

Implementasi Algoritma C4.5 Dalam Mengukur Tingkat Kepuasan Karyawan Terhadap Sistem Absensi Berbasis Pemindaian Wajah

Febriana Manalu¹, Yeni Rokhayati²

Teknik Informatika, Politeknik Negeri Batam
Teknologi Rekayasa Multimedia¹, Politeknik Negeri Batam²
febrimanalu39@gmail.com¹, yeni@polibatam.ac.id²

Article Info

Article history:

Received 20 January 2025

Revised 23 January 2025

Accepted 27 January 2025

Keyword:

Kepuasan Karyawan, Algoritma C4.5,
Pohon Keputusan.

ABSTRACT

PT Wasco Engineering Indonesia implemented a modern attendance system using badge scanning, but it is considered ineffective by management in accurately recording employee attendance. Therefore, the management decided to improve the attendance system by adopting face scanning technology and conducting employee satisfaction surveys to evaluate the effectiveness of the new system to measure the effectiveness of the system. The analysis of employee satisfaction results was carried out using the C4.5 Algorithm which produces a decision tree with easy-to-understand visualizations. The results of the analysis using the C4.5 Algorithm produced a decision tree model with an accuracy and testing value of 90.54%. The results of this analysis will be used by management for evaluation media and improving the quality of the attendance system based on facial recognition technology to improve the company operations.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Penggunaan sistem absensi dikategorikan menjadi dua jenis yaitu tradisional dan modern. Sistem absensi tradisional dengan cara menulis manual untuk mengolah kehadiran, dimana hal tersebut memakan waktu dan biaya yang banyak. Sedangkan, sistem absensi modern tidak menggunakan sistem menulis manual untuk mengelolanya melainkan dengan memanfaatkan teknologi yang sudah menjadi penggunaan sehari – hari pada abad ke-20 ini, seperti menggunakan komputer, *mobile*, *fingerprint*, *scan badge*, dan pemindaian wajah. PT Wasco Engineering Indonesia sudah menggunakan sistem absensi modern dengan menggunakan *scan badge*. Namun, evaluasi yang dilakukan oleh manajemen menyimpulkan bahwa sistem ini masih memiliki kekurangan dan tidak efektif dalam mencatat kehadiran karyawan. Penilaian ini berdasarkan pada temuan banyak karyawan yang memanfaatkan celah dalam sistem dengan cara menitipkan badge mereka kepada rekan kerja untuk melakukan absensi tanpa kehadiran fisik, kasus banyaknya karyawan yang melakukan double break serta kasus banyaknya absensi yang tidak tercatat oleh sistem. Dalam mengatasi permasalahan tersebut, manajemen telah memutuskan untuk meningkatkan sistem absensi dengan menggunakan teknologi pemindaian wajah.

Setelah mengimplementasikan sistem absensi menggunakan pemindaian wajah, manajemen memprioritaskan evaluasi tingkat kepuasan karyawan terhadap sistem baru untuk memastikan keberhasilan implementasinya. Sebagai bagian dari upaya untuk mengukur tingkat kepuasan karyawan terhadap sistem absensi menggunakan pemindaian wajah di PT Wasco Engineering Indonesia, manajemen telah merancang dan menyebarkan form kuesioner kepada karyawan. Dengan melibatkan sebagian karyawan dalam survei kepuasan akan menjadi landasan penting bagi manajemen untuk mengevaluasi kinerja sistem, mengidentifikasi area-area perbaikan yang diperlukan serta merancang strategi untuk meningkatkan efisiensi operasional perusahaan.

Algoritma C4.5 dikenal dengan kemampuan menghasilkan model pohon keputusan yang mudah diinterpretasikan. yang mudah diinterpretasikan. Algoritma ini dapat menangani atribut numerik maupun kategorikal,

mengatasi data yang tidak lengkap, dan menggunakan pendekatan gain ratio untuk mengidentifikasi faktor-faktor utama yang memengaruhi tingkat kepuasan karyawan. Dengan algoritma C4.5, faktor – faktor penentu dapat diurutkan berdasarkan pengaruhnya dan akan memberikan gambaran jelas kepada manajemen untuk mengevaluasi sistem absensi. Hasil dari implementasi algoritma C4.5 bertujuan untuk membantu manajemen memahami tingkat kepuasan karyawan secara mendalam, mengidentifikasi tingkat kepuasan karyawan dan diharapkan menjadi dasar bagi manajemen untuk mengevaluasi dan meningkatkan kualitas sistem absensi berbasis pemindaian wajah. Berdasarkan permasalahan diatas manajemen membutuhkan hasil analisis untuk meningkatkan kualitas sistem absensi berbasis pemindaian wajah. Maka muncul ide untuk implementasi algoritma C4.5 dalam mengukur tingkat kepuasan karyawan terhadap sistem absensi berbasis pemindaian wajah.

II. LANDASAN TEORI

A. Tinjauan Pustaka

Implementasi data mining untuk mengukur tingkat kepuasan mahasiswa terhadap pelayanan akademik dengan algoritma C4.5 sudah dilakukan oleh Tulus dan Ariesanto [13]. Proses pengolahan data akademik diawali dengan pengumpulan data kuesioner akademik mahasiswa, kemudian data tersebut dikategorikan menggunakan algoritma C4.5 dan menghasilkan pohon keputusan. Hasil dari pohon model keputusan terdapat rule yang digunakan untuk memprediksi kepuasan mahasiswa terhadap layanan akademik. Kemudian dilanjutkan dengan mengukur akurasi dalam pengolahan data pada tools RapidMiner. Maka setelah mendapatkan hasil tingkat kepuasan mahasiswa diharapkan pihak kampus dapat meningkatkan pelayanan akademik dengan meninjau aspek yang paling dominan.

