Implementasi Data Mining untuk Klasifikasi Pengaduan PMI Menggunakan Decision Tree

Muhammad Ali Ridla[1], Sofiatun Mukarroma[2], Nargis Yahya*[3]
Program Studi Sistem Informasi, Fakultas Sains dan Teknologi Komputer, Universitas Ibrahimy Situbondo
e-mail: [1]el.riedla@gmail.com, [2]sofiatulmukarromah085@gmail.com, [3]nargisyahya09@gmail.com
*Penulis Korespondensi

Diterima: 16 Juni 2026; Revisi ke-1: 2 Juli 2026; Revisi ke-2: 24 September 2026; Disetujui: 27 September 2026

Abstrak

Meningkatnya jumlah data pengaduan Pekerja Migran Indonesia (PMI) menyebabkan proses pengelolaan dan pengelompokan data menjadi semakin kompleks. Pengolahan data secara manual dinilai kurang efektif karena memerlukan waktu yang lebih lama dalam proses identifikasi dan tindak lanjut kasus. enelitian ini bertujuan menerapkan teknik data mining *menggunakan algoritma* Decision Tree *untuk klasifikasi data pengaduan PMI. Dataset yang digunakan berjumlah 130 data pengaduan yang telah melalui tahap* preprocessing. *Proses klasifikasi dilakukan menggunakan RapidMiner, sedangkan evaluasi model menggunakan* confusion matrix *berdasarkan nilai* accuracy, precision, *dan* recall. *Hasil penelitian menunjukkan bahwa parameter terbaik diperoleh menggunakan* criterion information gain, maximal depth *3, dan* minimal leaf size *1 dengan nilai* accuracy *sebesar 86,92%,* precision *sebesar 81,94% (BELUM), 93,10% (SELESAI) dan* recall *sebesar 93,65 (BELUM), 80,60% (SELESAI). Hasil tersebut menunjukkan bahwa algoritma* Decision Tree *mampu mengklasifikasikan data pengaduan PMI dengan baik sehingga dapat membantu proses pengelompokan data secara lebih cepat, terstruktur, dan efisien.*

Kata kunci: *data mining, klasifikasi, pengaduan pmi, decision tree.*

Abstract

The increasing number of Indonesian Migrant Worker (PMI) complaint records has made the process of data management and classification increasingly complex. Manual data processing is considered less effective because it requires more time for case identification and follow-up. This study aims to implement data mining techniques using the Decision Tree algorithm to classify PMI complaint data. The dataset used consists of 130 complaint records that have undergone the preprocessing stage. The classification process was carried out using RapidMiner, while model performance was evaluated using a confusion matrix based on accuracy, precision, and recall. The results show that the best model was obtained using the information gain criterion, a maximal depth of 3, and a minimal leaf size of 1, achieving an accuracy of 86,92%, precision of 81,94% for the PENDING class and 93,10% for the COMPLETED class, and recall of 93,65% for the PENDING class and 80,60% for the COMPLETED class. These findings indicate that the Decision Tree algorithm is capable of accurately classifying PMI complaint data, thereby supporting a faster, more structured, and more efficient complaint classification process.

Keywords: *data mining, classification, pmi complaints, decision tree*

I. PENDAHULUAN

Perkembangan teknologi informasi memberikan dampak yang signifikan terhadap berbagai aspek kehidupan, termasuk dalam pengelolaan informasi dan pelayanan publik. Salah satu sektor yang menghasilkan data dalam jumlah besar adalah layanan pengaduan Pekerja Migran Indonesia (PMI). Pekerja Migran Indonesia merupakan warga negara Indonesia yang bekerja di luar negeri dengan memperoleh upah di luar wilayah Indonesia serta memiliki hak untuk memperoleh perlindungan sebelum, selama, dan setelah masa penempatan sebagaimana diatur dalam Undang-Undang Nomor 18 Tahun 2017 [1]. Meningkatnya jumlah

This is an open access article under the CC–BY-SA license.

PMI beserta kompleksitas permasalahan yang dihadapi menyebabkan volume data pengaduan terus bertambah, sehingga dibutuhkan pengelolaan data yang lebih efektif dan efisien untuk mendukung proses pelayanan dan pengambilan keputusan. Data pengaduan PMI mencakup berbagai jenis permasalahan, seperti dokumen, upah, komunikasi, kekerasan, penempatan kerja, hingga persoalan hukum. Besarnya jumlah data yang tersimpan sering kali hanya dimanfaatkan sebagai arsip sehingga informasi yang terkandung didalamnya belum dimanfaatkan secara optimal. Kondisi tersebut menyebabkan proses identifikasi pola permasalahan dan pengambilan keputusan menjadi kurang efektif. Oleh karena itu, diperlukan suatu pendekatan yang mampu mengolah data pengaduan menjadi informasi yang lebih bernilai guna mendukung identifikasi pola serta pengelompokan kasus secara sistematis [2].

