

# Klasifikasi Kepuasan Siswa Terhadap Kinerja Tentor Pengajar di CV. Ibsi Edutama Menggunakan *Algoritma C5.0*

Amalia Fandaning Tyas<sup>1</sup>, Andi Iwan Nurhidayat<sup>2</sup>

D4 Manajemen Informatika, Universitas Negeri Surabaya  
Kampus Ketintang, Jalan Ketintang Surabaya 60231

[1amalia.19012@mhs.unesa.ac.id](mailto:1amalia.19012@mhs.unesa.ac.id)

[2andyl34k5@unesa.ac.id](mailto:2andyl34k5@unesa.ac.id)

**Abstrak**— CV Ibsi Edutama ini adalah sebuah Lembaga bimbingan belajar yang terletak di Surabaya. Dalam melakukan proses pembelajaran, perusahaan ini menggunakan tentor atau pengajar sebagai Pendidik bagi para siswa. Mengukur kepuasan siswa akan sangat membantu CV. Ibsi Edutama dalam meningkatkan kualitas pelayanan yang ada untuk meningkatkan jumlah siswa. Tmentor yang berpengalaman dan berkompeten dalam kemampuannya dalam bidang tersebut tentu mudah dalam menyampaikan materi kepada para siswa/siswi dengan lancar. Tidak semua kinerja tentor di CV. Ibsi Edutama berkinerja dengan baik, terkadang ada beberapa tentor yang memiliki kualitas belajar mengajar yang buruk. Oleh karena itu, untuk mendapatkan tentor yang berkualitas, pihak CV. Ibsi Edutama tentu saja melakukan evaluasi terhadap setiap tentor yang mengajar di CV. Ibsi Edutama. Hasil penelitian ini mengindikasikan bahwa *Algoritma Decision tree C5.0* dapat mengklasifikasikan tingkat kepuasan siswa dengan akurasi yang tinggi, mencapai 90%.

**Kata kunci**— *Algoritma C5.0, Decision Tree, Klasifikasi, Kepuasan*

**Abstract**— CV Ibsi Edutama is a tutoring institution located in Surabaya. In carrying out the learning process, this company uses tutors or instructors as educators for students. Measuring student satisfaction will greatly help CV. Ibsi Edutama in improving the quality of service which will ultimately lead to an increase in the number of students. Tutors who are experienced and competent in their fields of expertise certainly can easily convey material to students. Not all of the mentor's performance on the CV. Ibsi Edutama is performing well, sometimes there are some tutors who have poor teaching and learning quality. Therefore, to get qualified tentors, CV. Ibsi Edutama of course evaluates every tutor who teaches at CV. Ibsi Edutama. The results of this study indicate that the *Decision Tree Algorithm C5.0* is able to classify the level of student satisfaction with a fairly high level of accuracy, namely 90%.

**Keywords**— *C5.0 Algorithm, Decision Tree, Classification, Satisfaction*

## I. PENDAHULUAN

Saat ini, banyak perusahaan yang bergerak di bidang jasa pendidikan, termasuk bimbingan belajar, yang menghadapi persaingan yang sangat ketat. Oleh karena itu, diperlukan analisis untuk mengidentifikasi faktor-faktor yang dapat menyebabkan kepuasan siswa yang mengikuti bimbingan belajar terhadap kinerja tentor. Salah satu usaha yang bergerak

dibidang Pendidikan adalah CV. Ibsi Edutama. CV. Ibsi Edutama ini adalah Lembaga bimbingan belajar yang terletak di daerah Surabaya. Dalam melakukan proses pembelajaran, perusahaan ini menggunakan tentor atau pengajar sebagai tenaga pendidik bagi para siswa/siswi. Tmentor, sebagai tenaga pendidik, harus memiliki peran ganda, bukan hanya sebagai pengajar, tetapi juga sebagai pembimbing yang memberikan contoh positif kepada siswa. Kemajuan dan perkembangan siswa/siswi tidak hanya tergantung pada fasilitas yang memadai, tetapi juga sangat dipengaruhi oleh kualitas pengajaran dan pembimbingan yang diberikan oleh tentor.

Tmentor yang memiliki pengalaman dan kompetensi dalam bidangnya mampu dengan mudah mengkomunikasikan materi kepada para siswa/siswi. Meskipun demikian, tidak semua tentor di CV. Ibsi Edutama memiliki kinerja yang baik, dan beberapa di antaranya menghadapi tantangan dalam kualitas pembelajaran. Dampak dari proses pembelajaran yang kurang baik akan berimbas negatif pada kemajuan siswa, namun sebaliknya, ketika pengajaran dilakukan dengan baik, hal tersebut akan memberikan dampak positif pada perkembangan siswa.

Dengan demikian, untuk mendapatkan tentor yang berkualitas, pihak CV. Ibsi Edutama tentu saja melakukan evaluasi terhadap setiap tentor yang mengajar di CV. Ibsi Edutama tersebut dikarenakan banyaknya pesaing di bidang usaha jasa bimbingan belajar CV. Ibsi Edutama memerlukan sebuah analisis kepuasan dari siswa mengenai layanan dari tentor CV. Ibsi Edutama.

Judul Tugas Akhir ini yaitu "Klasifikasi Kepuasan Siswa Terhadap Kinerja Tentor Pengajar di CV. Ibsi Edutama." Penelitian ini diharapkan dapat membantu CV. Ibsi Edutama untuk meningkatkan kualitas layanan yang diberikan. Maka dari itu penulis ingin membuat klasifikasi kepuasan siswa dengan mengkategorikan beberapa kriteria dari kinerja tentor untuk mengetahui hasil yang akurat yang berdampak bagi kepuasan siswa berdasarkan kriteria yang ditentukan menggunakan *Algoritma C5.0* yang nanti diimplementasikan di sebuah website agar memudahkan CV. Ibsi Edutama untuk melihat hasil dari kepuasan siswa. Hasil analisis dari penelitian ini menunjukkan bahwa *Algoritma C5.0*. Dengan tingkat akurasi 90%, model klasifikasi yang telah dikembangkan dapat membantu CV. Ibsi Edutama untuk

memprediksi kepuasan siswa dan meningkatkan layanan yang diberikan kepada siswa/siswi.

