

Penerapan Algoritma Id3 Dalam Klasifikasi Kelayakan Penerima Bantuan Langsung Tunai (BLT) Studi Kasus: Desa Somi Kecamatan Gido

Iren Priscilla Bu'ulolo¹, Samuel V.B. Manurung², Arina Prima Silalahi³

^{1,2,3}Fakultas Ilmu Komputer, Universitas Methodist Indonesia

Info Artikel

Histori Artikel:

Received, July 28, 2025

Revised, Sept 20, 2025

Accepted, Nov 12, 2025

Keywords:

BLT, Algoritma ID3, klasifikasi, pohon keputusan, kelayakan penerima.

ABSTRAK

Penyaluran Bantuan Langsung Tunai (BLT) di Desa Somi masih dilakukan secara manual, sehingga rawan subjektivitas dan ketidaktepatan sasaran. Penelitian ini bertujuan untuk menerapkan algoritma ID3 dalam mengklasifikasikan kelayakan penerima BLT secara objektif berdasarkan data sosial ekonomi masyarakat. Sebanyak 333 data digunakan sebagai dataset yang dianalisis menggunakan metode ID3 dengan perhitungan entropy dan information gain untuk membentuk pohon keputusan. Model diuji menggunakan 67 data uji dan menghasilkan akurasi sebesar 95,52%, dengan precision 96%, recall 93,1%, dan F1-score 94,5%. Hasil penelitian menunjukkan bahwa algoritma ID3 efektif dalam membantu proses seleksi penerima BLT secara sistematis dan akurat.

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



Penulis Koresponden:

Iren Priscilla Bu'ulolo,
Fakultas Ilmu Komputer,
Universitas Methodist Indonesia, Medan,
Jl. Hang Tuah No.8, Medan - Sumatera Utara.
Email: irenpricilabuulolo@gmail.com

1. PENDAHULUAN

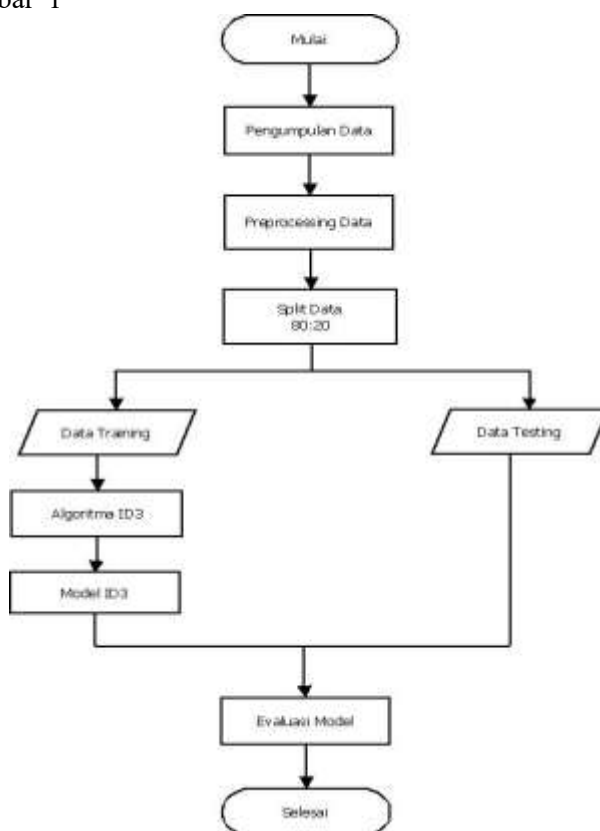
Bantuan Langsung Tunai (BLT) merupakan program bantuan sosial krusial yang diluncurkan oleh pemerintah untuk membantu masyarakat yang mengalami kesulitan ekonomi. Di Desa Somi, Kecamatan Gido, program ini memiliki peran vital, terutama pasca-pandemi COVID-19 yang memicu krisis ekonomi. Namun, proses penentuan penerima BLT di desa ini masih menghadapi tantangan serius. Keputusan yang ada selama ini cenderung bersifat manual dan subjektif, sangat bergantung pada pandangan individu atau pihak berwenang tanpa adanya dasar perhitungan yang objektif dan konsisten. Kondisi ini sering kali berujung pada ketidaktepatan dalam penyaluran bantuan, di mana mereka yang seharusnya layak tidak menerimanya, sementara yang tidak memenuhi kriteria malah mendapatkan bantuan. Untuk mengatasi masalah ini, diperlukan sistem klasifikasi yang dapat menentukan kelayakan penerima BLT secara lebih akurat dan efisien. Salah satu metode yang relevan untuk tujuan ini adalah algoritma ID3 (Iterative Dichotomiser 3). Algoritma ini, sebagai bagian dari cabang machine learning dan kecerdasan buatan, berfokus pada pembangunan pohon keputusan dengan memilih atribut paling informatif untuk mengklasifikasi data. Berdasarkan penelitian sebelumnya [1], [2], penggunaan algoritma ini dapat membantu mengatasi ketidaktepatan dan inefisiensi waktu dalam proses penyaluran bantuan. Semakin baik algoritma machine learning