Dalam penerapannya, data mining memiliki berbagai teknik analisis, salah satunya adalah klasifikasi. Teknik klasifikasi digunakan untuk mengelompokkan data ke dalam kategori tertentu berdasarkan karakteristik yang dimiliki sehingga memungkinkan proses identifikasi dan prediksi terhadap suatu kelas data. Metode yang digunakan dalam penelitian ini adalah *Decision Tree*. Pemilihan metode tersebut didasarkan pada kemampuannya dalam menghasilkan model klasifikasi yang mudah dipahami, memiliki proses pengambilan keputusan yang jelas, serta mampu menghasilkan aturan klasifikasi dalam bentuk pohon keputusan yang mudah diinterpretasikan[3]. Selain itu, metode *Decision Tree* juga mampu mengidentifikasi atribut yang paling berpengaruh terhadap hasil klasifikasi sehingga sesuai digunakan pada data pengaduan PMI yang didominasi atribut kategorikal.

Hasil penelitian menunjukkan bahwa model Decision Tree dengan criterion information gain, maximal depth 3, dan minimal leaf size 1 menghasilkan accuracy sebesar 86,92%. Nilai tersebut menunjukkan bahwa model mampu mengklasifikasikan sebagian besar data pengaduan PMI dengan tepat. Hasil ini kemudian dibandingkan dengan beberapa penelitian terdahulu yang menerapkan algoritma Decision Tree pada permasalahan klasifikasi yang berbeda. Penelitian yang dilakukan oleh Sheila menerapkan algoritma C4.5 menggunakan RapidMiner untuk melakukan klasifikasi penyakit diabetes dan menghasilkan *accuracy* terbaik sebesar 86,05%[4]. Selain itu, penelitian mengenai klasifikasi tingkat keparahan penyakit Demam Berdarah Dengue (DBD) dengan algoritma Decision Tree berbasis RapidMiner memperoleh *accuracy* sebesar 85%, dengan *precision* sebesar 84% dan *recall* sebesar 83%[5]. Berdasarkan kedua penelitian tersebut, nilai *accuracy* yang diperoleh pada penelitian ini, yaitu 86,92%, memiliki selisih sebesar 0,87 poin persentase dibandingkan penelitian klasifikasi diabetes dan 1,92 poin persentase dibandingkan penelitian DBD. Perbedaan hasil tersebut menunjukkan bahwa performa algoritma Decision Tree dapat dipengaruhi oleh karakteristik data, atribut yang digunakan, jumlah dataset, tahapan preprocessing, serta konfigurasi parameter pada proses pemodelan.

Perbandingan juga dapat dilakukan dengan penelitian Wardani et al. yang menerapkan algoritma Decision Tree C4.5 untuk klasifikasi penjualan barang terlaris menggunakan RapidMiner. Penelitian tersebut menghasilkan *accuracy* sebesar 90,85% pada pemodelan awal dan meningkat menjadi 95,68% setelah dilakukan proses *tuning* parameter [6] Penelitian lainnya mengenai klasifikasi kenaikan jabatan karyawan menggunakan algoritma C4.5 berbasis RapidMiner memperoleh *accuracy* sebesar 88,91%[7]. Jika dibandingkan dengan hasil penelitian tersebut, *accuracy* sebesar 86,92% yang diperoleh dalam penelitian ini masih menunjukkan nilai yang lebih rendah. Namun, perbedaan nilai *accuracy* tersebut tidak dapat dijadikan dasar untuk menentukan bahwa suatu model memiliki performa yang secara mutlak lebih baik, karena setiap penelitian menggunakan karakteristik dataset, jumlah data, atribut, pembagian dataset, serta skenario pengujian yang berbeda.

Berdasarkan hasil kajian terhadap penelitian-penelitian terdahulu, penerapan algoritma Decision Tree telah banyak dilakukan pada berbagai bidang, seperti klasifikasi penyakit, prediksi kemiskinan, klasifikasi pendidikan, dan klasifikasi kualitas udara. Namun, penerapan algoritma tersebut pada klasifikasi data pengaduan Pekerja Migran Indonesia (PMI) masih terbatas, khususnya penelitian yang melakukan pengujian beberapa kombinasi parameter criterion, maximal depth, dan minimal leaf size untuk memperoleh konfigurasi model yang sesuai dengan karakteristik data pengaduan [8].

