

Penerapan Algoritma C4.5 Dalam Melakukan Seleksi Pemain Tengah Pada Tim Nasional Sepakbola

Gabereal Pandapotan Pakpahan¹, Wilda Susanti²

^{1,2}Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Institut Bisnis dan Teknologi Pelita Indonesia

e-mail: gabereal@student.pelitaindonesia.ac.id, wilda@lecturer.pelitaindonesia.ac.id

Abstrak

Seleksi pemain tengah pada tim nasional sepak bola merupakan proses penting yang menentukan kualitas permainan tim. Namun, proses ini sering dilakukan tanpa dukungan analisis data yang objektif, sehingga berpotensi menimbulkan keputusan yang tidak tepat. Hal ini mendorong perlunya sistem seleksi berbasis data untuk membantu pelatih dalam mengambil keputusan yang akurat. Penelitian ini bertujuan menerapkan algoritma C4.5 untuk membangun model klasifikasi dalam menentukan kelayakan pemain tengah berdasarkan data performa. Data yang digunakan terdiri dari 1069 pemain tengah asal Spanyol yang memiliki atribut seperti *passing accuracy*, *ball recoveries*, *key passes per match*, *expected assists (xA)*, *dribbles*, dan *minutes played*. Metode yang digunakan meliputi perhitungan manual nilai *entropy* dan *information gain* untuk membentuk pohon keputusan, serta implementasi algoritma C4.5 menggunakan *Python* untuk validasi model. Hasil perhitungan menunjukkan bahwa atribut *ball recoveries* dan *key passes per match* memiliki nilai *gain* tertinggi sebesar 0,881 dan menjadi dasar pemisahan kelas pada decision tree. Hasil penelitian menunjukkan bahwa algoritma C4.5 mampu memprediksi seleksi pemain tengah dengan tingkat akurasi 100% berdasarkan *data testing* dan *training*. Sistem menampilkan 3 pemain terpilih yaitu Fabian Ruiz, Pedri, Rodri dan juga menampilkan rekomendasi 10 pemain terbaik yang layak dipertimbangkan untuk Tim Nasional Spanyol.

Kata Kunci: C4.5, Seleksi Pemain, Sepak Bola, Data

Abstract

The The selection of midfield players for the national football team is a crucial process that determines the overall quality of the team's performance. However, this process is often carried out without the support of objective data analysis, which may lead to inaccurate decisions. This highlights the need for a data-driven selection system to assist coaches in making accurate decisions. This study aims to apply the C4.5 algorithm to build a classification model for determining the eligibility of midfield players based on performance data. The dataset consists of 1,069 Spanish midfield players with attributes such as passing accuracy, ball recoveries, key passes per match, expected assists (xA), dribbles, and minutes played. The method involves manual calculation of entropy and information gain values to construct the decision tree, as well as implementing the C4.5 algorithm using Python for model validation. The results show that ball recoveries and key passes per match have the highest gain value of 0.881 and serve as the basis for class separation in the decision tree. The result reveal that the C4.5 algorithm is capable of predicting midfield player selection with 100% accuracy based on both training and testing data. The system outputs three selected players Fabián Ruiz, Pedri, Rodri and also provides recommendations for the top 10 players who are most suitable to be considered for the Spanish National Team.

Keywords: C4.5, Player Selection, Football, Data

1. Pendahuluan

Teknologi komputer telah berkembang pesat dan menjadi tulang punggung dalam berbagai bidang, termasuk olahraga [1]. Salah satu pemanfaatan teknologi komputer relevan dalam konteks olahraga modern adalah data mining, yaitu proses mengekstraksi informasi penting dari kumpulan data besar untuk mendukung pengambilan keputusan. Dalam dunia olahraga seperti sepak bola, data mining memungkinkan analisis mendalam terhadap performa atlet berdasarkan data statistik yang tersedia.

Tim nasional Spanyol dikenal dengan filosofi permainan tiki-taka yang menuntut kualitas tinggi di lini tengah. Melimpahnya jumlah gelandang profesional asal Spanyol menjadikan proses seleksi semakin kompleks, terlebih tanpa dukungan sistem berbasis data. Kasus dilema seleksi menjelang Piala Dunia 2022 menunjukkan pentingnya pendekatan yang lebih ilmiah dalam menentukan komposisi tim [2]. Pentingnya seleksi pemain dalam sepak bola, terutama pada posisi pemain tengah, tidak dapat diabaikan. Pemain tengah memiliki peran strategis sebagai penghubung antara lini pertahanan dan lini serang, serta menjadi pengatur ritme permainan di lapangan. Oleh karena itu, dibutuhkan pendekatan berbasis teknologi seperti data mining yang dapat memberikan gambaran objektif mengenai performa setiap pemain berdasarkan indikator yang relevan.

Sebagai solusi terhadap permasalahan tersebut, salah satu metode yang potensial untuk digunakan dalam hal ini adalah algoritma C4.5, menurut [3] yang mampu menyusun pohon keputusan berdasarkan atribut-atribut performa untuk menghasilkan klasifikasi yang dapat dijadikan dasar seleksi. Algoritma C4.5 memiliki kemampuan menyusun pohon keputusan yang jelas dan mudah dipahami, di mana setiap cabang mewakili kriteria penting seperti akurasi umpan, jumlah tekel sukses, jumlah assist, kehilangan bola, dan riwayat cedera. Dengan model ini, proses seleksi dapat dilakukan secara lebih ilmiah, objektif, dan transparan [4].

Penelitian yang dilakukan oleh [5] yang berjudul “Algoritma C4.5 untuk Penentuan Tim Pemain Utama Olahraga Voli”. Dalam studi ini, algoritma C4.5 digunakan untuk mendukung proses klasifikasi dalam pemilihan pemain inti dalam olahraga voli. Hasil penelitian menunjukkan kesesuaian antara perhitungan manual dan hasil dari program komputer, membuktikan keakuratan dan konsistensi algoritma dalam membantu pengambilan keputusan yang berbasis data.

Penelitian serupa juga dilakukan oleh [6] dengan judul “Sistem Informasi Pengelompokan Pemain Sepak Bola Menggunakan Data Mining Klasifikasi Pada PSSI Provinsi Gorontalo”. Dalam studi tersebut, digunakan metode algoritma C4.5 untuk mengelompokkan pemain berdasarkan performa, seperti akurasi umpan, intersep, assist, kehilangan bola, jumlah pelanggaran, hingga catatan cedera. Penelitian ini menitikberatkan pada proses klasifikasi pemain tengah untuk level tim nasional.