Implementasi data mining untuk menganalisis kepuasan pegawai terhadap pelayanan bidang SDM dengan algoritma C4.5 sudah dilakukan oleh Putrama dan Zakarias [3]. Proses pengolahan data kepuasan pegawai diawali dengan pengumpulan hasil angket karyawan, selanjutnya data diolah menggunakan algoritma C4.5 dan menghasilkan pohon keputusan. Maka setelah mendapatkan hasil klasifikasi data kepuasan pegawai diharapkan

Polsek Siantar Barat dapat menindak kejahatan sehingga meningkatkan pengamanan dalam menangani kejahatan.

Implementasi data mining untuk mengetahui tingkat kepuasan karyawan terhadap layanan akademik dan kemahasiswaan menggunakan algoritma C4.5 sudah dilakukan oleh Reni, Dadang dan Elviza [7]. Proses pengolahan data kuesioner mahasiswa diawali dengan pengumpulan data kuesioner melalui google form yang diisi oleh mahasiswa di lingkungan Universitas Prof Dr. Hazairin, S.H. Selanjutnya data diolah menggunakan algoritma C4.5 untuk membangun pohon keputusan, hasil tersebut akan diaplikasikan menggunakan tools Weka untuk melihat keakuratan yang diperoleh. Maka setelah mendapatkan hasil kepuasan mahasiswa diharapkan pihak perguruan tinggi dapat meningkatkan kualitas pelayanan pada bidang akademik dan kemahasiswaan.

Table 1 Perbandingan dengan Penelitian Sebelumnya

Perbandingan	Penelitian 1	Penelitian 2	Penelitian 3	Penelitian Sekarang
Tujuan Penelitian	Mengukur tingkat kepuasan mahasiswa terhadap pelayanan akademik	Menganalisis kepuasan pegawai terhadap pelayanan bidang SDM	Mengetahui tingkat kepuasan mahasiswa terhadap layanan akademik dan kemahasiswaan	Mengukur tingkat kepuasan karyawan terhadap sistem pemindaian wajah
Studi Kasus	Program Studi Informatika Fakultas Teknik dan Ilmu Komputer Universitas Pancasakti Tegal	STIKOM Tunas Bangsa Pematangsiantar	Biro Administrasi Akademik dan Kemahasiswaan Universitas Prof Dr Hazairin, SH	PT Wasco Engineering Indonesia
Algoritma	C4.5	C4.5	C4.5	C4.5
Software Pengembangan	RapidMiner	RapidMiner	Weka	Python

B. Dasar Teori

B.A Data Mining

Data mining adalah proses semi – otomatis yang menggunakan statistik, matematika, kecerdasan buatan dan teknik pembelajaran mesin untuk mengekstraksi dan mengidentifikasi informasi dan pengetahuan berguna yang terkait dengan berbagai basis data [15]. Data mining sering juga disebut sebagai knowledge discovery in database (KDD). KDD merupakan suatu proses menentukan informasi yang berguna dalam data dalam bentuk basis data yang berukuran besar [5]. Berikut ini merupakan beberapa tahapan dalam data mining sebagai berikut :

1. Pembersihan data (data cleaning) adalah proses dimana data – data yang tidak konsisten dan noise dibuang dari koleksi data.
2. Integrasi data (data integration) adalah proses penyeleksian data dari data sampel karena tidak semua sampel akan digunakan
3. Transformasi data (data transformation) adalah proses penentuan data yang sudah dikumpulkan sehingga data dapat diubah kedalam bentuk mining.
4. Evaluasi pola (pattern evaluation) adalah hasil dari teknik data mining berupa pola prediksi dievaluasi untuk menilai apakah hipotesis yang diinginkan tercapai.
5. Persentase pengetahuan (knowledge presentation) adalah proses tahapan terakhir dengan teknik visualisasi bertujuan untuk membantu user agar mengerti dan mengiterpresentasikan hasil dari pengolahan data.

B.B Klasifikasi

Klasifikasi adalah salah satu metode dari data mining yang berfungsi untuk mengelompokkan data berdasarkan keterkaitan terhadap data sampel. Tujuan utama dari klasifikasi adalah membangun model yang mampu membedakan atribut ke dalam kategori atau kelas

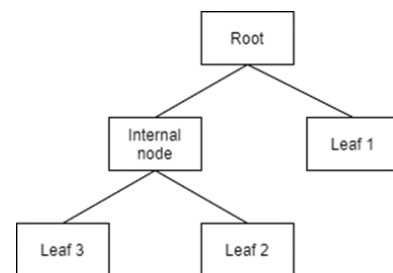
yang sesuai. Model tersebut kemudian digunakan untuk mengklasifikasikan atribut yang belum diketahui kelasnya. Metode klasifikasi dibagi menjadi beberapa metode yaitu sebagai berikut :