Penelitian ini menerapkan metode *Algoritma C5.0*, sebuah perkembangan dari *Algoritma C4.5* dan *ID3*. *Algoritma C5.0* menunjukkan tingkat akurasi yang lebih tinggi dalam perhitungan dibandingkan dengan algoritma lain. Dan mampu menciptakan pohon keputusan dengan akurasi yang lebih optimal karena secara terus-menerus melakukan pembaruan pembobotan pada data training yang digunakan sebagai acuan, sehingga pohon keputusan selalu diperbarui dengan menggunakan data training yang terkini. Hal ini memastikan bahwa model pohon keputusan yang dihasilkan selalu merupakan yang terbaru.

## II. KAJIAN PUSTAKA

### A. Data Mining

*Data mining* merupakan suatu proses yang bertujuan untuk menggali informasi penting dari kumpulan data besar. *Data mining* merupakan bagian integral dari proses keseluruhan dalam penemuan pengetahuan dalam database (knowledge discovery in database). Bidang data mining menyatukan teknik-teknik dari pengenalan pola, pembelajaran mesin, statistik, manajemen database, dan visualisasi untuk mengatasi permasalahan dalam mengambil informasi dari database yang berskala besar. [1]

### B. Decision Tree

"*Decision tree* adalah struktur flowchart yang menyerupai tree (pohon), di mana setiap simpul internal menandakan suatu tes atribut, setiap cabang merepresentasikan hasil tes, dan simpul daun merepresentasikan kelas atau distribusi kelas". Struktur flowchart terdiri dari node internal yang menggambarkan tes atau atribut pada setiap tahap. Setiap cabang merepresentasikan hasil dari atribut tersebut, sementara jalur dari daun ke akar mewakili aturan untuk klasifikasi.[3]

### C. Algoritma C5.0

*Algoritma C5.0* merupakan perbaikan/pengembangan dari *algoritma C4.5*. dimana pengembangan dilakukan pada penggunaan memori, pemisahan atribut dan akselerasi. *C4.5* merupakan algoritma yang mengikuti aturan *ID3*, sama seperti *C5.0* juga mengikuti aturan *C4.5* sehingga memiliki aturan yang hampir sama dengan *C4.5*. *Algoritma C5.0*. *Algoritma C5.0* memiliki banyak fitur seperti pohon keputusan yang besar dapat dilihat sebagai seperangkat aturan yang mudah untuk dipahami. Dan juga *algoritma C5.0* ini memberikan pengakuan noise dan missing data. Begitupun masalah over fitting dan error dapat diselesaikan oleh *Algoritma C5.0*. Pengklasifikasian *C5.0* dapat mengantisipasi atribut mana yang relevan dan tidak.[4]

TABEL I  
PERBEDAAN ALGORITMA C4.5 DAN ALGORITMA C5.0

| No | Kategori | Algoritma C4.5 | Algoritma C5.0 |
|----|----------|----------------|----------------|
|----|----------|----------------|----------------|

|   |         |  |  |
|---|---------|--|--|
| 1 | Speed   | Membutuhkan waktu yang lebih lama dalam menentukan rules tree, yaitu 8 jam | Membutuhkan waktu yang cepat dalam menemukan rules tree, yaitu 3 menit         |
| 2 | Memory  | Memori yang digunakan mencapai 3 gb  | Memori yang digunakan hanya 200mb  |
| 3 | Akurasi | Memiliki tingkat error punning yang tinggi                                 | Memiliki error Purning lebih Rendah karena Terdapat fase menghitung gain ratio |

**Langkah-langkah dalam membuat pohon keputusan menggunakan *Algoritma C5.0* adalah sebagai berikut:**

1. Menyiapkan data training untuk mengambil informasi dari kuisioner dan dikelompokkan ke dalam kelas-kelas sesuai atribut.
2. Perhitungan dilakukan untuk menentukan akar pohon. Akar dari pohon dipilih berdasarkan atribut dengan nilai *gain* tertinggi setelah menghitung nilai *gain* dari masing-masing atribut dan atribut ini akan menjadi akar pertama dari pohon.
3. Sebelum menghitung nilai *gain* dari atribut, nilai *entropy* dihitung terlebih dahulu [5]. Rumus yang digunakan untuk menghitung nilai *entropy* adalah:

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

Keterangan:

- S : Himpunan kasus
- n : Jumlah partisi S
- pi : Proporsi dari Si terhadap S

4. Setelah mendapatkan nilai *entropy* dari setiap atribut, langkah berikutnya adalah menghitung information gain dengan menggunakan rumus:

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

Keterangan:

- S : Himpunan Kasus
- A : Atribut
- n : Jumlah partisi Atribut A

- $|S_i|$  : Jumlah kasus pada partisi ke-i
- $|S|$  : Jumlah kasus dalam S

5. Pada algoritma C4.5, jika perhitungan *information gain* sudah selesai, maka algoritma C5.0 akan melanjutkan ke tahap perhitungan *gain ratio* dengan menggunakan rumus:

$$\text{Gain Ratio} = \frac{\text{Information Gain (S,A)}}{\sum_{i=1}^n \text{Entropy (S}_i\text{)}}$$

6. Kemudian, proses perhitungan *gain ratio* diulangi hingga setiap cabang memiliki kelas masing-masing. Atribut yang telah dipilih sebelumnya tidak akan diikutsertakan lagi dalam perhitungan nilai *gain ratio* pada iterasi selanjutnya.