Dengan mengadopsi pendekatan berbasis algoritma C4.5 dan memanfaatkan data performa aktual dari sumber resmi, penelitian ini diharapkan dapat memberikan kontribusi nyata dalam pengambilan keputusan seleksi pemain secara lebih objektif dan ilmiah. Sistem ini diharapkan dapat menjadi alat bantu bagi pelatih tim nasional dalam menyusun skuad yang optimal, khususnya dalam menghadapi turnamen besar seperti Piala Dunia.

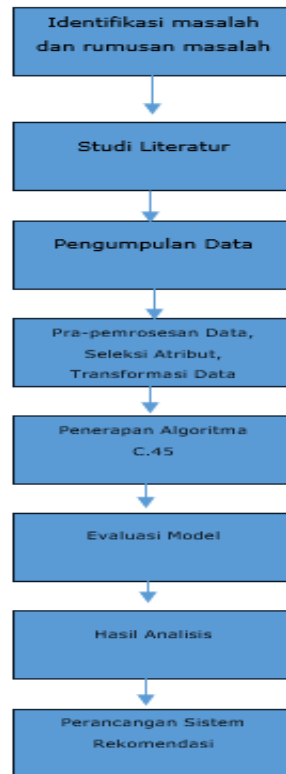
2. Metode Penelitian

2.1. Tempat dan Waktu Penelitian

Penelitian ini dilakukan dengan mengambil data di *Real Federación Española de Fútbol* (RFEF) dan juga beberapa situs penyedia data sepak bola. Pemilihan lokasi ini didasarkan pada ketersediaan data pemain timnas dan data seleksi yang relevan dengan penelitian. Data pemain yang dikumpulkan sejumlah 2769 selama kurun waktu tiga tahun terakhir yaitu dari tahun 2022 hingga 2025 yang mencakup berbagai umur pemain guna mendapatkan variasi data yang cukup

untuk melakukan prediksi seleksi pemain. Pengumpulan data berlangsung selama data dan variable yang berkaitan terjadi nya seleksi pemain dengan fokus pada data historis.

2.2. Kerangka Penelitian



Gambar 1. Kerangka Penelitian

2.2.1 Identifikasi Variabel

Penelitian ini berfokus pada seleksi pemain tengah Tim Nasional Sepak Bola Spanyol dengan menerapkan metode Algoritma C.45. Pemilihan pemain dilakukan berdasarkan berbagai indikator performa yang relevan dengan peran pemain tengah dalam tim sepak bola. Variabel-variabel yang digunakan dalam penelitian ini terdiri dari :

1. Variabel Independen :
 - a. *Passing Accuracy (%)* : Persentase akurasi operan yang dilakukan oleh pemain dalam pertandingan. Akurasi umpan sangat penting bagi pemain tengah dalam menjaga ritme permainan dan membangun serangan.
 - b. *Key Passes per Match* : Jumlah umpan kunci yang berkontribusi terhadap peluang mencetak gol. Pemain tengah yang sering memberikan umpan kunci cenderung lebih kreatif dalam membangun serangan.
 - c. *Ball Recoveries per Match* : Kemampuan pemain dalam mendapatkan kembali bola setelah kehilangan penguasaan. Pemain yang memiliki tingkat pemulihan bola tinggi dapat membantu tim dalam mempertahankan kontrol permainan.
 - d. *Expected Assists (xA)* : Perkiraan jumlah assist yang dapat diberikan oleh pemain berdasarkan kualitas umpan yang dilakukan.
 - e. *Heatmap & Positional Data* : Data mengenai pergerakan pemain di lapangan untuk memahami peran dan kontribusi mereka dalam sistem permainan tim.
 - f. *Minutes Played* : Data mengenai sebanyak apa menit bermain seorang pemain
 - g. *Dribbles* : Jumlah *dribble* sukses

2. Variabel Dependen : Status Seleksi (Terpilih/Tidak Terpilih) – Keputusan akhir apakah pemain masuk dalam tim nasional atau tidak berdasarkan evaluasi performa dan hasil pemodelan Algoritma C.45.

2.2.2 Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari berbagai sumber terpercaya, antara lain :

1. Statistik pertandingan resmi dari liga-liga top Eropa seperti *La Liga*, *Premier League*, *Bundesliga*, *Serie A*, dan *Ligue 1*.
2. Situs analisis sepak bola seperti *Opta*, *Whoscored*, *Transfermarkt*, dan *StatsBomb* yang menyediakan data mendalam mengenai performa pemain.
3. Rekaman pertandingan yang digunakan untuk analisis kualitatif dalam memahami gaya bermain pemain.
4. Jurnal dan penelitian terdahulu yang membahas metode seleksi pemain berdasarkan data statistik.

Data yang dikumpulkan mencakup statistik performa pemain tengah dari berbagai liga dalam periode tiga tahun terakhir (2022-2025). Selain itu, referensi tambahan dari buku dan publikasi ilmiah digunakan untuk memperkuat landasan teori dan metode penelitian.

2.2.3 Pengolahan Data

Setelah data terkumpul, langkah selanjutnya adalah pengolahan data yang terdiri dari beberapa tahap :

- a. Pembersihan Data (*Data Cleaning*) : Tahap ini melibatkan penghapusan data yang tidak valid, penanganan nilai yang hilang (*missing values*), serta eliminasi *outlier* (anomali) yang dapat mengganggu hasil analisis dan akurasi model prediksi.
- b. Normalisasi Data : Variabel-variabel independen seperti *Passing Accuracy* dan *Ball Recovery* dinormalisasi agar berada pada skala yang seimbang. Hal ini diperlukan agar algoritma C4.5 tidak berat sebelah terhadap atribut tertentu yang memiliki nilai ekstrem.
- c. *Feature Engineering* : Pada tahap ini, dilakukan pembuatan fitur baru atau transformasi fitur lama untuk meningkatkan kualitas informasi yang digunakan dalam proses seleksi. Misalnya, menggabungkan variabel *Key Pass* dan *Assist* menjadi satu variabel baru jika diperlukan.
- d. Pembagian Data : Data dibagi menjadi dua bagian, 80% digunakan untuk pelatihan (*training set*) dan 20% digunakan untuk pengujian (*testing set*)