yang digunakan, semakin baik pula keputusan yang dihasilkan[3]. Oleh karena itu, penelitian ini bertujuan untuk mengimplementasikan algoritma ID3 untuk mengklasifikasi kelayakan penerima BLT di Desa Somi berdasarkan kriteria yang telah ditentukan, serta mengevaluasi tingkat akurasi. Penelitian ini memiliki batasan yang jelas agar lebih terarah, yaitu berfokus pada penerima BLT di Desa Somi dengan menggunakan 333 data dan 14 atribut yang diperoleh dari Kepala Desa. Metode yang diterapkan secara khusus adalah algoritma ID3 dalam bentuk pohon keputusan (*decision tree*), dan evaluasi akurasi dilakukan menggunakan Confusion Matrix dengan implementasi menggunakan bahasa pemrograman Python. Melalui penelitian ini, diharapkan dapat dihasilkan solusi berbasis algoritma ID3 yang mampu meningkatkan ketepatan dan objektivitas dalam proses klasifikasi kelayakan penerima BLT. Hal ini tidak hanya akan menjadikan penyaluran bantuan lebih adil dan sesuai dengan kriteria yang ditetapkan, tetapi juga memberikan gambaran yang jelas mengenai efektivitas algoritma ID3 sebagai dasar pengembangan sistem pendukung keputusan di masa mendatang agar hasil seleksi semakin optimal.

2. METODE PENELITIAN

2.1. Framework Penelitian

Framework merupakan kerangka kerja yang terdiri dari berbagai komponen, aturan, dan standar yang dirancang untuk mempermudah proses pengembangan aplikasi atau sistem. Dengan menggunakan *framework*, pengembang dapat bekerja dengan lebih cepat, terstruktur, dan efisien. *Framework* memberikan fondasi yang diperlukan, sehingga pengembang tidak perlu memulai dari nol, melainkan dapat mengikuti pola-pola yang telah disediakan. Adapun *framework* pada penelitian ini dapat dilihat pada Gambar 1



Gambar 1. *Framework* Penelitian

2.2. Teknik Pengumpulan Data

Data Bantuan Langsung Tunai (BLT) yang digunakan dalam penelitian ini diambil dari Desa Somi, Kecamatan Gido, dengan total sebanyak 333 data yang terdiri dari 14 atribut dan merupakan data kategori.