Table 2 Metode Klasifikasi

Metode	Pengertian	Kelebihan	Kekurangan
Pohon Keputusan (Decision Tree)	Metode struktur yang mengubah data menjadi struktur pohon berdasarkan aturan keputusan dari fitur-fitur dalam data	- Mudah diinterpretasi dan divisualisasikan	- Cenderung overfitting jika pohonnya terlalu dalam
		- Tidak memerlukan normalisasi data	- Kurang akurat untuk data kompleks
		- Efisien untuk data dengan variabel yang kecil	- Sensitif terhadap perubahan data
Rule-Based Classification	Metode yang menggunakan aturan "if-then" untuk klasifikasi data berdasarkan pola atau hubungan antar fitur dalam data.	- Mudah diinterpretasi	- Memerlukan aturan yang baik
		- Fleksibel dan mudah dimodifikasi	- Kurang efektif pada data besar atau kompleks
		- Cepat dalam pengambilan keputusan	- Tidak cocok untuk data tidak terstruktur
Neural Network	Metode yang meniru cara kerja otak dengan menghubungkan lapisan-lapisan.	- Kemampuan tinggi pada data besar dan kompleks	- Sulit diinterpretasi (kotak hitam)
		- Sangat baik untuk pengenalan pola	- Membutuhkan data besar
		- Dapat mempelajari hubungan non-linear antar fitur	- Memerlukan daya komputasi tinggi
Support Vector Machine (SVM)	Metode yang mengubah data menjadi dimensi lebih tinggi melalui pemetaan non-linear.	- Cocok untuk data dengan dimensi tinggi	- Kurang efisien pada data besar
		- Efektif untuk klasifikasi biner	- Sulit diinterpretasi
		- Robust terhadap overfitting pada margin yang besar	- Sensitif terhadap pemilihan parameter kernel
Naïve Bayes	Metode yang berakar pada teorema Bayes dengan asumsi independensi antar fitur	- Cepat dan efisien, terutama pada data besar	- Asumsi independensi antar fitur sering tidak realistis
		- Efektif pada klasifikasi teks	- Kurang akurat untuk data kompleks
		- Tidak memerlukan data yang banyak	- Sulit menangani fitur numerik tanpa transformasi tambahan

B.C Pohon Keputusan (Decision Tree)

Pohon keputusan merupakan salah satu metode sederhana dari metode klasifikasi pada data mining yang dapat mengubah sekumpulan data besar menjadi pohon keputusan yang merepresentasikan aturan yang mudah dipahami. Dalam pohon keputusan, setiap node merepresentasikan atribut yang telah diuji dan node daun (leaf) merepresentasikan kelompok kelas tertentu [5]. Node akar (root) merupakan level node teratas dari sebuah pohon keputusan dan memiliki pengaruh besar pada suatu kelas tertentu. Konsep dari pohon keputusan adalah mengubah data menjadi struktur pohon keputusan, kemudian mengonversinya menjadi aturan rule menjadi lebih sederhana dan mudah diinterpretasikan. Pohon keputusan dapat dilihat seperti Gambar 1 dibawah ini :



Gambar 1 Pohon Keputusan

Terdapat berbagai jenis algoritma yang dapat mendukung pembuatan pohon keputusan, salah satunya adalah algoritma ID3 yang merupakan algoritma dasar dalam membentuk pohon keputusan. Algoritma C4.5 merupakan pengembangan dari konsep dasar algoritma ID3, dimana pembentukan pohon dilakukan dengan menghitung nilai entropi dan information gain, serta nilai gain ratio sebagai patokan dalam membentuk akar dan simpul daun pada decision tree, namun keduanya menghasilkan tingkat akurasi yang cukup tinggi[14]. Algoritma CART merupakan algoritma yang menggunakan gini index untuk mengukur klasifikasinya. Berikut ini merupakan penjelasan dari algoritma tersebut sebagai berikut :

Table 3 Algoritma Pendukung Pohon Keputusan

Algoritma	Pengertian	Kelebihan	Kekurangan
ID3	Algoritma pohon keputusan yang menggunakan information gain untuk memilih atribut terbaik pada setiap node.	- Sederhana dan mudah dipahami.	- Rentan terhadap overfitting.
		- Efisien untuk dataset kecil dengan atribut kategori.	- Tidak menangani missing values.
C4.5	Pengembangan dari ID3 yang menggunakan gain ratio sebagai kriteria pemilihan atribut dan mendukung atribut kontinu.	- Mendukung atribut kontinu dan kategori.	- Lebih kompleks dibanding ID3, sehingga membutuhkan waktu komputasi lebih lama.
		- Dapat menangani missing values.	- Sensitif terhadap dataset kecil.
CART	Algoritma pohon keputusan yang menggunakan Gini Index untuk klasifikasi atau Mean Squared Error untuk regresi.	- Mendukung klasifikasi dan regresi.	- Hanya menghasilkan pohon biner.
		- Membentuk pohon biner yang sederhana dan mudah diimplementasikan.	- Rentan terhadap overfitting tanpa pruning.