#### D. Python

Python adalah bahasa pemrograman yang populer dengan berbagai manfaat dalam Python mendukung pemrograman berorientasi objek dan dapat dijalankan di berbagai platform sistem operasi seperti PC, Macintosh, dan UNIX. Beberapa kelebihan Python antara lain:

1. Dapat membangun program dengan cepat dengan menggunakan kode yang lebih sedikit. Dukungan multiplatform.
2. Kemudahan dalam pembelajaran Python.
3. Sistem pengelolaan memori otomatis.
4. Python mendukung pemrograman berorientasi objek.[5]

#### E. Akurasi Pengujian

Pengujian akurasi sangat penting untuk memverifikasi kevalidan data yang telah dihitung. Pengujian akurasi dilakukan menggunakan *confusion matrix*, juga dikenal sebagai *cross-validation*, adalah metode pengujian akurasi yang melibatkan pembagian dataset menjadi bagian data *training* dan data *testing*. Dalam *confusion matrix* terdapat empat istilah yang merepresentasikan hasil klasifikasi, yaitu: [6]

TABEL II  
CONFUSION MATRIX

|                  |              | Actual Values |              |
|------------------|--------------|---------------|--------------|
|                  |              | Positive (1)  | Negative (0) |
| Predicted Values | Positive (1) | TP            | FP           |
|                  | Negative (0) | FN            | TN           |

Berikut adalah beberapa istilah dalam evaluasi klasifikasi:

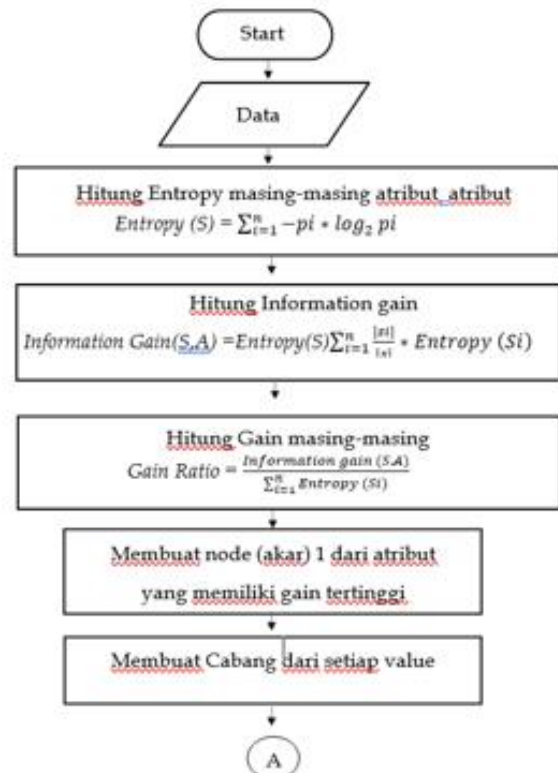
- True Positive (TP)  
Merupakan prediksi positif yang benar.
- True Negative (TN)  
Merupakan prediksi negatif yang benar.
- False Positive (FP)  
Merupakan prediksi positif yang salah.
- False Negative (FN)  
Merupakan prediksi negatif yang salah.

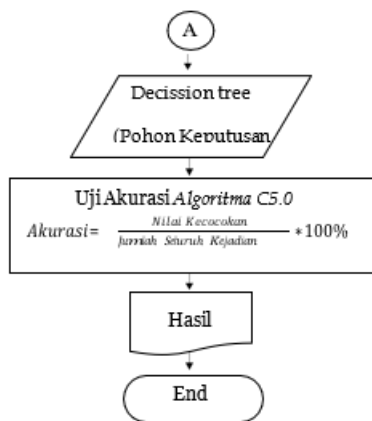
Dari tabel *confusion matrix* diatas, dapat dihitung akurasi sebagai berikut:

$$\text{Akurasi} : \frac{TP+TN}{(TP+TN+FP+FN)} \times 100\%$$

### III. METODOLOGI

#### A. Rancangan Penelitian





Gambar. 1 Flowchart Analisis Algoritma C5.0

### 1. Pengumpulan Data

Pada tahap pengambilan data penulis mengumpulkan beberapa data melalui kuisioner yang disebarakan kepada para siswa CV. Ibsi Edutama berupa data kepuasan siswa terhadap kualitas pengajaran oleh tentor, kedisiplinan tentor, sikap tentor saat mengajar, penampilan dari tentor, penguasaan materi oleh tentor, penjelasan saat mengajar, suasana belajar, kemudian hasil belajar yang didapat oleh siswa apa sudah sesuai dengan keinginan. Data set ini diperoleh melalui penyebaran kuisioner kepada siswa/siswi CV. Ibsi Edutama dan kemudian diproses menggunakan MS Excel. Tujuan dari pengolahan data set ini adalah untuk mengembangkan perhitungan pohon keputusan dengan melalui perhitungan *Entropy*, *information gain*, dan *gain ratio* pada masing-masing atribut.

### 2. Teknik Kategorisasi Kepuasan

Kuesioner ini terdiri dari 8 pernyataan yang dinilai dengan skor 1-4. Untuk memvalidasi kuesioner, dilakukan pengecekan kesesuaian setiap pernyataan yang akan diujicobakan dengan indikator tertentu. Skala Likert digunakan sebagai rentang skor dalam penelitian ini dengan model empat pilihan, yang mewakili nilai-nilai sebagai berikut: [7]

- 4: Sangat Baik
- 3: Baik
- 2: Kurang Baik
- 1: Tidak Baik

TABEL III  
KATEGORI KEPUASAN 2 JENJANG

| Rumus        | Kategori   |
|--------------|------------|
| $X \geq \mu$ | Puas       |
| $X < \mu$    | Tidak Puas |

Keterangan :