2.3. Preprocessing Text

Preprocessing/cleaning merupakan proses yang dilakukan dengan cara memeriksa kembali data antara lain untuk pembersihan data-data redundansi, outlier, dan bernilai null (data yang kosong)[3][4]. Tahapan preprocessing data yang akan dijelaskan meliputi[4]

1. *Cleaning* data: proses pembersihan dan perbaikan dataset dengan tujuan untuk meningkatkan kualitas data sehingga siap untuk dianalisis. *Cleaning* data set mendapat 14 atribut dan 2 label dengan jumlah 333 data.
 - a. Mengubah *symbol* V menjadi “*TRUE*” dan format *symbol* X menjadi “*FALSE*”.
2. *Encoding*: Dalam analisis data, agar data dapat diproses dengan algoritma pembelajaran mesin data kategorik harus diubah menjadi data numerik. Proses mengubah data kategorik menjadi data numerik disebut *encoding*.
 - a. Mengubah nilai “*TRUE*” menjadi 1 dan “*FALSE*” menjadi 0.

2.4. Split Data 80:20

Pembagian data merupakan langkah penting dalam pengujian model *machine learning*. Dalam penelitian ini, jumlah data yang digunakan untuk klasifikasi kelayakan penerima BLT adalah sebanyak 333 data, yang dibagi dengan skema 80:20, yaitu 266 data untuk pelatihan dan 67 data untuk pengujian. Setelah data dibagi, langkah berikutnya adalah melatih model menggunakan algoritma *ID3*. Model dilatih dengan memproses data training untuk membentuk pohon keputusan (*decision tree*) berdasarkan atribut-atribut yang ada. Data latih digunakan untuk membangun model, sementara data uji menilai akurasi prediksi terhadap data baru. Tujuannya adalah menghindari *overfitting* dan memastikan model dapat melakukan generalisasi dengan baik.

2.4. Evaluasi Model

Evaluasi model yang digunakan dalam penelitian ini adalah confusion matrix, *Confusion matrix* adalah tabel yang menyatakan klasifikasi jumlah data uji yang benar dan jumlah data uji yang salah[5]. Pada pengukuran kinerja menggunakan *Confusion Matrix*, terdapat 4 (empat) istilah sebagai representasi hasil proses klasifikasi. Keempat istilah tersebut adalah *True Positive* (TP), *True Negative* (TN), *False Positive* (FP) dan *False Negative* (FN). Dimana Nilai *True Negative* (TN) merupakan jumlah data negative yang terdeteksi dengan benar, sedangkan *False Positive* (FP) merupakan data negatif namun terdeteksi sebagai data positif. Sementara itu, *True Positive* (TP) merupakan data positif yang terdeteksi benar. *False Negative* (FN) merupakan kebalikan dari *True Positive*, sehingga data positif, namun terdeteksi sebagai data negative[6]. Dari *confusion matrix* tersebut dapat dihitung beberapa metrik evaluasi sebagai berikut:

$$accuracy = \frac{\text{Jumlah Prediksi Benar}}{\text{Jumlah Keseluruhan data}}$$

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

2.5 Algoritma Iterative Dichomiser 3 (ID3)

Algoritma *ID3* merupakan algoritma yang digunakan untuk membangun sebuah *decision tree* atau pohon keputusan. *Decision tree* menggunakan struktur hierarki untuk *supervised learning*[7], [8].

Proses dari *decision tree* dimulai dari *root Node* hingga *leaf Node* yang dilakukan secara rekursif. Dimana setiap pencabangan menyatakan suatu kondisi yang harus terpenuhi di setiap ujung pohon menyatakan kelas dari suatu data. Proses dari *decision tree* ini mengubah bentuk data (tabel) menjadi sebuah pohon (*tree*) kemudian mengubahnya menjadi suatu

aturan (*rule*)[7]. Alur kerja algoritma ID3 dalam membangun pohon keputusan dimulai dengan tahap pengumpulan atribut. Setelah itu, dilakukan perhitungan *entropy* total dan *Gain* untuk masing-masing atribut. Untuk menentukan *entropy* dan *Gain*, digunakan rumus sebagai berikut[8].