B.D Algoritma C4.5

Algoritma C4.5 adalah metode pengembangan dari algoritma ID3 yang di ciptakan oleh J. Rose Quinlan. Algoritma C4.5 merupakan salah satu algoritma yang digunakan untuk membentuk pohon keputusan menggunakan gain ratio sebagai kriteria pemilihan atribut. Sedangkan pohon keputusan merupakan suatu cara untuk memprediksi atau mengelompokkan yang sangat kuat. Secara umum algoritma C4.5 digunakan untuk membangun pohon keputusan (Bramer & Max, 2007) adalah sebagai berikut:

1. Memilih atribut sebagai akar
2. Membuat cabang untuk masing – masing nilai
3. Membagi kasus dalam masing – masing cabang
4. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama

Untuk memilih atribut sebagai akar berdasarkan pada nilai gain tertinggi dari atribut - atribut yang ada. Pembentukan algoritma C4.5 terbagi menjadi 2 jenis yaitu entropy dan gain dengan rumus mencari nilai gain dan nilai entropy (Yuli, 2018) sebagai berikut :

1. Entropy

Entropy adalah ukuran ketidakpastian, keacakan atau ketidakteraturan dalam sekumpulan data. Dengan rumus Entropy sebagai berikut :

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

Keterangan :

- S = Himpunan data
- n = jumlah kelas
- pi = Proporsi dari data yang termasuk dalam kelas ke-i

2. Gain

Gain merupakan ukuran seberapa pengurangan ketidakpastian entropy setelah data dipartisi berdasarkan suatu atribut. Dengan rumus Gain sebagai berikut :

$$Gain(S, A) = Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} \cdot Entropy(S_v)$$

Keterangan :

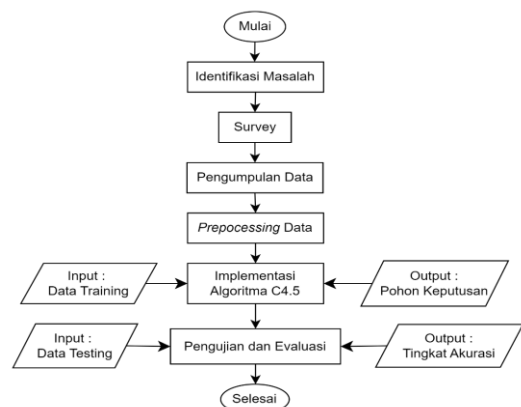
- S = Himpunan data
- A = Atribut yang dipilih
- n = Jumlah kelas
- |S_v| = Jumlah data pada subset S_v yang dihasilkan dengan mempartisi S berdasarkan nilai v dari atribut A
- |S| = Total jumlah data

C. Sistem Absensi Pemindaian Wajah di PT Wasco Engineering Indonesia

PT Wasco Engineering Indonesia (sebelumnya dikenal sebagai Wah Seong Corporation Berhad) didirikan pada tahun 1999 merupakan sebuah grup energi terintegrasi terkemuka global yang menyediakan pelapisan pipa, layanan teknik & fabrikasi, pembuatan pipa dan lainnya untuk sektor minyak dan gas. Wasco telah memanfaatkan teknologi pemindaian wajah sebagai system absensi untuk meningkatkan efisiensi dan akurasi dalam pencatatan kehadiran karyawan. Dengan teknologi pemindaian wajah, system mampu mengenali setiap karyawan secara unik dan dapat mencegah kecurangan saat proses absensi. Selain itu dengan implementasi sistem absensi ini, proses absensi menjadi lebih cepat dan nyaman, karena karyawan hanya perlu berdiri di depan kamera untuk dideteksi. Penggunaan sistem ini juga mendukung upaya perusahaan dalam mengelola data kehadiran dengan lebih terstruktur dan mudah diakses untuk meningkatkan produktivitas dan efisiensi operasional perusahaan.

III. METODE PENELITIAN

Penelitian ini diawali dengan identifikasi masalah yang dilakukan dengan survey pada PT Wasco Engineering Indonesia. Setelah melakukan survey dilanjutkan dengan mengumpulkan data kuesioner karyawan terkait kepuasan karyawan terhadap sistem absensi berbasis pemindaian wajah. Sebelum data diolah, data akan diperiksa terlebih dahulu agar tidak noise dan tidak konsisten untuk mempermudah proses pengolahan data. Kemudian setelah data diolah dan mendapatkan format data yang sesuai untuk di analisis, dilanjutkan dengan klasifikasi menggunakan algoritma C4.5 yang disebut juga pohon keputusan (*decision tree*). Setelah mendapatkan hasil kepuasan karyawan terhadap sistem absensi berbasis sistem pemindaian wajah terdapat pola yang didapatkan berdasarkan data yang ada. Untuk penjelasan lebih lanjut mengenai metode penelitian yang digunakan dapat dilihat pada Gambar 2 dibawah ini :



Gambar 2 Metode Penelitian

Penjelasan rinci pada setiap tahapan di metode penelitian akan dibahas sebagai berikut :

A. Tahap Pengumpulan Data

Penelitian ini memperoleh data dari *Google Form* yang disebar ke karyawan PT Wasco Engineering Indonesia mengenai kualitas sistem absensi berbasis sistem pemindaian wajah yang dilakukan oleh HR pada periode Oktober 2024 sebanyak 100 sampel. Berikut ini merupakan form kuesioner yang disebar ke karyawan, dapat dilihat pada Tabel 4 dibawah ini :