Oleh karena itu, kebaruan penelitian ini terletak pada penerapan optimasi parameter Decision Tree untuk memperoleh konfigurasi model yang sesuai dalam mengklasifikasikan status pengaduan PMI. Penelitian ini tidak hanya mengukur accuracy model, tetapi juga mengevaluasi precision dan recall melalui confusion matrix serta menganalisis struktur pohon keputusan yang dihasilkan. Pendekatan tersebut diharapkan dapat menghasilkan model klasifikasi yang mudah diinterpretasikan dan dapat mendukung proses pengelompokan status pengaduan PMI secara lebih terstruktur [9].

II. METODE PENELITIAN

Adapun Metode penelitian yang digunakan pada penelitian ini terdiri dari beberapa tahapan yang bertujuan untuk memperoleh model klasifikasi terbaik pada data pengaduan Pekerja Migran Indonesia (PMI). Alur penelitian yang digunakan dapat dilihat pada Gambar 1:



Gambar 1. Metode penelitian yang dilakukan

A. Pengumpulan Data

Pada tahap ini data penelitian diperoleh dari data pengaduan Pekerja Migran Indonesia (PMI) dengan dataset terdiri atas 130 data periode November–Desember 2024. Dataset memiliki 11 atribut prediktor dan 1 atribut target *(label)*, yaitu status Progres Pengaduan yang terdiri atas dua kelas, yaitu BELUM dan SELESAI. Distribusi kelas pada dataset meliputi 72 data berstatus BELUM dan 58 data berstatus SELESAI.

Dataset yang digunakan terdiri dari berbagai atribut pengaduan seperti gender, kabupaten/kota, media aduan, negara penempatan, klasifikasi permasalahan, pengguna jasa, serta status tindak lanjut pengaduan [10]. Data dikumpulkan dari laporan pengaduan yang masuk melalui sistem pelayanan maupun media pengaduan yang tersedia. Tahap pengumpulan data dilakukan untuk memperoleh dataset yang relevan dengan kebutuhan penelitian sehingga proses klasifikasi dapat dilakukan secara optimal. Rincian atribut dan contoh data penelitian ditampilkan pada Tabel 1**.**

Tabel 1. Data set

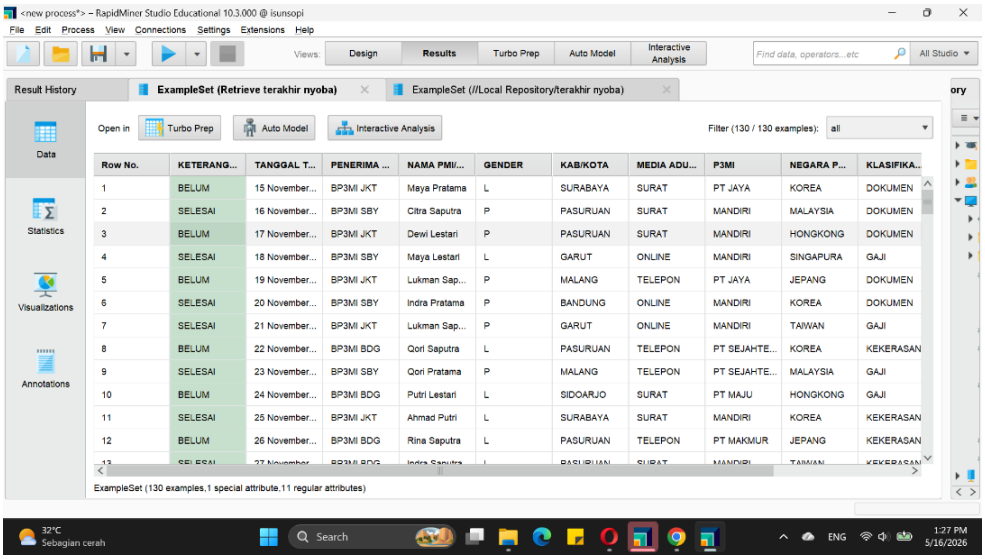
Tanggal Terima Aduan	Penerima Aduan	.	Negara Penempatan	Klasifikasi Permasalahan	Media Tl	Pengguna Jasa	Keterangan Progres
15 November 2024	BP3MI JKT	.	Korea	Dokumen	Online	Badan Hukum	Belum
16 November 2024	BP3MI SBY	.	Malaysia	Dokumen	Online	Perseorangan	Selesai
17 November 2024	BP3MI JKT	.	Hongkong	Dokumen	Online	Perseorangan	Belum
18 November 2024	BP3MI SBY	.	Singapura	Gaji	Online	Perseorangan	Selesai
19 November 2024	BP3MI JKT	.	Jepang	Dokumen	Online	Perseorangan	Belum
20 November 2024	BP3MI SBY	.	Korea	Dokumen	Online	Badan Hukum	Selesai
21 November 2024	BP3MI JKT	.	Taiwan	Gaji	Langsung	Badan Hukum	Selesai
22 November 2024	BP3MI BDG	.	Korea	Kekerasan	Langsung	Perseorangan	Belum
23 November 2024	BP3MI SBY	.	Malaysia	Gaji	Online	Badan Hukum	Selesai
24 November 2024	BP3MI BDG	.	Hongkong	Gaji	Online	Badan Hukum	Belum
25 November 2024	BP3MI JKT	.	Korea	Kekerasan	Online	Perseorangan	Selesai
26 November 2024	BP3MI BDG	.	Jepang	Kekerasan	Online	Perseorangan	Belum
27 November 2024	BP3MI BDG	.	Taiwan	Kekerasan	Online	Perseorangan	Selesai
28 November 2024	BP3MI BDG	.	Taiwan	Gaji	Langsung	Perseorangan	Belum
29 November 2024	BP3MI JKT	.	Jepang	Dokumen	Online	Badan Hukum	Selesai
30 November 2024	BP3MI BDG	.	Taiwan	Dokumen	Online	Perseorangan	Belum
01 December 2024	BP3MI SBY	.	Hongkong	Dokumen	Online	Badan Hukum	Selesai
02 December 2024	BP3MI JKT	.	Taiwan	Kekerasan	Online	Perseorangan	Selesai
03 December 2024	BP3MI JKT	.	Jepang	Dokumen	Langsung	Badan Hukum	Belum
04 December 2024	BP3MI JKT	.	Jepang	Sakit	Online	Badan Hukum	Belum
05 December 2024	BP3MI JKT	.	Hongkong	Sakit	Langsung	Perseorangan	Selesai
06 December 2024	BP3MI SBY	.	Korea	Gaji	Online	Badan Hukum	Selesai
07 December 2024	BP3MI BDG	.	Korea	Gaji	Langsung	Perseorangan	Belum