2.2.4 Penerapan Algoritma C.45

Algoritma C4.5 digunakan untuk menyusun pohon keputusan berdasarkan atribut performa pemain, seperti akurasi umpan, *assist*, tekel sukses, intersep, kehilangan bola, dan riwayat cedera. Dengan menghitung *entropy* dan *gain*, algoritma ini menentukan atribut terbaik untuk klasifikasi sehingga menghasilkan kategori “layak dipilih” atau “tidak layak dipilih.” Dalam konteks tim nasional, khususnya posisi pemain tengah, penerapan C4.5 memberi seleksi yang lebih objektif, transparan, dan mudah dipahami. Penelitian sebelumnya [7] juga membuktikan efektivitas C4.5 pada olahraga, sehingga relevan untuk mendukung seleksi pemain berbasis data. Ada beberapa tahap dalam membuat sebuah pohon keputusan dengan algoritma C4.5 yaitu:

- a. Mempersiapkan data training, dapat diambil dari data histori yang pernah terjadi sebelumnya dan sudah dikelompokkan dalam kelas - kelas tertentu.
- b. Menentukan akar dari pohon dengan menghitung nilai *gain* yang tertinggi dari masing-masing atribut atau berdasarkan nilai index *entropy* terendah. Sebelumnya dihitung terlebih dahulu nilai index *entropy*, dengan rumus :

$$Entropy(i) = \sum_{j=1}^m f(i,j) \cdot 2f[(i,j)] \quad (1)$$

- c. Hitung nilai *Gain* dengan rumus:

$$gain = \sum_{i=1}^p \frac{n_i}{n} \cdot ie(i) \quad (2)$$

- d. Selanjutnya menghitung Gain ratio :
- $$Gainratio(S, A) = \frac{Gain(S, A)}{Splitinformation(S, A)} \quad (3)$$
- e. Ulangi langkah ke-2 hingga semua record terpartisi. Proses partisi pohon keputusan akan berhenti disaat :
1. Semua tupel dalam record dalam simpul N mendapat kelas yang sama,
 2. Tidak ada atribut dalam record yang dipartisi lagi,
 3. Tidak ada record dalam cabang yang kosong.
- f. *Pruning*

2.2.5 Evaluasi Model

Evaluasi model dilakukan untuk mengukur tingkat keberhasilan model dalam memprediksi seleksi pemain. Beberapa teknik evaluasi yang digunakan antara lain :

- a) Akurasi Model : Mengukur seberapa akurat prediksi model dibandingkan dengan data aktual seleksi pemain [8].
- b) *Confusion Matrix* : Menganalisis kesalahan klasifikasi dalam prediksi model untuk mengidentifikasi area yang perlu diperbaiki [9].
- c) *Feature Importance* : Mengidentifikasi variabel yang paling berpengaruh dalam keputusan seleksi pemain berdasarkan hasil Algoritma C.45 [10].
- d) *Cross-Validation* : Menggunakan teknik *k-fold cross-validation* untuk memastikan model tidak mengalami overfitting dan dapat bekerja dengan baik pada data baru [11].

2.2.6 Hasil Analisis

Hasil dari model Algoritma C.45 dianalisis untuk menentukan pemain yang layak masuk ke dalam Tim Nasional Spanyol. Pemain yang terpilih dibandingkan dengan daftar pemain yang sebelumnya dipanggil oleh pelatih untuk melihat tingkat kesesuaian model dengan seleksi manual. Jika model memiliki tingkat akurasi tinggi, metode ini dapat digunakan sebagai alat bantu seleksi pemain yang lebih objektif dan berbasis data.

2.2.7 Perancangan Sistem Rekomendasi

Perancangan sistem rekomendasi bertujuan untuk membantu proses pengambilan keputusan dengan menyajikan alternatif terbaik berdasarkan data yang tersedia. Dalam konteks seleksi pemain sepak bola, khususnya pemain tengah tim nasional, sistem rekomendasi dirancang untuk mengolah data performa pemain melalui algoritma klasifikasi, seperti C4.5, sehingga menghasilkan rekomendasi pemain yang layak dipilih. Proses ini meliputi tahap pengumpulan data, *preprocessing*, penerapan algoritma, hingga penyajian hasil dalam bentuk visualisasi yang mudah dipahami.

Dengan adanya sistem rekomendasi, pelatih dan manajemen tim dapat memperoleh dasar keputusan yang lebih objektif, transparan, serta mengurangi risiko bias subjektif. Selain itu, sistem ini juga mempercepat proses seleksi karena dapat mengolah jumlah data yang besar secara efisien dan konsisten.

2.2.8 Perhitungan Manual

Data *training* yang digunakan,

Jumlah total dataset	: 15 data
Jumlah klasifikasi	
Layak	: 3 data
Tidak Layak	: 7 data

Tabel 1. Data *Training* Digunakan

No	Nama	Passing Acc.	Ball Rec.	Key Passes / Match	Expected Assists (xA)	Dribbles	Minutes Played	Kelayakan
1	Jaime Z.	81.2	4.4	4.4	0.47	35	2250	✗ Tidak
2	Manuel Z.	85.6	4.5	1.7	4.5	0.54	35	✗ Tidak
3	David Z.	87.4	2.2	1.4	2.2	0.57	30	✗ Tidak
4	Rodri	86.9	6.3	5.5	0.8	6.3	3204	✓ Layak
5	Pedri	89.6	6.0	4.5	5.9	0.67	3308	✓ Layak
6	Fabian Ruiz	89.5	6.1	5.5	6.1	0.67	3100	✓ Layak
7	Mikel Ziganda	85.4	3.6	1.3	3.6	0.27	35	✗ Tidak
8	Claudia Zornoza	83.2	2.7	1.8	2.7	0.23	35	✗ Tidak
9	Elady Zorrilla	82.8	4.3	1.9	4.3	0.57	38	✗ Tidak
10	Gonzalo Zorrilla	78.6	2.8	1.7	2.8	0.26	38	✗ Tidak
11	Sergio Santos	81.9	1.2	3.4	0.51	35	2250	✗ Tidak
12	Aitor Sanz	80	1.4	4.4	0.45	35	2250	✗ Tidak
13	Alejandro Sanz	85.6	2.3	4.7	0.51	35	2250	✗ Tidak
14	Alvaro Sanz	80.2	1.3	3.8	0.54	35	2250	✗ Tidak
...	...	...	...	...	...	...	...	...
1069	Martin Zubimendi	82.2	0.3	6.5	0.18	35	2250	✗ Tidak

Kriteria klasifikasi (asumsi):

Layak jika Passing Accuracy ≥ 86 dan Ball Recovery ≥ 6.0

Selain itu: Tidak Layak

Perhitungan Entropy:

- Split Berdasarkan Passing Accuracy (≥ 86)

Kategori:

$\geq 86 \rightarrow$ Rodri, David, Pedri, Fabian $\rightarrow$ 2 Layak, 2 Tidak

$< 86 \rightarrow$ Jaime, Manuel $\rightarrow$ 0 Layak, 2 Tidak

Hitung Entropy masing-masing subset:

$\geq 86 \rightarrow$ 4 data: 2 Layak, 2 Tidak

$$Entropy = -(0.5 \log_2 0.5 + 0.5 \log_2 0.5) = 1.0$$

$< 86 \rightarrow$ 2 data: 0 Layak, 2 Tidak

$$Entropy = -(1 \cdot \log_2 1) = 0$$

$$Entropy\ Passing = \frac{4}{6}(1.0) + \frac{2}{6}(0) = 0.667$$

$$Gain\ Passing = Entropy(s) - entropy\ passing = 0.981 - 0.667 = 0.251$$

Tabel 2. *Split Passing*

Kategori	Jumlah	Layak	Tidak	Entropy
≥ 86	4	2	2	1.000
< 86	6	1	5	0.650

2. Split Berdasarkan Ball Recovery (≥ 6.0)

$$Entropy\ Passing = 4/6 (1.0) + 2/6 (0) = 0.667$$

$$Gain\ Passing = Entropy(s) - entropy\ passing = 0.981 - 0.667 = 0.251$$

[Ball Recovery ≥ 6.0]

/ \
Yes No
Layak Tidak Layak

Tabel 3. *Split Recovery*

Kategori	Jumlah	Layak	Tidak	Entropy
≥ 6.0	3	3	0	0.000
< 6.0	7	0	7	0.000

3. Key Passes / Match (Threshold ≥ 4.5)

a. ≥ 4.5 : Rodri, Pedri, Fabian $\rightarrow$ 3 Layak

b. < 4.5 : 7 data $\rightarrow$ 0 Layak, 7 Tidak

c. $\geq 4.5 \rightarrow$ 3 Layak $\rightarrow$ Entropy = 0

d. $< 4.5 \rightarrow$ 0 Layak, 7 Tidak $\rightarrow$ Entropy = 0

$$Entropy\ KeyPass = 0\ Gain = 0.881$$

4. Expected Assists (xA) (Threshold ≥ 4.5)

a. ≥ 4.5 : Manuel, Pedri, Fabian, Mikel, Elady $\rightarrow$ 2 Layak, 3 Tidak

b. < 4.5 : Rodri, Jaime, David, Claudia, Gonzalo $\rightarrow$ 1 Layak, 4 Tidak

Entropy:

a) $\geq 4.5 \rightarrow$ 2 Layak, 3 Tidak

$$Entropy = -\left(\frac{2}{5}\log_2\frac{2}{5} + \frac{3}{5}\log_2\frac{3}{5}\right) \approx 0.971$$

b) $< 4.5 \rightarrow$ 1 Layak, 4 Tidak

$$Entropy = -\left(\frac{1}{5}\log_2\frac{1}{5} + \frac{4}{5}\log_2\frac{4}{5}\right) \approx 0.722$$

$$Entropy\ xA = -\left(\frac{5}{10}(0.971) + \frac{5}{10}(0.722)\right) = 0.846$$

$$Gain\ xA = 0.881 + 0.84 = 0.035$$

5. Dribbles (Threshold ≥ 0.6)

a. ≥ 0.6 : Rodri, Pedri, Fabian, David $\rightarrow$ 3 Layak, 1 Tidak

b. < 0.6 : Sisanya $\rightarrow$ 0 Layak, 6 Tidak

c. $\geq 0.6 \rightarrow$ 3 Layak, 1 Tidak

$$Entropy = 0.811$$

a. $< 0.6 \rightarrow$ 0 Layak, 6 Tidak $\rightarrow$ Entropy = 0

$$Entropy\ Dribbles = \frac{4}{10}(0.811) + \frac{6}{10}(0) = 0.324$$

$$Gain\ Dribbles = 0.881 - 0.324 = 0.557$$

6. Minuted Played (Threshold ≥ 40)

a. ≥ 40 : Pedri, Fabian $\rightarrow$ 2 Layak

b. < 40 : 8 data (1 Layak, 7 Tidak)

c. $\geq 40 \rightarrow$ 2 Layak, 0 Tidak $\rightarrow$ Entropy = 0

d. $< 40 \rightarrow$ 1 Layak, 7 Tidak

$$\text{Entropy} = -\left(\frac{1}{8}\log_2\frac{1}{8} + \frac{7}{8}\log_2\frac{7}{8}\right) \approx 0.543$$

$$\text{EntropyMinutes} = \frac{2}{10}(0) + \frac{8}{10}(0.543) = 0.435$$

$$\text{GainMinutes} = 0.881 - 0.435 = 0.446$$

Dari hasil di bawah, atribut Ball Recovery dan Key Passes / Match memiliki Gain tertinggi (0.881) dan akan menjadi atribut pemisah utama pada akar decision tree.

Tabel 4. Atribut *Gain* Tertinggi

Atribut	Entropy	Gain
Ball Recovery	0.000	0.881
Key Passes / Match	0.000	0.881
Passing Accuracy	0.324	0.557
Dribbles	0.324	0.557
Minutes Played	0.435	0.446
Expected Assists (xA)	0.846	0.035

2.2.9 Penerapan Algoritma

Hasil dari model Algoritma C.45 dianalisis untuk menentukan pemain yang layak masuk ke dalam Tim Nasional Spanyol. Pemain yang terpilih dibandingkan dengan daftar pemain yang sebelumnya dipanggil oleh pelatih untuk melihat tingkat kesesuaian model dengan seleksi manual. Jika model memiliki tingkat akurasi tinggi, metode ini dapat digunakan sebagai alat bantu seleksi pemain yang lebih objektif dan berbasis data.

Kesimpulan dari penelitian ini akan memberikan wawasan mengenai efektivitas metode Algoritma C.45 dalam seleksi pemain tengah dan bagaimana pendekatan berbasis data dapat membantu pelatih dalam pengambilan keputusan strategis untuk meningkatkan performa tim nasional.