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

Keterangan:

$Entropy(S)$ = Tingkat ketidakpastian dalam *dataset S*

n = Jumlah total kelas dalam *dataset*

p_i = Proporsi data pada kelas i

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

Keterangan:

$Entropy(S)$ = Tingkat ketidakpastian dalam *dataset S*

n = Jumlah total kelas dalam *dataset*

p_i = Proporsi data pada kelas i

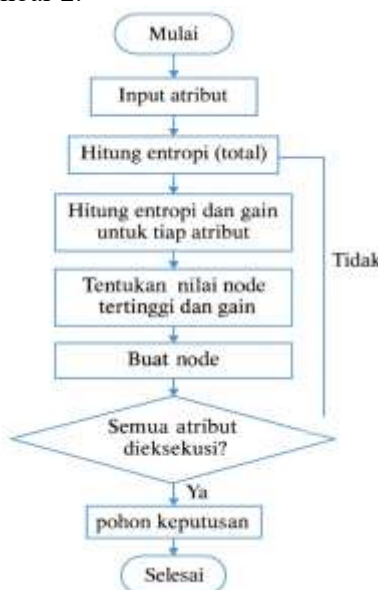
$Gain(S, A)$ = Pengurangan *entropy* setelah *dataset S* dibagi oleh atribut A

n = Jumlah *subset* yang dihasilkan oleh atribut A

$|S_i|$ = Jumlah data dalam *subset S_i*

$|S|$ = Jumlah total data dalam *dataset S*

Proses ini akan diulang hingga semua atribut telah digunakan atau data sudah terklasifikasi sepenuhnya. Hasil akhirnya adalah pohon keputusan yang siap digunakan untuk proses klasifikasi. Flowchart ID3 dapat dilihat pada Gambar 2.



Gambar 2. Flowchart ID3

3. HASIL DAN PEMBAHASAN

3.1. Implementasi Algoritma ID3 Berbasis Python

Dalam proses pengerjaan penelitian ini membangun pohon keputusan dengan Python, khususnya menggunakan pustaka scikit-learn, melibatkan beberapa tahapan utama: pertama, menyiapkan data yang akan digunakan; kedua, menghitung entropi dan gain untuk memilih atribut terbaik; ketiga, melatih model menggunakan algoritma Decision Tree; dan terakhir, memvisualisasikan pohon yang telah terbentuk agar mudah diinterpretasi.

3.1.1 Preprocessing

1. Cleaning data

Proses *cleaning* data dilakukan terhadap semua data dengan cara mengubah simbol “V” menjadi *TRUE* dan simbol “X” menjadi *FALSE*. Langkah ini dilakukan untuk memastikan bahwa seluruh data berada dalam format yang seragam dan siap untuk dianalisis secara logis