Tanggal Terima Aduan	Penerima Aduan	.	Negara Penempatan	Klasifikasi Permasalahan	Media TI	Pengguna Jasa	Keterangan Progres
08 December 2024	BP3MI SBY	.	Korea	Gaji	Online	Badan Hukum	Selesai
09 December 2024	BP3MI BDG	.	Jepang	Gaji	Online	Perseorangan	Belum
10 December 2024	BP3MI SBY	.	Singapura	Kekerasan	Online	Badan Hukum	Selesai
11 December 2024	BP3MI BDG	.	Malaysia	Kekerasan	Langsung	Perseorangan	Belum
12 December 2024	BP3MI JKT	.	Taiwan	Sakit	Online	Perseorangan	Selesai
13 December 2024	BP3MI BDG	.	Singapura	Sakit	Online	Badan Hukum	Belum
14 December 2024	BP3MI JKT	.	Malaysia	Dokumen	Online	Perseorangan	Selesai
15 December 2024	BP3MI SBY	.	Singapura	Sakit	Online	Badan Hukum	Belum
16 December 2024	BP3MI SBY	.	Taiwan	Gaji	Langsung	Badan Hukum	Selesai
17 December 2024	BP3MI BDG	.	Korea	Gaji	Langsung	Badan Hukum	Belum
18 December 2024	BP3MI JKT	.	Taiwan	Sakit	Langsung	Perseorangan	Selesai
19 December 2024	BP3MI BDG	.	Malaysia	Kekerasan	Online	Badan Hukum	Belum
20 December 2024	BP3MI BDG	.	Singapura	Gaji	Online	Badan Hukum	Selesai
21 December 2024	BP3MI JKT	.	Jepang	Dokumen	Online	Perseorangan	Belum

B. Preprocessing Data

Tahap *preprocessing* dilakukan untuk meningkatkan kualitas data sebelum digunakan dalam proses klasifikasi. Data yang telah dikumpulkan kemudian diolah menggunakan RapidMiner agar siap digunakan pada proses data mining. Pada tahap ini, dataset yang digunakan terdiri dari **130 data *(examples)*** dengan 11 atribut reguler dan 1 atribut khusus ***(label/class)*** yang akan digunakan sebagai target klasifikasi.