Tabel 5. Data *Testing* dan Hasil Prediksi

No	Nama Pemain	Passing Acc.	Ball Rec.	Key Passes	xA	Dribbles	Minutes	Aktual	Prediksi
1	Rodri	90.1	7.1	4.9	0.62	21	2880	✓ Layak	✓ Layak
2	Pedri	88.7	6.2	4.7	0.58	27	2760	✓ Layak	✓ Layak
3	Fabian Ruiz	86.9	5.6	4.4	0.54	22	2710	✓ Layak	✓ Layak
4	Dani Olmo	83.2	4.9	4.3	0.50	20	2490	✓ Layak	✓ Layak
5	Zubimendi	81.2	4.4	4.4	0.47	35	2250	✓ Layak	✓ Layak
6	Mikel Merino	81.7	4.2	4.2	0.46	29	2200	✓ Layak	✓ Layak
7	Oyarzabal	81.2	3.9	6.2	0.27	38	2075	✓ Layak	✓ Layak
8	Koke	79.7	3.6	3.7	0.52	38	2075	✓ Layak	✓ Layak
9	Marc Bernal	78.2	1.9	3.2	0.14	35	2250	✗ Tidak	✗ Tidak
10	Naim Garcia	79.3	0.7	8.9	0.13	35	2250	✗ Tidak	✗ Tidak
11	Javi Perez	79.2	1.2	3.7	0.25	35	2250	✗ Tidak	✗ Tidak

12	Dani Rigo	79.2	0.4	7.5	0.14	35	2250	✗ Tidak	✗ Tidak
13	David G	79.1	0.7	8.1	0.18	35	2250	✗ Tidak	✗ Tidak
14	Jose Miguel	79.1	1	7.7	0.06	13	2110	✗ Tidak	✗ Tidak
...	...	...	...	...	...	...	...	...	...
1069	Aitor Sanz	79.1	0.5	7.9	0.13	35	2250	✗ Tidak	✗ Tidak

Tabel 6. Prediksi dan Aktual

	Prediksi	Aktual
Layak	8	8
Tidak Layak	7	7

True Positives (TP) = 8

True Negatives (TN) = 7

False Positives (FP) = 8

False Negatives (FN) = 7

Evaluasi Model

a. Akurasi

$$= (TP + TN) / \text{total}$$

$$= (8 + 7) / 15 = 15/15 = 100\%$$

b. Precision (Presisi Layak)

$$= TP / (TP + FP) = 8 / (8 + 0) = 1.0 = 100\%$$

c. Recall (Recall Layak)

$$= TP / (TP + FN) = 8 / (8 + 0) = 1.0 = 100\%$$

d. F1-Score (Layak)

$$= 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$$

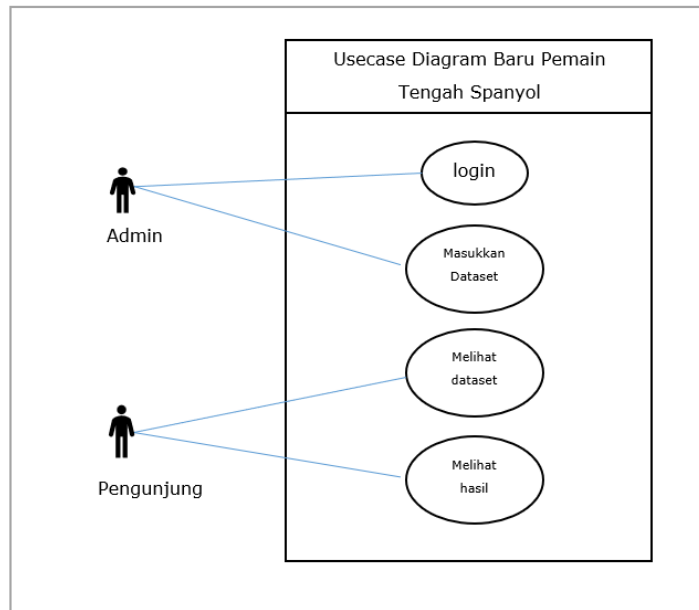
$$= 2 \times (1 \times 1) / (1 + 1) = 1.0 (100\%)$$

3. Hasil dan Pembahasan

3.1 Perancangan Sistem

a. *Use Case Diagram*

Use case diagram bekerja dengan cara mendeskripsikan tipikal interaksi antar pengguna dengan sistem melalui sebuah cerita yang menggambarkan bagaimana sebuah sistem dipakai. Diagram ini tidak hanya menunjukkan siapa saja aktor yang terlibat, tetapi juga menjelaskan alur kegiatan, tujuan yang ingin dicapai, serta hubungan antara pengguna dan fungsi-fungsi utama yang ada di dalam sistem, sehingga memudahkan dalam memahami kebutuhan serta ruang lingkup dari sistem yang akan dibangun.