Table 4 Form Kuesioner Karyawan

No	Pertanyaan	Penilaian				
		Sangat Tidak Setuju	Tidak Setuju	Netral	Setuju	Sangat Setuju
Akurasi						
1	Sistem absensi berbasis pemindaian wajah dapat mengenali wajah dengan waktu <10s					
2	Dengan menggunakan sistem absensi berbasis pemindaian wajah akurasi kehadiran karyawan meningkat					
3	Sistem absensi berbasis pemindaian wajah dapat rekam kehadiran dengan akurat					
4	Sistem absensi berbasis pemindaian wajah mampu mengenali jika ada perubahan pencahayaan maupun perubahan fisik pada wajah					
5	Sistem absensi berbasis pemindaian wajah lebih akurat dalam rekam kehadiran dibandingkan metode absensi lainnya					
Fungsionalitas						
6	Sistem absensi berbasis pemindaian wajah dapat berfungsi dengan baik					
7	Sistem absensi berbasis pemindaian wajah memudahkan karyawan untuk mencatat kehadiran					
8	Sistem absensi berbasis pemindaian wajah dapat mengenali wajah dengan akurat					
9	Sistem absensi berbasis pemindaian wajah meningkatkan efisiensi dalam mencatat kehadiran karyawan					
10	Sistem absensi berbasis pemindaian wajah mudah digunakan					
Manfaat						
11	Anda tidak mengalami masalah saat menggunakan sistem absensi berbasis pemindaian wajah					
12	Sistem absensi berbasis pemindaian wajah lebih efektif dalam mencatat kehadiran dibandingkan tapping badge					
13	Sistem absensi berbasis pemindaian wajah memudahkan HR dalam mengelola absensi karyawan					
14	Dengan kemajuan teknologi sistem absensi					

	berbasis pemindaian wajah digunakan					
15	Sistem absensi berbasis pemindaian wajah mengurangi risiko kecurangan atau penyalahgunaan dalam pencatatan jam kerja karyawan					

Setelah menyebarkan kuesioner pada karyawan PT Wasco Engineering Indonesia maka akan didapatkan gambaran hasil pengumpulan data, dapat dilihat pada Tabel 5 dibawah ini :

Table 5 Gambaran Hasil Pengumpulan Data

Responden	Pertanyaan 1	Pertanyaan 2	Pertanyaan 3	...	Pertanyaan 15	Tingkat Kepuasan
1	4	5	4	2	3	81%
2	4	3	3	5	5	77%
3	4	5	4	5	5	81%
...	4	5	4	2	3	84%
100	5	3	4	4	3	77%

B. Tahap Preprocessing Data

Tahap *Preprocessing* data dilakukan untuk mengukur tingkat kepuasan karyawan berdasarkan atribut Kategori. Oleh karena itu, berdasarkan pertanyaan – pertanyaan pada form kuesioner akan dibagi menjadi tiga kategori untuk menghasilkan pohon keputusan yang akurat. *Preprocessing* data dapat dilihat seperti Tabel 6 dibawah ini :

Table 6 Preprocessing Data

No	Pertanyaan	Kategori
1	Sistem absensi berbasis pemindaian wajah dapat mengenali wajah dengan waktu <10s	Akurasi
2	Dengan menggunakan sistem absensi berbasis pemindaian wajah akurasi kehadiran karyawan meningkat	Akurasi
3	Sistem absensi berbasis pemindaian wajah dapat rekam kehadiran dengan akurat	Akurasi
4	Sistem absensi berbasis pemindaian wajah mampu mengenali jika ada perubahan pencahayaan maupun perubahan fisik pada wajah	Akurasi
5	Sistem absensi berbasis pemindaian wajah lebih akurat dalam rekam kehadiran dibandingkan metode absensi lainnya	Akurasi
6	Sistem absensi berbasis pemindaian wajah dapat berfungsi dengan baik	Fungsionalitas
7	Sistem absensi berbasis pemindaian wajah memudahkan karyawan untuk mencatat kehadiran	Fungsionalitas
8	Sistem absensi berbasis pemindaian wajah dapat mengenali wajah dengan akurat	Fungsionalitas
9	Sistem absensi berbasis pemindaian wajah meningkatkan efisiensi dalam mencatat kehadiran karyawan	Fungsionalitas
10	Sistem absensi berbasis pemindaian wajah mudah digunakan	Fungsionalitas
11	Anda selalu mengalami masalah saat menggunakan sistem absensi berbasis pemindaian wajah	Manfaat
12	Sistem absensi berbasis pemindaian wajah lebih efektif dalam mencatat kehadiran dibandingkan tapping badge	Manfaat
13	Sistem absensi berbasis pemindaian wajah memudahkan HR dalam mengelola absensi karyawan	Manfaat
14	Dengan kemajuan teknologi sistem absensi berbasis pemindaian wajah nyaman digunakan	Manfaat
15	Sistem absensi berbasis pemindaian wajah mengurangi risiko kecurangan atau penyalahgunaan dalam pencatatan jam kerja karyawan	Manfaat

Analisis kategori kepuasan responden telah dilakukan dengan mengelompokkan data yang terbagi menjadi 5 kategori yang berbeda. Kategori tersebut dirangkum dalam Tabel 7 dibawah ini :

Table 7 Kategori Kepuasan

Range Nilai	Kategori Kepuasan
20% - 35%	Sangat Tidak Puas
36% - 51%	Tidak Puas
52% - 67%	Netral
68% - 83%	Puas
84% - 100%	Sangat Puas

Setelah melakukan *preprocessing* data pada Tabel 6, maka akan diperoleh bentuk data hasil *preprocessing* seperti terlihat pada Tabel 8 dibawah ini :