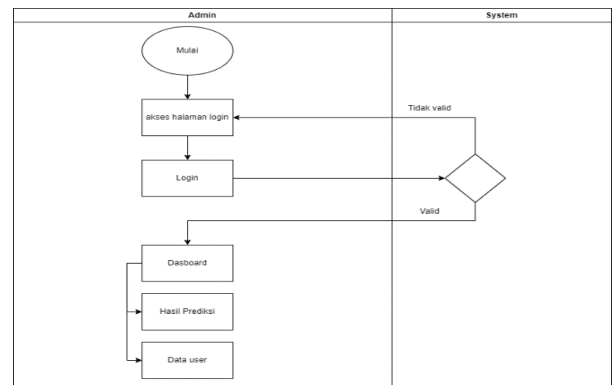
X : Skor mentah sampel

$\mu$  : rata-rata distribusi dalam populasi

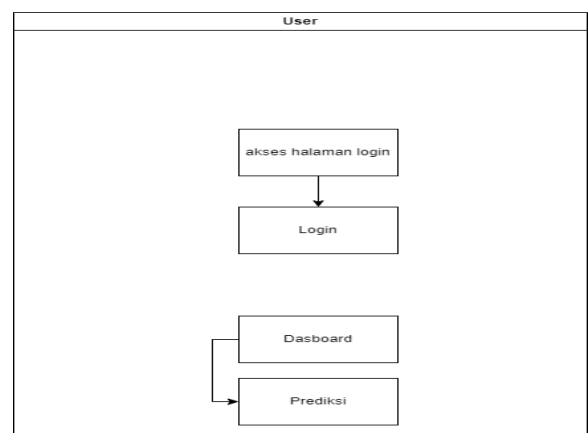
### B. Rancangan Sistem

Rancangan system yang akan dibangun adalah system klasifikasi kepuasan siswa menggunakan *Algoritma C5.0*. system akan dimulai dengan login terlebih dahulu lalu mengisi beberapa survey yang nantinya akan di proses menggunakan *Algoritma C5.0* untuk menghasilkan hasil puas dan tidak puas.

Maka untuk menggambarkan proses tersebut diperoleh *flowchart* sebagai berikut :



Gambar. 2 Flowchart System Admin



Gambar. 3 Flowchart System User

## IV. HASIL DAN PEMBAHASAN

### A. Hasil Penelitian

Hasil penelitian dari “Klasifikasi Tingkat Kepuasan Siswa Terhadap Kinerja Tentor di CV. Ibsi Edutama Menggunakan *Algoritma C5.0*.” Dalam penelitian ini, akan dibangun rules kepuasan siswa menggunakan pohon keputusan. Rules yang dihasilkan akan di implementasikan ke dalam system kepuasan siswa berbasis website.

Penelitian ini dilakukan melalui penyebaran kuisioner kepada siswa CV. Ibsi Edutama. Berikut adalah data yang akan digunakan untuk menganalisis kepuasan menggunakan *Algoritma C5.0*

Sebelum melakukan perhitungan *algoritma C5.0*, data pertama kali dibagi menggunakan split data untuk mencari pembagian yang memiliki tingkat akurasi tertinggi. Dalam hal ini, 80% data digunakan untuk data training dan 20% data digunakan untuk data testing, dari total 100 data hasil kuesioner. Setelah pembagian data selesai, dilakukan perhitungan *algoritma C5.0* secara manual. Berikut dijelaskan tahapan untuk menghitung *entropy*, *information gain*, dan *gain ratio* guna membentuk sebuah pohon keputusan secara manual.

#### 1. Menghitung Entropy masing-masing atribut

$$\text{Entropy Total} = \left( -\frac{62}{80} x \log_2 \left( \frac{62}{80} \right) \right) + \left( -\frac{18}{80} x \log_2 \left( \frac{18}{80} \right) \right) \\ = 0,769192829$$

Entropy A1:

Atribut "SB" =

$$\left( -\frac{28}{29} x \log_2 \left( \frac{28}{29} \right) \right) + \left( -\frac{1}{29} x \log_2 \left( \frac{1}{29} \right) \right) = 0,216396932$$

Atribut "B" =

$$= \left( -\frac{32}{36} x \log_2 \left( \frac{32}{36} \right) \right) + \left( -\frac{4}{36} x \log_2 \left( \frac{4}{36} \right) \right) =$$

0,503258335

Atribut "KB" =

$$\left( -\frac{2}{14} x \log_2 \left( \frac{2}{14} \right) \right) + \left( -\frac{12}{14} x \log_2 \left( \frac{12}{14} \right) \right) = 0,591672779$$

Atribut "TB" =

$$\left( -\frac{0}{1} x \log_2 \left( \frac{0}{1} \right) \right) + \left( -\frac{1}{1} x \log_2 \left( \frac{1}{1} \right) \right) = 0$$

Entropy A2:

Atribut "SB" =

$$= \left( -\frac{22}{22} x \log_2 \left( \frac{22}{22} \right) \right) + \left( -\frac{0}{22} x \log_2 \left( \frac{0}{22} \right) \right) = 0$$

Atribut "B" =

$$\left( -\frac{30}{32} x \log_2 \left( \frac{30}{32} \right) \right) + \left( -\frac{2}{32} x \log_2 \left( \frac{2}{32} \right) \right) = 0,337290067$$

Atribut "KB" =

$$\left( -\frac{8}{20} x \log_2 \left( \frac{8}{20} \right) \right) + \left( -\frac{12}{20} x \log_2 \left( \frac{12}{20} \right) \right) = 0,970950594$$

Atribut "TB" =

$$\left( -\frac{2}{6} x \log_2 \left( \frac{2}{6} \right) \right) + \left( -\frac{4}{6} x \log_2 \left( \frac{4}{6} \right) \right) = 0,918295834$$

Entropy A3:

Atribut "SB" =

$$\left( -\frac{26}{26} x \log_2 \left( \frac{26}{26} \right) \right) + \left( -\frac{0}{26} x \log_2 \left( \frac{0}{26} \right) \right) = 0$$

Atribut "B" =

$$\left( -\frac{32}{32} x \log_2 \left( \frac{32}{32} \right) \right) + \left( -\frac{0}{32} x \log_2 \left( \frac{0}{32} \right) \right) = 0$$

Atribut "KB" =

$$\left( -\frac{4}{17} x \log_2 \left( \frac{4}{17} \right) \right) + \left( -\frac{13}{17} x \log_2 \left( \frac{13}{17} \right) \right) = 0,787126586$$

Atribut "TB" =

$$\left( -\frac{0}{5} x \log_2 \left( \frac{0}{5} \right) \right) + \left( -\frac{5}{5} x \log_2 \left( \frac{5}{5} \right) \right) = 0$$

Entropy A4:

Atribut "SB" =

$$\left( -\frac{23}{23} x \log_2 \left( \frac{23}{23} \right) \right) + \left( -\frac{0}{26} x \log_2 \left( \frac{0}{26} \right) \right) = 0$$