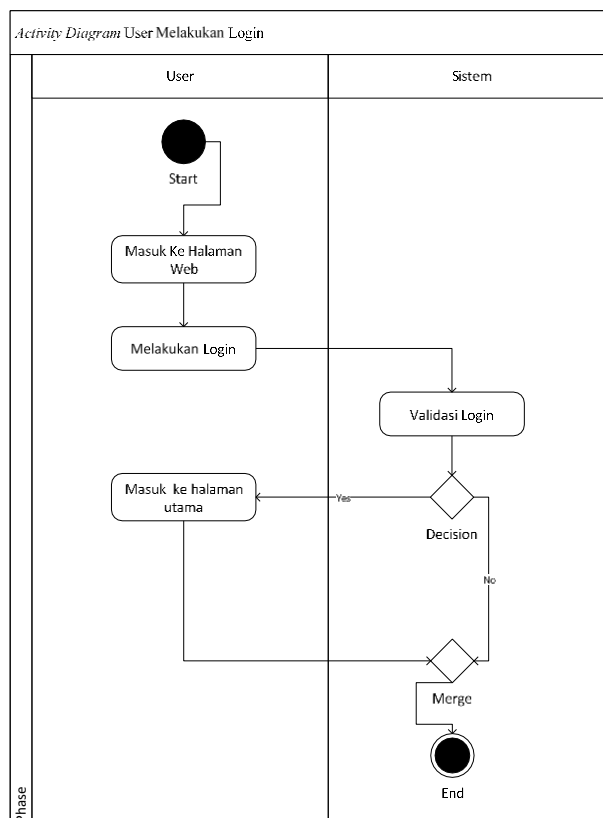


Gambar 2. Use Case Diagram

b. Activity Diagram

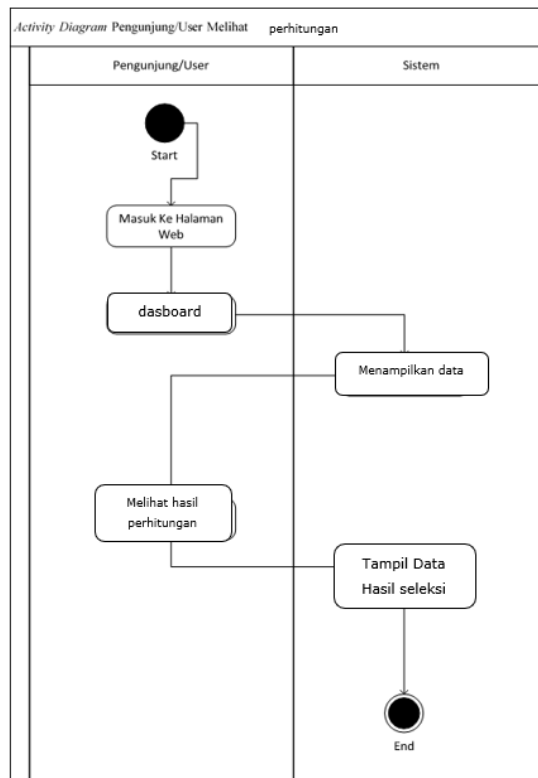
Activity diagram menggambarkan berbagai alur aktifitas dalam system yang sedang dirancang, bagaimana masing-masing alur berawal, decision yang mungkin terjadi dan bagaimana proses tersebut selesai. Berikut penggambaran activity diagram dari beberapa proses :

a. Activity Diagram Admin Melakukan Login



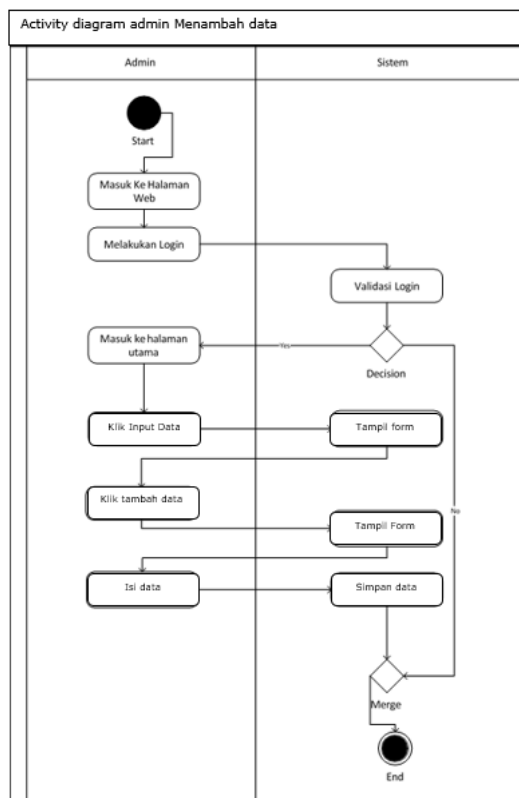
Gambar 3. Activity Diagram Login

b. *Activity Diagram* Pengunjung atau User Melihat perhitungan

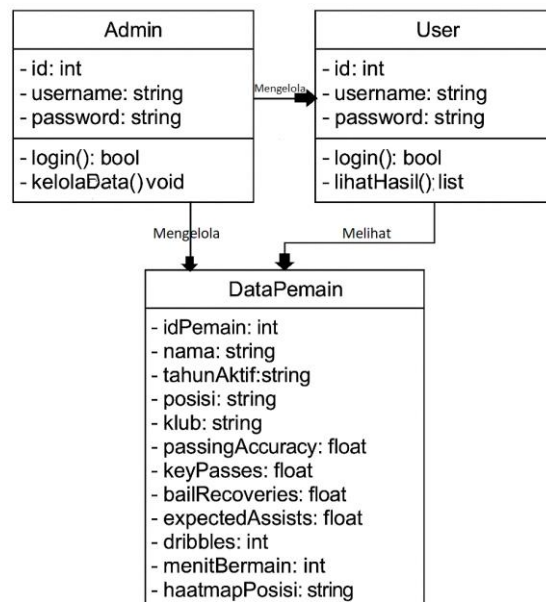
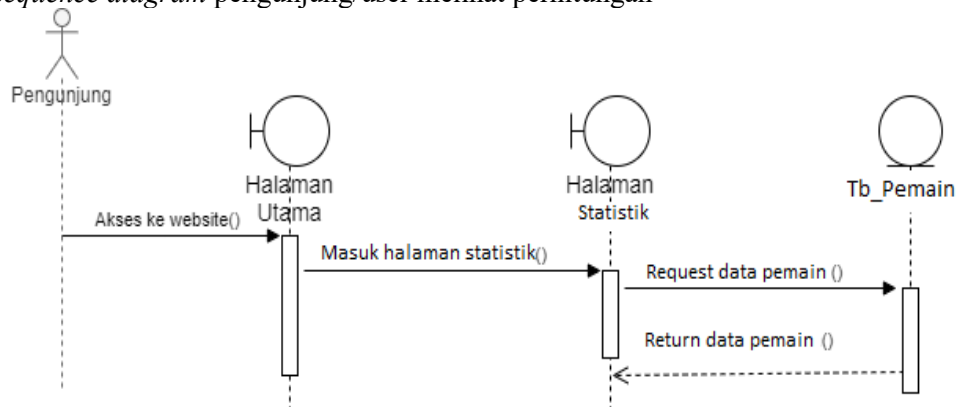
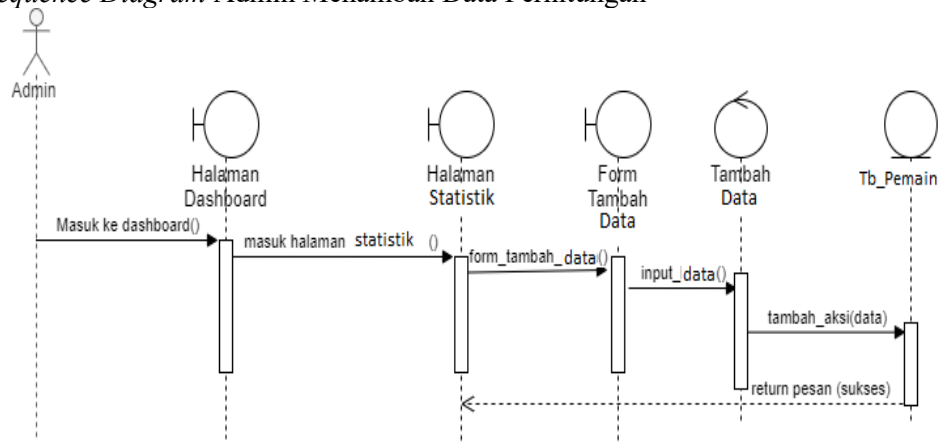