Table 8 Gambaran Data Hasil Preprocessing

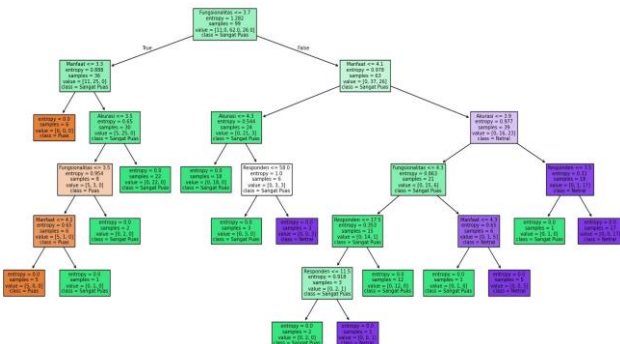
Responden	Akurasi	Fungsionalitas	Manfaat	Kategori Kepuasan
1	3.6	5	3.6	Puas
2	4	3.6	4	Puas
3	4.2	3.8	4.2	Puas
...	4.4	3.8	4.4	Sangat Puas
100	4	3.8	2.8	Puas

C. Tahap Klasifikasi Pohon Keputusan



Gambar 3 Klasifikasi Pohon Keputusan

Tahap klasifikasi pohon keputusan ini diawali dengan memasukkan hasil *preprocessing* data kuesioner karyawan. Kemudian menghitung nilai *entropy* dan *gain* pada setiap atribut berdasarkan data yang sudah dimasukkan, selanjutnya *gain* tertinggi akan menjadi akar. Kemudian dilanjutkan dengan proses penentuan cabang dan menghasilkan pohon keputusan berdasarkan Algoritma C4.5. Gambaran pohon keputusan yang akan didapatkan adalah seperti Gambar 4 dibawah ini :



Gambar 4 Gambaran Pohon Keputusan

D. Tahap Pengujian dan Evaluasi

Pada tahap pengujian dilakukan proses evaluasi terhadap seluruh dataset yang diperoleh dalam penelitian ini. Dataset tersebut akan diterapkan pada data kepuasan karyawan terhadap sistem absensi berbasis pemindaian wajah menggunakan bahasa pemrograman Python. Proses pengujian ini bertujuan

untuk mengukur kinerja model yang dihasilkan. Semakin tinggi nilai *accuracy* maka semakin baik pula model yang dihasilkan, kemudian hasil yang didapat akan dianalisa dan dievaluasi [10]. Untuk mendapatkan nilai akurasi dari rule yang dihasilkan maka data sampel terbagi menjadi 2 bagian yaitu data training dan data testing.

Data training adalah data sebenarnya yang sesuai dengan fakta sedangkan data testing adalah data yang digunakan untuk mengukur keakuratan hasil klasifikasi dengan benar [13]. Pembagian data training dan data testing merupakan salah satu faktor yang menentukan akurasi, sehingga jika terjadi kesalahan dalam menentukan komposisi kedua tipe data tersebut akan mempengaruhi nilai akurasi dan presisi yang diperoleh [17]. Pembagian ini bertujuan untuk memastikan bahwa model yang dibangun dapat diuji secara efektif, penelitian ini melakukan percobaan klasifikasi dengan persentase data training dan data testing yang berbeda. Gambaran hasil pengujian dapat dilihat pada Table 9 dibawah ini :

Table 9 Gambaran Hasil Pengujian

Data Training	Data Testing	Nilai Akurasi
95%	5%	95.71%
90%	10%	92%
85%	15%	88.53%
80%	20%	86.25%
75%	25%	80.33%

IV. HASIL DAN PEMBAHASAN

Implementasi Algoritma C4.5 untuk mengukur tingkat kepuasan karyawan terhadap sistem absensi berbasis pemindaian wajah dengan bahasa pemrograman Python terdiri dari beberapa tahapan sebagai berikut :

A. Hasil Tahap Pengumpulan Data

Hasil tahap pengumpulan data yang disebar menggunakan Google Form kepada karyawan PT Wasco Engineering Indonesia sebanyak 100 sampel, dapat dilihat hasil pengumpulan data kuesioner pada Tabel 10 dibawah ini :

Table 10 Hasil Pengumpulan Data

Pertanyaan	Sangat Tidak Setuju	Tidak Setuju	Netral	Setuju	Sangat Setuju
1	0	3	7	36	37
2	0	2	7	34	45
3	1	4	4	38	50
4	2	13	12	40	48
5	0	3	15	45	49
6	0	2	7	46	46
7	0	1	10	39	34
8	0	4	7	41	50
9	0	2	11	38	46
10	0	4	7	43	53
11	2	9	15	40	59
12	3	6	8	33	0
13	2	1	13	38	0
14	1	2	10	34	2
15	2	0	4	35	1

B. Hasil Tahap Preprocessing Data

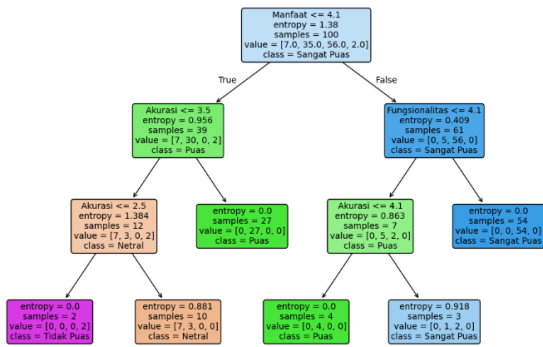
Hasil tahap *preprocessing* data terdiri dari beberapa atribut yaitu Responden, Akurasi, Fungsionalitas, Manfaat dan Kategori Kepuasan. Hasil *preprocessing* dapat kita lihat pada Tabel 11 dibawah ini :

Table 11 Hasil Preprocessing Data

Responden	Akurasi	Fungsionalitas	Manfaat	Kategori Kepuasan
1	5	4.8	4.6	Sangat Puas
2	3.6	4	4	Puas
3	3.6	4	4	Puas
4	4.4	4.6	4.6	Sangat Puas
...
100	3.8	3	4	Puas