nama	alamat	jenis kelamin	umur	pendidikan	jenis pekerjaan
1	...	...	...	...	...
2	...	...	...	...	...
3	...	...	...	...	...
4	...	...	...	...	...
5	...	...	...	...	...
6	...	...	...	...	...
7	...	...	...	...	...
8	...	...	...	...	...
9	...	...	...	...	...
10	...	...	...	...	...
11	...	...	...	...	...
12	...	...	...	...	...
13	...	...	...	...	...
14	...	...	...	...	...
15	...	...	...	...	...
16	...	...	...	...	...
17	...	...	...	...	...
18	...	...	...	...	...
19	...	...	...	...	...
20	...	...	...	...	...
21	...	...	...	...	...
22	...	...	...	...	...
23	...	...	...	...	...
24	...	...	...	...	...
25	...	...	...	...	...
26	...	...	...	...	...
27	...	...	...	...	...
28	...	...	...	...	...
29	...	...	...	...	...
30	...	...	...	...	...
31	...	...	...	...	...
32	...	...	...	...	...
33	...	...	...	...	...
34	...	...	...	...	...
35	...	...	...	...	...
36	...	...	...	...	...
37	...	...	...	...	...
38	...	...	...	...	...
39	...	...	...	...	...
40	...	...	...	...	...
41	...	...	...	...	...
42	...	...	...	...	...
43	...	...	...	...	...
44	...	...	...	...	...
45	...	...	...	...	...
46	...	...	...	...	...
47	...	...	...	...	...
48	...	...	...	...	...
49	...	...	...	...	...
50	...	...	...	...	...
51	...	...	...	...	...
52	...	...	...	...	...
53	...	...	...	...	...
54	...	...	...	...	...
55	...	...	...	...	...
56	...	...	...	...	...
57	...	...	...	...	...
58	...	...	...	...	...
59	...	...	...	...	...
60	...	...	...	...	...
61	...	...	...	...	...
62	...	...	...	...	...
63	...	...	...	...	...
64	...	...	...	...	...
65	...	...	...	...	...
66	...	...	...	...	...
67	...	...	...	...	...
68	...	...	...	...	...
69	...	...	...	...	...
70	...	...	...	...	...
71	...	...	...	...	...
72	...	...	...	...	...
73	...	...	...	...	...
74	...	...	...	...	...
75	...	...	...	...	...
76	...	...	...	...	...
77	...	...	...	...	...
78	...	...	...	...	...
79	...	...	...	...	...
80	...	...	...	...	...
81	...	...	...	...	...
82	...	...	...	...	...
83	...	...	...	...	...
84	...	...	...	...	...
85	...	...	...	...	...
86	...	...	...	...	...
87	...	...	...	...	...
88	...	...	...	...	...
89	...	...	...	...	...
90	...	...	...	...</	

[illegible]

2. Encoding

Tahap ini merupakan proses pemberian nilai terhadap label dari masing-masing atribut

dalam dataset. Pada tahap ini, setiap nilai kategorikal dikonversi menjadi bentuk numerik melalui teknik *one-hot encoding*. Teknik ini bekerja dengan cara membuat kolom-kolom biner baru yang mewakili setiap kategori sehingga setiap data akan diberikan nilai 0 atau 1, tergantung pada keberadaan kategori tersebut dalam data. Proses *one-hot encoding* sangat penting karena membantu algoritma *machine learning* dalam mengolah data kategori secara efektif tanpa mengasumsikan adanya hubungan ordinal antar kategori. Hasil penerapan teknik ini pada dataset dapat dilihat pada Gambar 4.

[illegible]

Split data 80:20 adalah proses pembagian dataset menjadi dua bagian, yakni 80% untuk pelatihan

(*training*) dan 20% untuk pengujian (*testing*). Data pelatihan digunakan untuk membangun model, sedangkan data pengujian berfungsi mengevaluasi performa model pada data yang belum pernah dilihat sebelumnya. Proses ini membantu memastikan model dapat belajar dengan cukup data sekaligus diuji secara objektif. Pada Gambar 5.