Adapun tahapan *preposcessing* yang dilakukan meliputi: Data *Cleaning* (menghapus data kosong, *(missing value))*, data duplikat, serta data yang tidak lengkap), Data *Selection* (memilih atribut yang relevan dengan penelitian), Data *Transformation* (menyesuaikan format data agar dapat diproses menggunakan algoritma klasifikasi atribut kategorikal diubah ke format yang sesuai dengan kebutuhan proses klasifikasi pada RapidMiner [11]. Hasil preprocessing data pengaduan PMI menggunakan RapidMiner ditunjukkan pada Gambar 2**.**

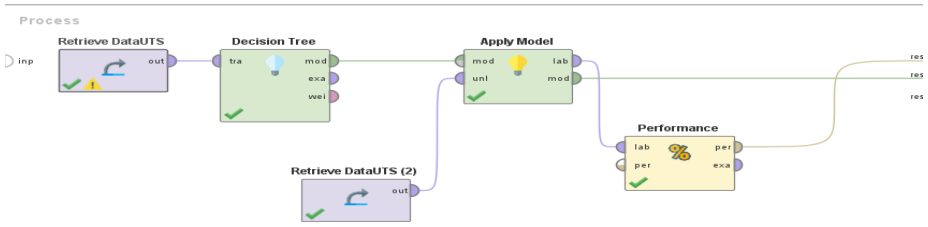


Gambar 2. Hasil *preprocessing* data pengaduan PMI yang telah melalui proses pembersihan, pemilihan atribut, dan transformasi data menggunakan *rapidminer*

Hasil *preprocessing* menunjukkan bahwa dataset telah tersusun secara terstruktur dan siap digunakan pada tahap pembentukan model klasifikasi menggunakan *Decision Tree*. *Proses preprocessing* yang dilakukan berhasil meningkatkan kualitas data dengan menghilangkan data yang tidak lengkap, memperbaiki *inkonsistensi* data, serta memastikan setiap atribut berada pada format yang sesuai untuk proses klasifikasi.

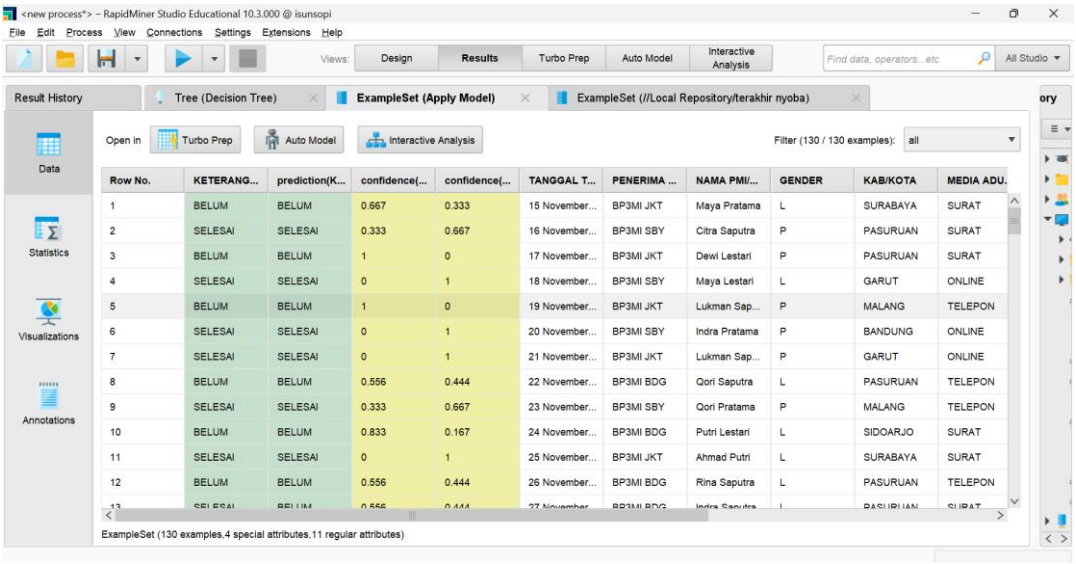
C. Penerapan Metode

Setelah tahap *preprocessing* selesai dilakukan, langkah berikutnya adalah proses pemodelan menggunakan algoritma *Decision Tree*. Algoritma *Decision Tree* bekerja dengan membangun struktur pohon keputusan melalui proses pembagian data ke dalam beberapa cabang berdasarkan atribut yang memiliki pengaruh terbesar terhadap kelas data. Setiap percabangan pada pohon keputusan merepresentasikan aturan keputusan yang digunakan untuk menentukan hasil klasifikasi [12]. Adapun tahapan proses klasifikasi terdiri dari beberapa operator sebagaimana ditujukan pada Gambar 3.