C. Hasil Tahap Klasifikasi Pohon Keputusan

Pada tahap klasifikasi pohon keputusan menggunakan bahasa pemrograman Python dengan target atribut yaitu kategori kepuasan. Berikut ini merupakan hasil pohon keputusan dapat kita lihat pada Gambar 5 dibawah ini :



Gambar 5 Hasil Pohon Keputusan

Hasil tahap klasifikasi pohon keputusan terdiri dari beberapa rule sebagai berikut :

1. Root
Dengan jumlah 100 sampel menghasilkan atribut Manfaat <= 4.1, nilai entropy = 1.38, value nilai 7 sampel Netral, 35 sampel Puas, 56 sampel Sangat Puas dan 2 sampel Tidak Puas dan kelas Sangat Puas.
2. Node 1
Dengan jumlah 39 sampel menghasilkan atribut Akurasi <= 3.5, nilai entropy = 0.956, value nilai 7 sampel Netral, 30 sampel Puas, 0 sampel Sangat Puas dan 2 sampel Tidak Puas dan kelas Puas.
3. Node 2
Dengan jumlah 12 sampel menghasilkan atribut Akurasi <= 2.5, nilai entropy = 1.384, value nilai 7 sampel Netral, 3 sampel Puas, 0 sampel Sangat Puas dan 2 sampel Tidak Puas dan kelas Netral.
4. Leaf 1
Dengan jumlah 27 sampel menghasilkan nilai entropy = 0.0, value nilai 0 sampel Netral, 5 sampel Puas, 56 sampel Sangat Puas dan 0 sampel Tidak Puas dan kelas Sangat Puas.
5. Leaf 2
Dengan jumlah 10 sampel menghasilkan nilai entropy = 0.881, value nilai 7 sampel Netral, 3 sampel Puas, 0 sampel Sangat Puas dan 0 sampel Tidak Puas dan kelas Netral.
6. Leaf 3
Dengan jumlah 2 sampel menghasilkan nilai entropy = 0.0, value nilai 0 sampel Netral, 0 sampel Puas, 0 sampel Sangat Puas dan 2 sampel Tidak Puas dan kelas Tidak Puas.
7. Node 3
Dengan jumlah 61 sampel menghasilkan atribut Akurasi <= 4.1, nilai entropy = 0.409, value nilai 0 sampel Netral, 5 sampel Puas, 56 sampel Sangat Puas dan 0 sampel Tidak Puas dan kelas Sangat Puas.
8. Node 4
Dengan jumlah 7 sampel menghasilkan atribut Akurasi <= 4.1, nilai entropy = 0.863, value nilai 0 sampel Netral, 5 sampel Puas, 2 sampel Sangat Puas dan 0 sampel Tidak Puas dan kelas Puas.
9. Leaf 1
Dengan jumlah 54 sampel menghasilkan nilai entropy = 0.0, value nilai 0 sampel Netral, 0 sampel Puas, 54 sampel Sangat Puas dan 0 sampel Tidak Puas dan kelas Sangat Puas.
10. Leaf 2
Dengan jumlah 3 sampel menghasilkan nilai entropy = 0.918, value nilai 0 sampel Netral, 1 sampel Puas, 2 sampel Sangat Puas dan 0 sampel Tidak Puas dan kelas Sangat Puas.

11. Leaf 3

Dengan jumlah 4 sampel menghasilkan nilai entropy = 0.0, value nilai 0 sampel Netral, 4 sampel Puas, 0 sampel Sangat Puas dan 0 sampel Tidak Puas dan kelas Puas.

Maka dapat ditarik kesimpulan rule dari pohon keputusan sebagai berikut :

1. Jika Manfaat <= 4.1 dan Akurasi <= 3.5 serta Akurasi <= 2.5 = True, maka class = Netral.
2. Jika Manfaat <= 4.1 dan Akurasi <= 3.5 serta Akurasi > 2.5, maka class = Puas.
3. Jika Manfaat > 4.1 dan Fungsionalitas <= 4.1 serta Akurasi <= 4.1 = False, maka class = Puas.
4. Jika Manfaat > 4.1 dan Fungsionalitas > 4.1, maka class = Sangat Puas

D. Hasil Tahap Pengujian dan Evaluasi

Hasil tahap pengujian dan evaluasi menggunakan bahasa pemrograman Python dengan 100 sampel terbagi menjadi 2 bagian yaitu data training dan data testing. Berikut ini merupakan hasil pengujian sebagai berikut :

Table 12 Hasil Pengujian dan Evaluasi

Data Training	Data Testing	Nilai Akurasi
95%	5%	90.54%
90%	10%	88.89%
85%	15%	88.23%
80%	20%	86.25%
75%	25%	85.24%

V. KESIMPULAN

Dari hasil pembahasan mengenai Implementasi Algoritma C4.5 Dalam Mengukur Tingkat Kepuasan Karyawan Terhadap Sistem Absensi Berbasis Pemindaian Wajah, dapat diambil kesimpulan sebagai berikut :

1. Implementasi Algoritma C4.5 dapat mengukur tingkat kepuasan karyawan terhadap sistem absensi berbasis pemindaian wajah dengan nilai akurasi sebesar 90.54%
2. Sistem absensi berbasis pemindaian wajah sudah berjalan dengan baik terlihat pada hasil dari pohon keputusan dan pengujian model pohon keputusan.
3. Dengan pengujian model sebanyak 5 kali yang sudah dilakukan, dapat dikatakan sangat baik karena terlihat jelas pada nilai akurasi semakin besar di setiap perbandingannya.