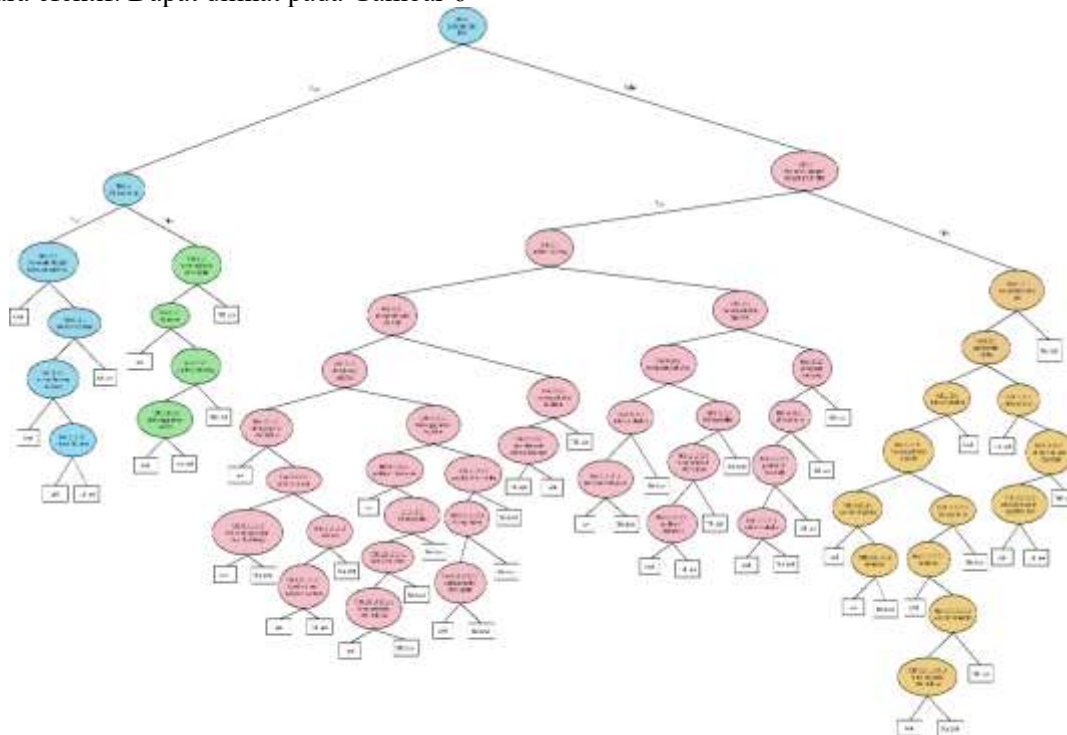
```
# Split data: 88% training, 12% testing
X_train, X_test, y_train, y_test = train_test_split(X_noisy, y, test_size=0.2, random_state=42)
```

Gambar 5. Hasil Split data

Proses pelatihan dilakukan menggunakan algoritma ID3 yang diimplementasikan melalui

fungsi *Decision tree Classifier* pada library *scikit-learn*. Model ini dibangun dengan membentuk

tree, di mana setiap pohon dilatih menggunakan kriteria *Gain* untuk menentukan fitur terbaik dalam melakukan pemisahan data. Yang menampilkan struktur pohon dan cara model melakukan klasifikasi secara efektif. Dapat dilihat pada Gambar 6

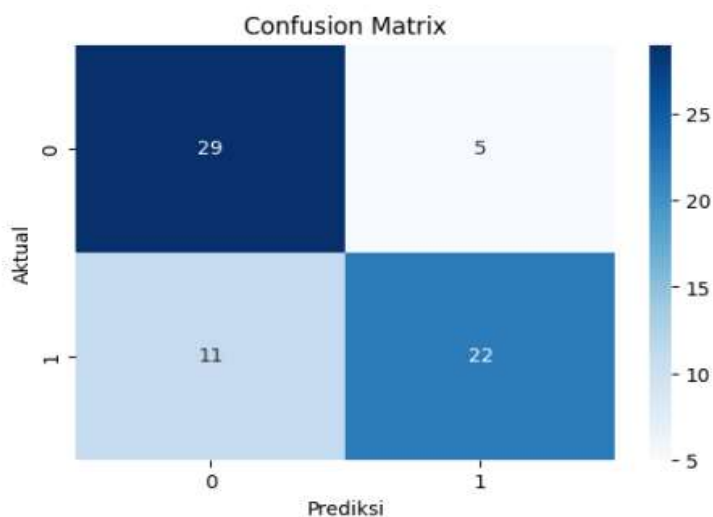


Gambar 6. *Traning ID3*

Sumber : <https://acesse.one/pohonkeputusan>

3.1.4 Evaluasi Model

Proses evaluasi model dilakukan dengan menguji performa *model algoritma* ID3 menggunakan data uji. Hasil evaluasi dengan *confusion matrix* dapat dilihat pada gambar 7.



Gambar 7. *Confusion Matrix*

Berdasarkan *confusion matrix*, hasil perhitungan nilai *accuracy*, *precision*, *recall*, dan *f1-score* dapat dilihat pada gambar 8.