Gambar 3. Alur proses klasifikasi data *decision tree* pada rapidminer

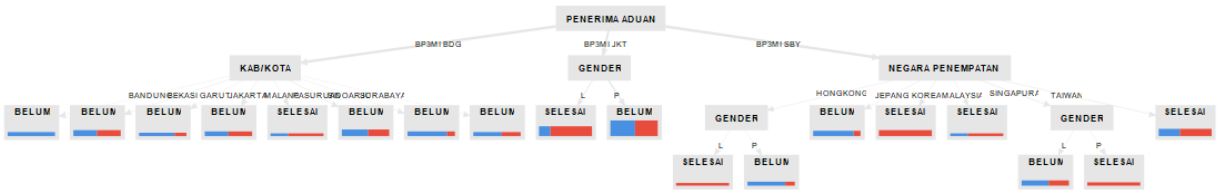
1) *Retrieve* DataUTS: Tahap pertama adalah mengambil dataset pengaduan PMI dari *repository* RapidMiner menggunakan operator ***Retrieve* DataUTS.** Dataset yang digunakan merupakan hasil *preprocessing* yang telah dibersihkan dan disiapkan sebelumnya. Data ini berisi atribut-atribut yang digunakan dalam proses klasifikasi seperti *gender*, kabupaten/kota, media aduan, negara penempatan, klasifikasi permasalahan, dan atribut lainnya.
2) *Decision Tree*: Pada tahap ini algoritma melakukan proses pembentukan pohon keputusan berdasarkan atribut yang memiliki pengaruh terbesar terhadap *label* klasifikasi. Hasil dari tahap ini berupa model pohon keputusan yang digunakan untuk menentukan status pengaduan PMI.
3) *Retrieve* DataUTS (2): Operator ini digunakan untuk mengambil dataset yang akan digunakan pada tahap pengujian model. Data ini diteruskan menuju proses penerapan model agar dapat dilakukan evaluasi performa klasifikasi.
4) *Apply* Model: Tahapan untuk menerapkan model *Decision Tree* yang telah terbentuk ke dataset pengujian menggunakan operator *Apply* Model.



Gambar 4. Hasil setelah *apply* model di rapidminer

Proses ini menghasilkan prediksi kelas untuk setiap data pengaduan berdasarkan pola yang dipelajari sebelumnya. Hasil penerapan model Decision Tree terhadap data pengujian menggunakan operator Apply Model ditunjukkan pada Gambar 4**.** Hasil tersebut menunjukkan prediksi kelas yang dihasilkan oleh model terhadap setiap data pengaduan.

5) *Perfomance:* Tahap terakhir adalah evaluasi model menggunakan operator *Performance.* Evaluasi dilakukan untuk mengetahui kemampuan model dalam melakukan klasifikasi data pengaduan PMI. Pengukuran performa dilakukan menggunakan *confusion matrix* dengan parameter *accuracy, precision, recall.* Melalui tahapan tersebut, selanjutnya, model *Decision Tree* dilakukan untuk proses klasifikasi data pengaduan PMI secara otomatis sehingga membantu proses pengelompokan data menjadi lebih cepat dan terstruktur[13]. Berdasarkan struktur pohon keputusan pada Gambar 5, atribut Penerima Aduan menjadi root node atau atribut utama yang paling berpengaruh dalam menentukan hasil klasifikasi status pengaduan.



Gambar 5. Hasil pohon keputusan *(decision tree)*

Gambar 5 tersebut menunjukkan hasil pembentukan model *Decision Tree* untuk klasifikasi data pengaduan PMI menggunakan parameter terbaik yang telah diperoleh sebelumnya. Berdasarkan hasil pohon keputusan, atribut Penerima Aduan menjadi *root node* atau atribut utama yang paling berpengaruh dalam menentukan hasil klasifikasi status pengaduan.

Hasil visualisasi menunjukkan bahwa atribut Penerima Aduan, *Gender,* Kab/Kota, dan Negara Penempatan memiliki kontribusi terbesar dalam proses klasifikasi pengaduan PMI. Penggunaan ***maximal depth* 3** menghasilkan struktur pohon yang lebih sederhana sehingga lebih mudah dipahami dan diinterpretasikan tanpa mengurangi performa model. Secara keseluruhan, model *Decision Tree* mampu menghasilkan aturan klasifikasi yang jelas sehingga dapat membantu proses pengelompokan data pengaduan PMI secara lebih cepat, terstruktur, dan otomatis[14].

D. Evaluasi Hasil

Tahap evaluasi dilakukan untuk mengukur performa model klasifikasi yang telah dibangun menggunakan algoritma *Decision Tree.* Evaluasi model dilakukan menggunakan *confusion matrix* untuk mengetahui tingkat ketepatan model dalam melakukan klasifikasi data pengaduan PMI.Pengukuran performa model dilakukan menggunakan beberapa parameter evaluasi, yaitu: *accuracy, precision, dan recall.* Berdasarkan hasil pengujian, metode *Decisison tree* dalam mengelompokkan data pengaduan PMI. Model mampu mengenali pola pengaduan dalam lebih cepat dan terstruktur dibandingkan proses manual. valuasi model dilakukan menggunakan confusion matrix untuk mengetahui kemampuan model dalam mengklasifikasikan status pengaduan PMI. Hasil confusion matrix yang meliputi nilai prediksi benar dan salah untuk kelas BELUM dan SELESAI ditampilkan pada Tabel 2**.**