DAFTAR PUSTAKA

[1] A. Muzakir And R. A. Wulandari, "Model Data Mining Sebagai Prediksi Penyakit Hipertensi Kehamilan Dengan Teknik Decision Tree," Sci. J. Informatics, Vol. 3, No. 1, Pp. 19–26, 2016, Doi: 10.15294/Sji.V3i1.4610

[2] Adriansa, M., Yulianti, L., & Elfianty, L. (2022). Analisis Kepuasan Pelanggan Menggunakan Algoritma C4.5. Jurnal Teknik Informatika UNIKA Santo Thomas, 07(21), 115–121. <https://doi.org/10.54367/jtiust.v7i1.1983>

[3] Alkhairi, P., & Situmorang, Z. (2022). Penerapan Data Mining Untuk Menganalisis Kepuasan Pegawai Terhadap Pelayanan Bidang SDM dengan Algoritma C4.5. Jurassic (Jurnal Riset Sistem Informasi Dan Teknik Informatika), 7(1), 40. <https://doi.org/10.30645/jurasik.v7i1.414>

[4] Bramer, Max (2007) Principles of Data Mining, Springer Science

[5] Han, J. & M. Kamber. (2001). Data Mining Concept and Techniques. Morgan Kaufmann Publishers, San Francisco. Publishers, San Francisco.

[6] J. N. L. (2016). Konversi Data Training Tentang Pemilihan Kelas. 145–154.

- [7] Jefri, 2013, "Implementasi Algoritma C4.5 Dalam Aplikasi Untuk Memprediksi Jumlah Mahasiswa Yang Mengulang Mata Kuliah Di STM IK AMIKOM Yogyakarta", Yogyakarta.
- [8] Kurniah, R., Surya Putra, D. Y., & Diana, E. (2022). Penerapan Data Mining *Decision Tree* Algoritma C4.5 Untuk Mengetahui Tingkat Kepuasan Mahasiswa Terhadap Layanan Akademik Dan Kemahasiswaan (Studi Kasus Universitas Prof.Dr. Hazairin,SH). Infotek : Jurnal Informatika Dan Teknologi, 5(2), 316–326. <https://doi.org/10.29408/jit.v5i2.5910>
- [9] M F. Arifin And D. Fitriannah, "Rekomendasi Penerimaan Mitra Penjualan Studi Kasus : Pt Atria Artha Persada," J. Telekomun. Dan Komputer., Vol. 8, No. 2, Pp. 87–101, 2018, Doi: 10.22441/Incomtech.V8i1.2198.
- [10] Mandasari, V., Tama, B. A., & Sriwijaya, U. (2011). Analisis Kepuasan Konsumen Terhadap
- [11] N Azwanti, "Algoritma C4.5 Untuk Memprediksi Mahasiswa Yang Mengulang Mata Kuliah (Studi Kasus Di Amik Labuhan Batu)," Simetris J. Tek. Mesin, Elektro dan Ilmu Komputer., vol. 9, no. 1, pp. 316-326 e-ISSN 2614-8773 DOI: 10.29408/jit.v5i2.5910 Link : <https://dx.doi.org/10.29408/jit.v5i2.5910> Infotek : Jurnal Informatika dan Teknologi – Vol.5 No. 2 Juli 2022 326 11–22, 2018, doi: 10.24176/simet.v9i1.1627.
- [12] Quinlan, J. R., (1993). C4.5: Programs for Machine Learning. San Mateo: Morgan Kaufmann Publishers.
- [13] R. A. Anggraini, G. Widagdo, A. S. Budi, and M. Qomaruddin, "Penerapan Data Mining Classification untuk Data Blogger Menggunakan Metode Naïve Bayes," J. Sist. dan Teknol. Inf., vol. 7, no. 1, p. 47, 2019, doi: 10.26418/justin.v7i1.30211.
- [14] Sari, R., D., I., dan Sindunata, Y., 2014, Penerapan Data Mining untuk Analisa Pola Perilaku Nasabah dalam Pengkreditan Menggunakan Metode C.45 Studi Kasus pada KSU Insan Kamil Demak, <https://lp2m.asia.ac.id/wp-content/uploads/2015/05/JURNAL-RINA-DEWI.pdf>, diakses 20 Juni 2018
- [15] Tulus Ujjianto, N., & Ramdhan, N. A. (2022). Implementasi Data Mining C4.5 Dalam Mengukur Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Akademik. *Jurnal Ilmiah Intech: Information Technology Journal of UMUS*, 4(01), 105–111.
- [16] Turban, E., dkk. (2005). *Decision Support Systems and Intelligent Systems*. Yogyakarta: Andi.
- [17] W. Musu, A. Ibrahim, and Heriadi, "Pengaruh Komposisi Data Training dan Testing terhadap Akurasi Algoritma C4.5," in *Seminar Sistem Informasi dan Teknologi Informasi (SISITI)*, 2021, pp. 186–195.
- [18] Walangitan, J., Sompie, S. R. U. A., & Naj Joan, X. B. N. (2024). Sistem Absensi Pengenalan Wajah Bermasker. *Jurnal Teknik Informatika*, 19(01), 21–30. <https://doi.org/10.35793/jti.v19i01.51327>