Atribut "B" =

$$\left( -\frac{34}{37} x \log_2 \left( \frac{34}{37} \right) \right) + \left( -\frac{3}{37} x \log_2 \left( \frac{3}{37} \right) \right) = 0,405977038$$

Atribut "KB" =

$$\left( -\frac{5}{20} x \log_2 \left( \frac{5}{20} \right) \right) + \left( -\frac{15}{20} x \log_2 \left( \frac{15}{20} \right) \right) = 0,811278124$$

Atribut "TB" =

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

Entropy A5 :

Atribut "SB" =

$$\left( -\frac{24}{24} x \log_2 \left( \frac{24}{24} \right) \right) + \left( -\frac{0}{24} x \log_2 \left( \frac{0}{24} \right) \right) = 0$$

Atribut "B" =

$$\left( -\frac{36}{37} x \log_2 \left( \frac{36}{37} \right) \right) + \left( -\frac{1}{37} x \log_2 \left( \frac{1}{37} \right) \right) = 0,179256067$$

Atribut "KB" =

$$\left( -\frac{2}{13} x \log_2 \left( \frac{2}{13} \right) \right) + \left( -\frac{11}{13} x \log_2 \left( \frac{11}{13} \right) \right) = 0,619382195$$

Atribut "TB" =

$$\left( -\frac{0}{6} x \log_2 \left( \frac{0}{6} \right) \right) + \left( -\frac{6}{6} x \log_2 \left( \frac{6}{6} \right) \right) = 0$$

Entropy A6:

Atribut "SB" =

$$\left(-\frac{18}{18}x \log_2\left(\frac{18}{18}\right)\right) + \left(-\frac{0}{18}x \log_2\left(\frac{0}{18}\right)\right) = 0$$

Atribut "B" =

$$\left(-\frac{37}{40}x \log_2\left(\frac{37}{40}\right)\right) + \left(-\frac{3}{40}x \log_2\left(\frac{3}{40}\right)\right) = 0,384311544$$

Atribut "KB" =

$$\left(-\frac{7}{16}x \log_2\left(\frac{7}{16}\right)\right) + \left(-\frac{9}{16}x \log_2\left(\frac{9}{16}\right)\right) = 0,619382195$$

Atribut "TB" =

$$\left(-\frac{0}{6}x \log_2\left(\frac{0}{6}\right)\right) + \left(-\frac{6}{6}x \log_2\left(\frac{6}{6}\right)\right) = 0$$

Entropy A7:

Atribut "SB" =

$$\left(-\frac{21}{21}x \log_2\left(\frac{21}{21}\right)\right) + \left(-\frac{0}{21}x \log_2\left(\frac{0}{21}\right)\right) = 0$$

Atribut "B" =

$$\left(-\frac{37}{38}x \log_2\left(\frac{37}{38}\right)\right) + \left(-\frac{1}{38}x \log_2\left(\frac{1}{38}\right)\right) = 0,175565026$$

Atribut "KB" =

$$\left(-\frac{4}{18}x \log_2\left(\frac{4}{18}\right)\right) + \left(-\frac{14}{18}x \log_2\left(\frac{4}{18}\right)\right) = 0,764204507$$

Atribut "TB" =

$$\left(-\frac{0}{3}x \log_2\left(\frac{0}{3}\right)\right) + \left(-\frac{3}{3}x \log_2\left(\frac{3}{3}\right)\right) = 0$$

Entropy A8:

Atribut "SB" =

$$\left(-\frac{17}{17}x \log_2\left(\frac{17}{17}\right)\right) + \left(-\frac{0}{17}x \log_2\left(\frac{0}{17}\right)\right) = 0$$

Atribut "B" =

$$\left(-\frac{38}{41}x \log_2\left(\frac{38}{41}\right)\right) + \left(-\frac{3}{41}x \log_2\left(\frac{3}{41}\right)\right) = 0,37764321$$

Atribut "KB" =

$$\left(-\frac{7}{15}x \log_2\left(\frac{7}{15}\right)\right) + \left(-\frac{8}{15}x \log_2\left(\frac{8}{15}\right)\right) = 0,996791632$$

Atribut "TB" =

$$\left(-\frac{0}{7}x \log_2\left(\frac{0}{7}\right)\right) + \left(-\frac{7}{7}x \log_2\left(\frac{7}{7}\right)\right) = 0$$

## 2. Menghitung Information Gain

**Nilai Information Gain A1**

= 0,769192829 -

$$\left(\left(\frac{1}{80}x 0\right)\right) + \left(\left(\frac{14}{80}x 0,103542736\right) + \left(\frac{36}{80}x 0,226466251\right) + \left(\frac{29}{80}x 0,078443888\right)\right) = 0,360739954$$

**Nilai Information Gain A2**

= 0,769192829 -

$$\left(\left(\frac{6}{80}x 0,918295834\right)\right) + \left(\left(\frac{20}{80}x 0,970950594\right) + \left(\frac{32}{80}x 0,337290067\right) + \left(\frac{22}{80}x 0\right)\right) = 0,447495945$$

**Nilai Information Gain A3**

= 0,769192829 -

$$\left(\left(\frac{5}{80}x 0\right)\right) + \left(\left(\frac{17}{80}x 0,787126586\right) + \left(\frac{32}{80}x 0\right) + \left(\frac{26}{80}x 0\right)\right) = 0,601928429$$

**Nilai Information Gain A4**

= 0,769192829 -

$$\left(\left(\frac{0}{80}x 0\right)\right) + \left(\left(\frac{20}{80}x 0,811278124\right) + \left(\frac{37}{80}x 0,405977038\right) + \left(\frac{23}{80}x 0\right)\right) = 0,378608918$$

**Nilai Information Gain A5**

= 0,769192829 -

$$\left(\left(\frac{16}{80}x 0\right)\right) + \left(\left(\frac{13}{80}x 0,619382195\right) + \left(\frac{37}{80}x 0,179256067\right) + \left(\frac{24}{80}x 0\right)\right) = 0,585637291$$