Tabel 2. Hasil *confusion matrix*

accuracy: 86,92%			
	true BELUM	*true* SELESAI	*class precision*
pred. BELUM	59	13	81.94%
pred. SELESAI	4	94	93.10%
class recall	93.65%	80.60%	

Hasil evaluasi menunjukkan bahwa masih terdapat beberapa kesalahan klasifikasi pada data yang memiliki karakteristik hampir sama antar kategori pengaduan. Hal tersebut terjadi karena metode *Decision Tree* menggunakan batas nilai tertentu dalam proses pengambilan keputusan sehingga data dengan karakteristik yang mirip dapat menghasilkan klasifikasi yang berbeda. Namun secara keseluruhan, metode ini tetap mampu memberikan hasil klasifikasi yang baik dan mudah[15].

III. HASIL DAN PEMBAHASAN

Pengujian dilakukan dengan mencoba beberapa kombinasi parameter pada algoritma *Decision Tree,* yaitu *criterion, maximal depth, dan minimal leaf size.* Pengujian parameter dilakukan untuk mengetahui kombinasi parameter yang menghasilkan performa terbaik dalam proses klasifikasi data pengaduan PMI. Parameter *criterion* yang digunakan pada penelitian ini terdiri dari *information gain, gain ratio, dan gini index.* Selain itu, dilakukan variasi nilai *maximal depth* untuk mengatur tingkat kedalaman pohon keputusan serta *minimal leaf size* untuk menentukan jumlah minimum data pada setiap *leaf node*[16]. Variasi parameter tersebut digunakan untuk melihat pengaruh perubahan parameter terhadap performa model klasifikasi. Hasil pengujian kombinasi criterion, maximal depth, dan minimal leaf size secara keseluruhan ditampilkan pada Tabel 3.

Tabel 3. Hasil pengujian paramater

No	*Criterion*	*Maximal Depth*	*Minimal Leaf Size*	*Accuracy*
1	*gain_ratio*	1	1	51.54%
2	*gain_ratio*	2	1	81,54%
3	*gain_ratio*	2	2	60,74%
4	*gain_ratio*	2	3	60,77%
5	*gain_ratio*	2	4	60,77%
6	*gain_ratio*	2	5	60,77%
7	*gain_ratio*	2	6	60,77%
8	*gain_ratio*	3	1	86,15%
9	*information_gain*	3	1	86,92%
10	*gain_ratio*	3	2	66.15%
11	*information_gain*	3	2	67.69%
12	*gini_index*	3	2	67.69%
13	*gain_ratio*	3	3	67,69%
14	*gain_ratio*	3	4	63,08%
15	*information_gain*	3	4	65,38%
16	*gain_ratio*	3	5	63,38%
17	*information_gain*	3	5	65.38%
18	*gain_ratio*	3	6	63,08%
19	*information_gain*	3	6	65,38%
20	*gain_ratio*	3	7	63,08%
21	*information_gain*	3	7	64,62%
22	*gain_ratio*	3	8	63,08%
23	*information_gain*	3	8	62,31%
24	*gini_index*	3	8	63,08%
25	*gain_ratio*	3	9	61,54%
26	*information_gain*	3	9	60,77%
27	*gain_ratio*	3	10	61,54%
28	*information_gain*	3	10	61,54%
29	*gain_ratio*	4	1	86,15%
30	*information_gain*	4	1	86,92%
31	*gain_ratio*	4	2	78,46%
32	*information_gain*	4	2	72,23%
33	*gini_index*	4	2	79,23%
34	*gain_ratio*	4	3	69,23%
35	*information_gain*	4	3	70,77%

Berdasarkan hasil pengujian yang telah dilakukan, terlihat bahwa setiap kombinasi parameter menghasilkan nilai akurasi yang berbeda. Perbedaan tersebut menunjukkan bahwa pemilihan parameter pada algoritma *Decision Tree* memiliki pengaruh terhadap kemampuan model dalam mengklasifikasikan data pengaduan PMI. Parameter yang diuji meliputi ***criterion, maximal depth, dan minimal leaf size,*** di mana masing-masing parameter memberikan dampak terhadap pembentukan struktur pohon keputusan dan proses pengambilan keputusan model. Hasil pengujian menunjukkan bahwa semakin tepat kombinasi parameter yang digunakan, maka semakin baik kemampuan model dalam mengenali pola data pengaduan. Sebaliknya, penggunaan parameter yang kurang sesuai dapat menyebabkan penurunan performa model karena pohon

keputusan menjadi terlalu sederhana atau terlalu kompleks. Berdasarkan hasil pengujian pada Tabel 3**,** diperoleh beberapa konfigurasi yang menghasilkan accuracy tertinggi. Rangkuman konfigurasi tersebut ditampilkan pada Tabel 4**.**