Gambar 4. *Activity Diagram* Melihat Perhitungan

c. *Activity Diagram* Admin Menambah data



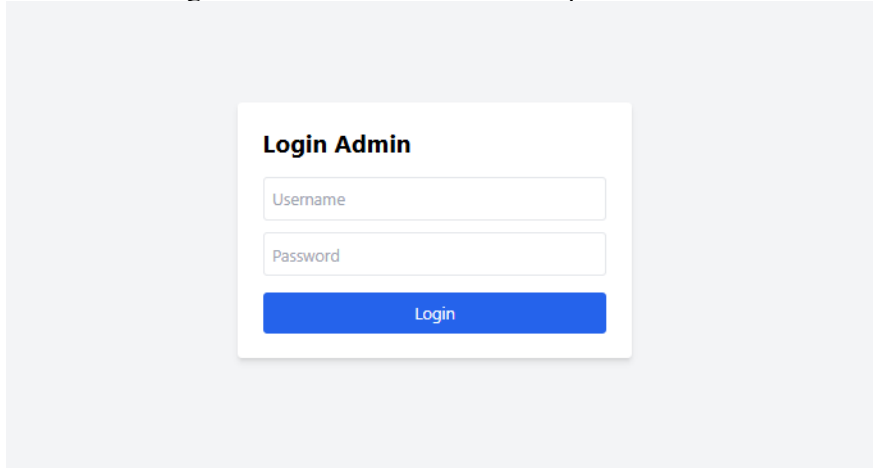
Gambar 5. *Activity Diagram* Menambah Data

c. *Class Diagram*Gambar 6. *Class Diagram*d. *Sequence Diagram*1. *Sequence diagram* pengunjung/user melihat perhitunganGambar 7. *Sequeunce Diagram* melihat perhitungan2. *Sequence Diagram* Admin Menambah Data PerhitunganGambar 8. *Sequeunce Diagram* menambah perhitungan

3.2 Implementasi Sistem

1. Login Admin

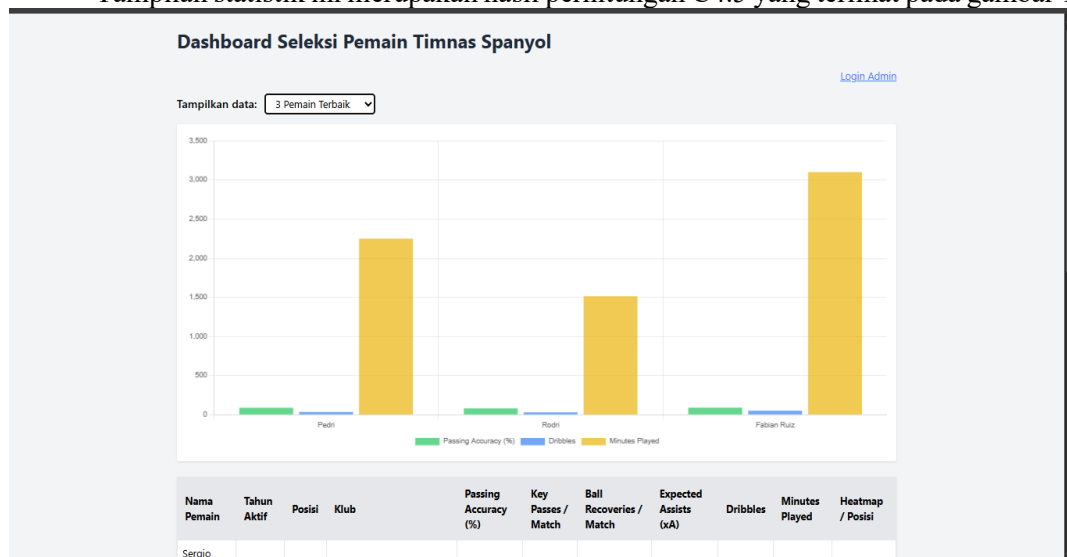
Gambar 9. menampilkan tampilan halaman Login Admin yang digunakan untuk mengakses sistem dengan memasukkan username dan password.



Gambar 9. *Login Admin*

2. Tampilan Statistik

Tampilan statistik ini merupakan hasil perhitungan C4.5 yang terlihat pada gambar 10.



Gambar 10. Tampilan Statistik

3. Tampilan Tambah Data

Tampilan Tambah data ini merupakan form input untuk mengisi, menambahkan data yang terlihat pada gambar 11.

Gambar 11. Tampilan Tambah Data

4. Tampilan Edit Data

Tampilan edit data ini merupakan form input untuk edit data yang sudah ada sebelumnya dan mau kita ubah yang terlihat pada gambar 12.

Gambar 12. Tampilan Edit Data

3.2 Pengujian Sistem

1. Pengujian *Black Box Testing* untuk Admin

Pengujian ini dilakukan untuk menguji fungsionalitas yang dapat diakses oleh admin pada sistem. Tabel berikut ini menunjukkan hasil pengujian dari berbagai fitur yang tersedia :

Tabel 1. Pengujian Admin

No	Fitur yang Diuji	Langkah Pengujian	Data Masukan	Hasil yang Diharapkan	Hasil Pengujian
1	Login Admin	Admin mengisi username dan password	Username & Password valid	Admin berhasil login	Berhasil

2	Input Data Pemain	Admin mengisi form input data pemain dan menyimpan	Data lengkap pemain	Data tersimpan dan muncul di tabel data pemain	Berhasil
3	Edit Data Pemain	Admin mengedit data pemain yang sudah ada	Data pemain yang ingin diubah	Data pemain berhasil diperbarui	Berhasil
4	Hapus Data Pemain	Admin memilih data dan menghapus pemain	Nama pemain	Data pemain berhasil dihapus	Berhasil
5	Proses Seleksi (Algoritma C.45)	Admin menekan tombol proses seleksi	-	Tiga pemain terbaik ditampilkan sesuai hasil C4.5	Berhasil
6	Logout	Admin menekan tombol logout	-	Kembali ke halaman login	Berhasil

2. Pengujian *Black Box Testing* untuk User

Pengujian ini dilakukan untuk menguji fungsionalitas yang dapat diakses oleh user pada sistem. Tabel berikut ini menunjukkan hasil pengujian dari berbagai fitur yang tersedia :