**Nilai Information Gain A6**

= 0,769192829 -

$$\left(\left(\frac{6}{80}x 0\right)\right) + \left(\left(\frac{16}{80}x 0,988699408\right) + \left(\frac{40}{80}x 0,384311544\right) + \left(\frac{18}{80}x 0\right)\right) = 0,379297175$$

**Nilai Information Gain A7**



$$= 0,769192829 -$$

$$\left( \left( \frac{3}{80} \times 0 \right) \right) + \left( \left( \frac{18}{80} \right) \times 0,764204507 \right) + \left( \left( \frac{38}{80} \right) \times 0,175565026 \right) + \left( \left( \frac{21}{80} \right) \times 0 \right)$$

$$= 0,513853428$$

Nilai Information Gain A8

$$= 0,769192829 -$$

$$\left( \left( \frac{7}{80} \times 0 \right) \right) + \left( \left( \frac{15}{80} \right) \times 0,996791632 \right) + \left( \left( \frac{41}{80} \right) \times 0,377646321 \right) + \left( \left( \frac{17}{80} \right) \times 0 \right)$$

$$= 0,388750658$$

### 3. Menghitung Nilai Gain Ratio

Nilai Gain Ratio A1

$$= \left( \frac{1,311328046}{0,360739954} \right) = 0,275095126$$

Nilai Gain Ratio A2

$$= \left( \frac{2,226536495}{0,447495945} \right) = 4,975545633$$

Nilai Gain Ratio A3

$$= \left( \frac{0,787126586}{0,601928429} \right) = 1,307674713$$

Nilai Gain Ratio A4

$$= \left( \frac{1,217255163}{0,378608918} \right) = 3,215072615$$

Nilai Gain Ratio A5

$$= \left( \frac{0,798638262}{0,585637291} \right) = 1,363708004$$

Nilai Gain Ratio A6

$$= \left( \frac{1,373010952}{0,379297175} \right) = 3,619881828$$

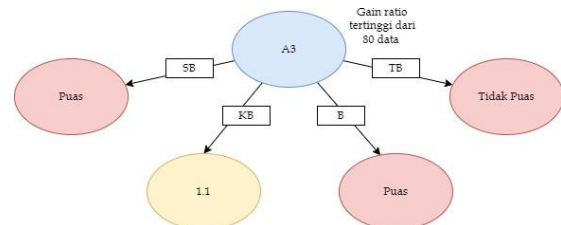
Nilai Gain Ratio A7

$$= \left( \frac{0,939769532}{0,513853428} \right) = 1,828866913$$

Nilai Gain Ratio A8

$$= \left( \frac{1,374437953}{0,388750658} \right) = 3,535535726$$

Dari perhitungan di atas, atribut A3 memiliki nilai *gain ratio* tertinggi dibandingkan dengan nilai *gain ratio* atribut A1, A2, A4, A5, A6, A7, dan A8. Oleh karena itu, atribut A3 dipilih sebagai akar dari *node 1*. Berikut adalah bentuk pohon keputusan untuk *node 1*:



Gambar. 4 Pohon Keputusan Node 1

Tahap berikutnya adalah mencari cabang dari setiap nilai (value) dengan menggunakan metode yang sama, yakni dengan menghitung *entropy*, *information gain*, dan *gain ratio*.

### B. Implementasi Decision Tree Pada Python

Berikut adalah langkah-langkah dalam membuat sebuah *decision tree* menggunakan *Python*:

#### 1. Langkah pertama yaitu import libraries python

```
# Import Library
import pandas as pd
```

Kode di atas menunjukkan beberapa pustaka yang diimport adalah:

- a. **Pandas** : Untuk memanipulasi, analisis dan untuk mengimpor library Pandas ke dalam program Python.
- b. **As pd** : Ini adalah bagian dari sintaksis `import` yang memberikan alias atau nama alternatif untuk modul yang diimpor. Dalam kasus ini, `pd` adalah alias yang diberikan untuk modul Pandas. Memberikan alias seperti ini memudahkan untuk memanggil fungsi atau kelas dari modul tersebut tanpa harus menuliskan nama lengkapnya.

#### 2. Setelah berhasil mengimport libraries, selanjutnya yaitu mengimport data dalam bentuk CSV seperti gambar dibawah ini

```
#Membaca dataset
data =
pd.read_csv("data_siswa_full.csv", sep=";
```

#### 3. Langkah selanjutnya yaitu melakukan cleaning data yaitu membuang data yang tidak diperlukan dalam perhitungan ini. Kemudian, melakukan pembagian data yaitu pembagian data training dan testing. Dalam perhitungan ini, data training diambil sebesar 80% dan data testing sebesar 20%.

```
#Membagi data menjadi data training dan
testing
data_clean = data.drop(columns=["Nama
Siswa"])
ratio_train = 0.8
split =
round(len(data_clean)*ratio_train)
data_train = data_clean[:split]
data_test = data_clean[split:]
```

#### 4. Kemudian melakukan perhitungan dengan *Algoritma C5.0* dengan pemrograman python.

```
# Menghitung entropi dari dataset yang
diberikan
def calculate_entropy(self, data,
weights):

# Menghitung jumlah masing-masing label
kelas
class_counts = {}

# Menghitung total bobot data
total_weight = 0.0

for i, record in enumerate(data):
# Mengambil label kelas dari setiap data
label = record[-1]

# Mengambil bobot dari setiap data
weight = weights[i]

entropy = 0.0

# Menghitung probabilitas masing-masing
label kelas
for count in class_counts.values():
probability = count /
total_weight

# Menghitung kontribusi entropi dari
masing-masing label
entropy -= probability *
math.log2(probability)

return entropy
```