Evaluasi Model:					
Akurasi	:	0.7612			
Presisi	:	0.7692			
Recall	:	0.7612			
F1-Score	:	0.7589			
Classification Report:					
		precision	recall	f1-score	support
0		0.72	0.85	0.78	34
1		0.81	0.67	0.73	33
accuracy				0.76	67
macro avg		0.77	0.76	0.76	67
weighted avg		0.77	0.76	0.76	67

Gambar 8. Hasil Evaluasi

3. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan Penerapan algoritma ID3 dalam klasifikasi kelayakan penerima BLT di Desa Somi telah berhasil dilakukan. Algoritma ini mampu membentuk pohon keputusan berdasarkan atribut sosial ekonomi yang relevan untuk menentukan status kelayakan. Dan hasil evaluasi model pada data uji menunjukkan bahwa algoritma ID3 memiliki kinerja yang cukup baik, dengan akurasi sebesar 79,10%. Nilai precision, recall, dan f1-score yang tinggi pada masing-masing kelas menunjukkan bahwa model ini cukup efektif dalam mengidentifikasi status kelayakan penerima bantuan secara tepat.

REFERENSI

- [1] S. Widaningsih, "Penerapan Data Mining untuk Memprediksi Siswa Berprestasi dengan Menggunakan Algoritma K Nearest Neighbor," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 9, no. 3, pp. 2598–2611, 2022, doi: 10.35957/jatisi.v9i3.859.
- [2] A. P. Silalahi and H. G. Simanullang, *DECISSION TREE DAN PENERAPANNYA*. 2025.
- [3] A. Agung, A. Daniswara, I. Kadek, and D. Nuryana, "Data Preprocessing Pola Pada Penilaian Mahasiswa Program Profesi Guru," *J. Informatics Comput. Sci.*, vol. 05, pp. 97–100, 2023.
- [4] H. G. Simanullang, S. Wijono, B. Kristianto, W. Sulisty, and A. P. Silalahi, "Detecting Cyberbullying Using Machine Learning Approaches," in *2024 Ninth International Conference on Informatics and Computing (ICIC)*, 2024, pp. 1–8. doi: 10.1109/ICIC64337.2024.10956842.
- [5] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," *J. Sains Komput. Inform. (J-SAKTI)*, vol. 5, no. 2, pp. 697–711, 2021, [Online]. Available: <https://ejurnal.tunasbangsa.ac.id/index.php/jsakti/article/view/369/348>
- [6] E. Fauziningrum, M. Pd and M. P. Encis Indah Suryaningsih, S.T., "Evaluasi Dan Prediksi Penguasaan Bahasa Inggris Maritim Menggunakan Metode Decision Tree Dan Confusion Matrix (Studi Kasus Di Universitas Maritim Amni)," *Angew. Chemie Int. Ed. 6(11)*, 951–952., pp. 5–24, 2021.
- [7] H. G. Simanullang, A. P. Silalahi, and M. I. Hutapea, *Pendukung Keputusan Seleksi Indeks LQ-45*. Malang: Madza Media, 2023.
- [8] A. P. Silalahi and H. G. Simanullang, *Decission Tree dan Penerapannya dengan Algoritma Iterative Dichotomizes versi 3 (ID3)*. 2025.
- [9] J. V. Reonaldho, D. Saepudin, and D. Adytia, "Prediksi Gelombang Ekstrim Air Laut di Pelabuhan Tanjung Priok Menggunakan Algoritma ID3," *e-Proceeding Eng.*, vol. 7, no. 1, pp. 2700–2713, 2020.
- [10] A. Sifaunajah and R. D. Wahyuningtyas, "Penggunaan Algoritma ID3 Untuk Klasifikasi Data Calon Peserta Didik," *CSRID (Computer Sci. Res. Its Dev. Journal)*, vol. 14, no. 2, p. 103, 2022, doi: 10.22303/csr.14.2.2022.103-112.