik. *Jurnal Teknologi dan Sistem Informasi (JURTEKSI ROYAL)*. Vol 3 No. 1
- DEFIYANTI, SOFI & MOHAMAD JAJULI. (2015). Integrasi Metode Klasifikasi Dan Clustering Dalam Data Mining.” *Konferensi Nasional Informatika (KNIF)*, 39–44.
- FADILLAH, W. R. et al. (2020). Implementasi Data Mining C4.5 Dalam Mengukur Tingkat Kepuasan Mahasiswa Terhadap Kinerja Asisten Laboratorium Komputer’, *Prosiding Seminar Nasional Riset Dan Information Science (SENARIS)*, 2, pp. 403–414
- GERARDO, B. G., SAIFULLAH, S. & IRAWAN, E. (2019). Teknik Data Mining Dalam Penilaian Pengajaran Guru Berdasarkan Indeks Kepuasan Siswa, *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, 3(1), pp. 508–514. doi: 10.30865/komik.v3i1.1634..
- IRNANDA. (2020). Penerapan Klasifikasi C4.5 Dalam Meningkatkan Kecakapan Berbahasa Inggris dalam Masyarakat. *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS)*, pp. 304–308..
- PRIBADI, E, S., PONINGSIH. & TAMBUNAN, H. S (2020). Analisa Tingkat Kepuasan Masyarakat Terhadap Pelayanan Pengadilan Agama Pematangsiantar Menggunakan Algoritma C4.5.