Fungsi ini digunakan untuk menghitung *entropy* dari setiap atribut kemudian dilanjutkan dengan perhitungan *gain ratio* yang tertinggi yang digunakan untuk menentukan label

```
# Memilih atribut terbaik untuk membagi
dataset menggunakan algoritma C5.0
def select_best_attribute_c50(self,
data, attributes, weights):
total_entropy =
self.calculate_entropy(data, weights)
best_attribute = None
best_gain_ratio = 0.0
split_info = 0.0
attribute_entropy = 0.0
summary_entropy = 0.0
for value in attribute_values:
subset, subset_weights =
self.split_data(data, attribute_index,
value, weights)
subset_entropy =
self.calculate_entropy(subset,
subset_weights)
subset_probability =
sum(subset_weights) / sum(weights)
attribute_entropy +=
subset_probability * subset_entropy
summary_entropy += subset_entropy
split_info -= subset_probability *
math.log2(subset_probability)
```

```
gain = total_entropy -
attribute_entropy

if summary_entropy != 0.0:
gain_ratio = gain /
summary_entropy
else:
gain_ratio = 0.0

if gain_ratio > best_gain_ratio:
best_gain_ratio = gain_ratio
best_attribute = attribute_index
```

#### 5. Visualisasi pohon keputusan siswa menggunakan *Algoritma C5.0*

```
# Membuat Pohon Keputusan
dt = DecisionTreeClassifier()

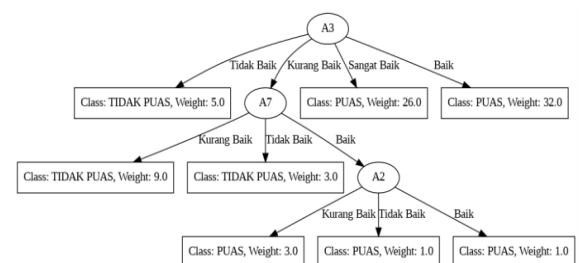
*Melatih Pohon Keputusan
dt.train(data_train)

from IPython.display import image

#Menghasilkan diagram pohon
keputusan
Image(dt.generate_tree_diagram("Dia
gram"))
```

Fungsi tersebut membuat sebuah objek *Decision tree classifier*, dan digunakan untuk melatih model pohon keputusan dengan data pelatihan, dan menampilkan diagram pohon keputusan.

Dibawah ini adalah bentuk pohon keputusan dari kepuasan siswa terhadap kinerja tutor pengajar di CV Ibsi Edutama menggunakan python:



Gambar. 5 Decision tree python

#### 6. Memanggil Model Menggunakan Library Pickle

```
# Load the model
Model_classifier =

pickle.load(open('model.pkl', 'rb'))
```



Pada bagian ini, model klasifikasi yang telah dilatih sebelumnya dimuat ke dalam variable 'model\_classifier' menggunakan modul 'pickle'.  
Pickle.load (open('model.pkl', 'rb')) digunakan untuk membaca file yang disimpan sebelumnya ( dalam format pickle) yang berisi model klasifikasi yang telah dilatih.

## 7. Tampilan Export Model

```
From model import
DecisionTreeClassifier
Import pickle
From pandas import read_csv

data =
read_csv('dataset_siswa.csv').drop(col
umns=['Nama Siswa'])
model_classifier =
DecisionTreeClassifier(max_depth=3)
model_classifier.train(data)

#Mengubah object model classifier
menjadi model.pkl wb = untuk save
jadi file
Pickle.dump(model_classifier,open('mo
del.pkl', 'wb'))
```

## 8. Membuat Tabel Confussion Matrix untuk uji akurasi model pohon keputusan yang dibuat

```
pred =
dt.predict_all(x_Test.values.tolist())

cf_matrix =
confusion_matrix(y_Test,pred)
group_names = ['True Puas', 'False Tdk
Puas','False Puas', 'True Tdk Puas']
```

Hasil dari tabel confusion matrix menunjukkan terdapat 14 kondisi true positive, 4 kondisi true negative, 2 kondisi false positive, dan 0 kondisi false negative. Setelah mendapatkan tabel confusion matrix, langkah selanjutnya adalah melakukan perhitungan akurasi.

## 9. Menghitung Akurasi Algoritma C5.0

```
tp = cf_matrix[0][0]
fp = cf_matrix[0][1]
tp = cf_matrix[1][0]
fp = cf_matrix[1][1]

accuracy = (tp+tn) / (tp+tn+fp+fn)
*100
print (accuracy,"%")
```

## C. Implementasi Sistem Kepuasan Siswa

### 1. User Interface

Setelah melakukan serangkaian rancangan penelitian dan pengumpulan data, selanjutnya dilakukan Implementasi metode *algoritma C5.0* kedalam system. Hasil implementasinya yaitu berupa aplikasi berbasis website sederhana yang mudah digunakan untuk user. Berikut merupakan tampilan user interface aplikasi.

#### a) Halaman Utama



Gambar. 6 Tampilan Utama

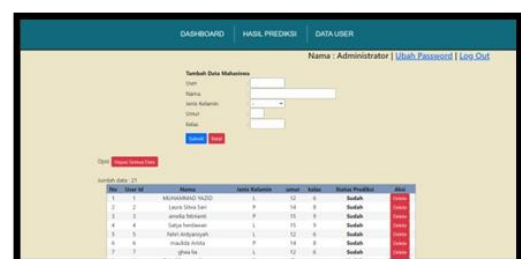
#### b) Halaman Admin



Gambar. 7 Tampilan Halaman Dashboard Admin



Gambar. 8 Tampilan Halaman Hasil Prediksi

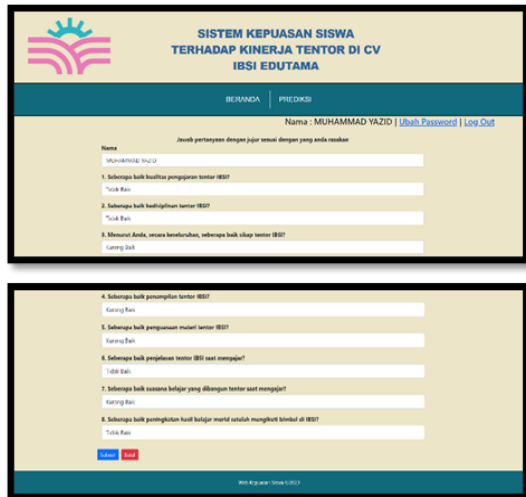


Gambar.9 Tampilan Halaman Data User

### c) Halaman User



Gambar.10 Tampilan Halaman Utama User



Gambar. 11 Tampilan Halaman Survey Prediksi



Gambar. 12 Tampilan Halaman Hasil Prediksi

### 2. Pengujian Sistem Admin

Pengujian Sistem Admin dilakukan menggunakan metode Black Box Testing adalah metode pengujian perangkat lunak. Tujuannya adalah untuk menguji perangkat lunak secara fungsional, sehingga dapat mengidentifikasi kesalahan pada berbagai aspek seperti kategori fungsi, antarmuka, struktur data, dan kinerja.[8]