Tabel 2. Pengujian User

No	Fitur yang Diuji	Langkah Pengujian	Data Masukan	Hasil yang Diharapkan	Hasil Pengujian
1	Akses Halaman Utama	User mengakses sistem melalui browser	-	Halaman utama tampil dengan baik	Berhasil
2	Melihat Hasil Seleksi	User melihat daftar 3 pemain terbaik	-	Daftar pemain terbaik ditampilkan dengan detail	Berhasil

4. Kesimpulan

1. Penerapan algoritma C4.5 mampu memberikan rekomendasi yang akurat dalam menentukan kelayakan pemain tengah berdasarkan data performa aktual. Pohon keputusan yang dihasilkan dari proses klasifikasi menggunakan atribut-atribut penting seperti passing accuracy, ball recoveries per match, dan expected assists menunjukkan bahwa algoritma C4.5 dapat mengelompokkan pemain ke dalam kategori Layak dan Tidak Layak secara sistematis, logis, dan ber-basis data. Hal ini membuktikan bahwa pendekatan berbasis algoritma mampu mengurangi unsur subjektivitas dalam proses seleksi dan menghasilkan output yang dapat dipercaya.
2. Sistem pengambilan keputusan yang dibangun dengan algoritma C4.5 terbukti mampu melakukan klasifikasi kelayakan pemain tengah secara efektif dan akurat. Implementasi perhitungan manual seperti entropy, information gain, hingga gain ratio berhasil diaplikasikan dan dibandingkan dengan hasil dari program, dengan menunjukkan konsistensi yang valid. Sistem ini menyediakan dasar pertimbangan objektif bagi pelatih atau tim pemandu bakat dalam menentukan pemain yang sesuai dengan kebutuhan tim berdasarkan kriteria yang terukur. Dari hasil pengujian, kriteria utama kelayakan pemain tengah ditentukan pada passing accuracy $\geq 86\%$ dan ball recoveries per match ≥ 6 , yang disusun melalui analisis korelasi performa antar variabel.

Daftar Pustaka

- [1] I. I. J. Rifka Alkhilyatul Ma'rifat, I Made Suraharta, *Teknologi Metaverse Dalam Ilmu Keolahragaan*, vol. 2, 2024. [Online]. Available: <https://doi.org/10.30646/tikomsin.v1i1i2.727>
- [2] A. C. Padmasari, ““Enrique Kesulitan Susun Skuad Spanyol,”” Kompas.com. Accessed: Sep. 08, 2025. [Online]. Available: <https://bola.kompas.com/read/2021/06/10/20300008/luis-enrique-soal-situasi-kusut-timnas-spanyol-jelang-euro-2020--ini-bukan-apa?page=all>
- [3] R. H. Pambudi, B. D. Setiawan, and Indriati, “Penerapan Algoritma C4.5 Untuk Memprediksi Nilai Kelulusan Siswa Sekolah Menengah Berdasarkan Faktor Eksternal,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 7, pp. 2637–2643, 2018, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [4] Dhea Halimah, Muhammad Ridwan Lubis, and Widodo Saputra, “Algoritma C4.5 Untuk Menentukan Klasifikasi Tingkat Pemahaman Mahasiswa Pada Matakuliah Bahasa Pemrograman,” *J. Tek. Mesin, Ind. Elektro Dan Inform.*, vol. 1, no. 3, pp. 24–38, 2022, doi: 10.55606/jtmei.v1i3.534.
- [5] R. Irvan Effendi, D. Remawati, and S. H. Fitriasih, “Algoritma C4.5 Untuk Penentuan Tim Pemain Utama Olahraga Voli (Studi Kasus : DX Junior Karangturi),” *J. TIKomSiN*, vol. 11, no. 2, pp. 13–19, 2023, [Online]. Available: <https://doi.org/10.30646/tikomsin.v1i1i2.727>
- [6] S. O. Sinta Oktaviani S Minu, Tajuddin Abdillah, “Sistem Informasi Pengelompokan Pemain Sepakbola Menggunakan Data Mining Klasifikasi Pada PSSI Provinsi Gorontalo,” *J. Syst. Inf. Technol.*, vol. 1, no. 2, 2021.
- [7] Andi Diah Kuswanto, Hotman Nicolas Badjo, Septian Kharist, Muhammad Zayyid Mubarak, Riski Saputra, and Rivaldi Muhamad Fitroh, “Penerapan Algoritma C4.5 Dalam Klasifikasi Prestasi Atlet: Studi Kasus Pada Daftar Nama Penerima Penghargaan Tahun 2023,” *Bridg. J. Publ. Sist. Inf. dan Telekomun.*, vol. 2, no. 3, pp. 103–113, 2024, doi: 10.62951/bridge.v2i3.115.
- [8] C. I. Ifaru *et al.*, “Penentuan Sparepart Genset Paling Sering Digunakan Pada Operator Indosat Ooredoo Hutchison Dengan Algoritma C4.5 Determination of the most frequently used generator spare parts at Indosat Ooredoo Hutchison Operator with C4.5 Algorithm,” *Jtsi*, vol. 5, no. 1, pp. 300–309, 2024.
- [9] A. Dwi Putri, F. Sholekhah, E. Dadynata, L. Efrizoni, and N. Sapina, “The Application of C4.5 Decision Tree Algorithm for Predicting the Survival Rate of Thyroid Cancer Patients Penerapan Algoritma Decesion Tree C4.5 untuk Memprediksi Tingkat Kelangsungan Hidup Pasien Kanker Tiroid,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 4, pp. 1485–1495, 2024.
- [10] B. Alnur, Mulyono, Fitri Amillia, and S. Sutoyo, “JITE (Journal of Informatics andTelecommunication Engineering),” *J. Informatics Telecommun. Eng.*, vol. 7, no. 1, pp. 102–111, 2023, [Online]. Available: https://www.researchgate.net/publication/335117624_Malang_City_Polytechnic_Web_Based_Student_Attendance_Information_System_Telecommunications_Engineering_Study_Program_Using_Fingerprint/fulltext/5d515fe34585153e594ef214/Malang-City-Polytechnic-Web-Based-S
- [11] A. H. Arif Hidayat, M. R. S. Rizky, M. D. Aryasatya, N. D. Marco, and I. Amien, “Implementasi Algoritma C4.5 Untuk Klasifikasi Tingkat Kerawanan Kekeringan di Kabupaten Cilacap,” *J. Komput. Antart.*, vol. 3, no. 2, pp. 71–78, 2025, doi: 10.70052/jka.v3i2.664.