Berikut merupakan table Hasil Pengujian Sistem Admin

TABEL IV  
HASIL PENGUJIAN SISTEM ADMIN INTERFACE

| Pengujian            | Inputan   | Diharapkan  | Hasil  |
|----------------------|---|---|--------|
| Field Sign In        | Memasukkan username dan pass-word                                     | Admin sukses masuk ke halaman dashboard admin                           | Sesuai |
| Field Hasil Prediksi | Admin import excel, delete data siswa dan download laporan            | Admin sukses Import excel, Delete data Siswa dan download hasil laporan | Sesuai |
| Field Data User      | Admin menambahkan data siswa Untuk login User serta delete data siswa | Admin sukses menambahkan data siswa dan delete data siswa               | Sesuai |
| Field Sign Out       | Admin Keluar dari halaman admin                                       | Admin sukses keluar dari halaman admin                                  | Sesuai |

### 3. Pengujian Sistem User

Testing yang dilakukan pada sistem Admin menggunakan metode Black Box Testing. Tujuannya adalah untuk menguji perangkat lunak secara fungsional, sehingga dapat mengidentifikasi kesalahan dalam kategori fungsi, antarmuka, struktur data, dan kinerja..[8]

## V. KESIMPULAN

### A. Kesimpulan

Berdasarkan hasil dari penelitian yang telah dilakukan, dapat disimpulkan bahwa:

1. Hasil penerapan klasifikasi kepuasan siswa ini berbentuk website yang mudah digunakan agar proses survey kepuasan lebih efektif karena tidak perlu lagi menyebarkan selebaran kertas door to door.
2. Atribut yang memiliki pengaruh dominan terhadap kinerja tensor di CV. Ibsi Edutama ada 3 yaitu Sikap tensor baik/tidak saat mengajar, kemudian suasana belajar yang diciptakan tensor apakah membuat siswa nyaman saat belajar, dan kedisiplinan tensor saat mengajar apakah datang tepat waktu ketika akan mengajar.
3. Algoritma C5.0 menunjukkan performa yang baik dalam studi kasus ini dengan tingkat akurasi mencapai 90%. Hal ini menyiratkan bahwa

metode ini adalah pilihan yang tepat untuk menghitung akurasi.

#### B. Saran

Penelitian ini jauh dari kesempurnaan, oleh karena itu, penulis memberikan beberapa saran untuk pengembangan lebih lanjut terhadap penelitian. Berikut adalah beberapa saran untuk mengembangkan penelitian ini:

1. Mengembangkan website dengan beberapa fitur yang lebih banyak dan relevan.
2. Untuk meningkatkan tingkat detail dalam proses klasifikasi, dapat dipertimbangkan untuk menambahkan atribut penelitian tambahan.
3. Menguji *Algoritma C5.0* dengan data yang lebih banyak

#### REFERENSI

- [1] W. D. Kusuma, Suhada, dan Saifullah, "Klasifikasi Kepuasan Siswa Terhadap Kinerja Guru SMK Satrya Budi 2 Perdagangan Menggunakan Algoritma C4.5," *TIN Terap. Inform. Nusan.*, vol. 2, no. 7, hal. 460–465, 2021, [Daring]. Tersedia pada: <https://ejurnal.seminar-id.com/index.php/tin>
- [2] Heliyanti Susana, "Penerapan Model Klasifikasi Metode Naive Bayes Terhadap Penggunaan Akses Internet," *J. Ris. Sist. Inf. dan Teknol. Inf.*, vol. 4, no. 1, hal. 1–8, 2022, doi: 10.52005/jursistekni.v4i1.96.
- [3] K. Aidi Saputra, J. Tata Hardinata, M. Ridwan Lubis, S. Retno Andani, dan I. Syahputra Saragih, "Klasifikasi Algoritma C4.5 Dalam Penerapan Tingkat Kepuasan Siswa Terhadap Media Pembelajaran Online," *Media Online*, vol. 1, no. 3, hal. 113–118, 2020, [Daring]. Tersedia pada: <https://djournals.com/klik>
- [4] R. Pandya, "C5. 0 Algorithm to Improved Decision Tree with Feature Selection and Reduced Error Pruning," *Int. J. Comput. Appl.*, vol. 117, no. 16, hal. 18–21, 2015, doi: 10.5120/20639-3318.
- [5] V. S. Ginting, K. Kusrini, dan E. Taufiq, "Implementasi Algoritma C4.5 untuk Memprediksi Keterlambatan Pembayaran Sumbangan Pembangunan Pendidikan Sekolah Menggunakan Python," *Inspir. J. Teknol. Inf. dan Komun.*, vol. 10, no. 1, hal. 36–44, 2020, doi: 10.35585/inspir.v10i1.2535.
- [6] I. W. Saputro dan B. W. Sari, "Uji Performa Algoritma Naive Bayes untuk Prediksi Masa Studi Mahasiswa," *Creat. Inf. Technol. J.*, vol. 6, no. 1, hal. 1, 2020, doi: 10.24076/citec.2019v6i1.178.
- [7] Puji Purnomo, "PENGEMBANGAN TES HASIL BELAJAR MATEMATIKA MATERI MENYELESAIKAN MASALAH YANG BERKAITAN DENGAN WAKTU, JARAK DAN KECEPATAN UNTUK SISWA KELAS V," 2016.
- [8] Y. D. Wijaya dan M. W. Astuti, "Pengujian Blackbox Sistem Informasi Penilaian Kinerja Karyawan Pt Inka (Persero) Berbasis Equivalence Partitions," *J. Digit. Teknol. Inf.*, vol. 4, no. 1, hal. 22, 2021, doi: 10.32502/digital.v4i1.3163.