Tabel 4. Hasil *accuracy* tertinggi

Criterion	***Maximal Depth***	***Minimal Leaf Size***	***Accuracy***
information_gain	*3*	*1*	*86,92%*
Information_gain	*4*	*1*	*86.92%*
Information_gain	*7*	*1*	*86.92%*

Dengan demikian, hasil pengujian parameter yang telah dilakukan, diperoleh beberapa kombinasi parameter yang menghasilkan nilai *accuracy* tertinggi sebesar **86,92%,** yaitu pada penggunaan ***criterion information gain*** dengan ***maximal depth* 3, 4, dan 7,** serta ***minimal leaf size* 1.** Meskipun menghasilkan nilai *accuracy* yang sama, penelitian ini memilih kombinasi ***maximal depth 3*** sebagai parameter terbaik karena menghasilkan struktur pohon keputusan yang lebih sederhana dan mudah diinterpretasikan dibandingkan kedalaman yang lebih besar.

Penggunaan kedalaman pohon yang lebih kecil juga dapat mengurangi kompleksitas model sehingga risiko terbentuknya pohon keputusan yang terlalu rumit dapat diminimalkan. Selain itu, nilai ***minimal leaf size* 1** memungkinkan model membentuk percabangan yang lebih detail sehingga pola pada data pengaduan PMI dapat dikenali dengan lebih baik [17].

Hasil tersebut menunjukkan bahwa pemilihan parameter memberikan pengaruh terhadap performa klasifikasi yang dihasilkan. Kombinasi parameter yang tepat mampu meningkatkan kemampuan model dalam mengenali pola data pengaduan PMI secara lebih optimal. Oleh karena itu, proses tuning parameter menjadi salah satu tahapan penting untuk memperoleh model *Decision Tree* yang efektif, mudah dipahami, serta memiliki performa klasifikasi yang baik dalam mengelompokkan data pengaduan PMI.

IV. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa metode *Decision Tree* dapat diterapkkan untuk mengklasifikasikan data pengaduan Pekerja Migran Indonesia (PMI) dengan baik. Proses klasifikasi dilakukan melalui beberapa tahapan, yaitu pengumpulan data, pembentukan model klasifikasi, serta evaluasi hasil menggunakan aplikasi RapidMiner.

Hasil pengujian menunjukkan bahwa setiap kombinasi parameter menghasilkan tingkat *accuracy* yang berbeda, sehingga pemilihan parameter berpengaruh terhadap performa model klasifikasi. Dari beberapa parameter yang diuji, diperoleh tiga kombinasi parameter yang menghasilkan *accuracy* tertinggi yang sama sebesar **86,92%,** yaitu pada penggunaan ***criterion information gain*** dengan ***maximal depth* 3, 4, dan 7**, serta ***minimal leaf size* 1**. Meskipun menghasilkan *accuracy* yang sama, penelitian ini memilih kombinasi ***maximal depth* 3** sebagai parameter terbaik karena menghasilkan *preprocessing* struktur pohon keputusan yang lebih sederhana, lebih mudah diinterpretasikan, dan memiliki kompleksitas model yang lebih rendah dibandingkan depth yang lebih besar. Pemilihan depth yang lebih kecil juga membantu mengurangi risiko terbentuknya pohon keputusan yang terlalu kompleks sehingga model lebih mudah dipahami dan diimplementasikan. Selain itu, penggunaan ***minimal leaf size 1*** memungkinkan model membentuk percabangan yang lebih detail sehingga pola pada data pengaduan PMI dapat dikenali dengan lebih baik.

Adapun saran untuk penelitian selanjutnya adalah melakukan eksperimen perbandingan antara metode *Decision Tree* dengan algoritma *Machine Learning* lain seperti *Random Forest* atau SVM, serta memperluas *tuning* parameter guna menemukan model yang lebih optimal. Selain itu, disarankan untuk menerapkan teknik *Natural Language Processing* (NLP) dan penanganan ketidakseimbangan data pada tahap *preprocessing* jika data pengaduan berbentuk teks naratif yang tidak seimbang. Terakhir, model terbaik yang telah dihasilkan sebaiknya diintegrasikan ke dalam sistem informasi berbasis *web* atau *dashboard* interaktif agar klasifikasi otomatis ini dapat diimplementasikan secara *real-time* oleh pihak pengelola untuk mempercepat pengambilan keputusan dalam pelayanan Pekerja Migran Indonesia (PMI).

DAFTAR PUSTAKA

[1] F. O. Wuisan, N. Mohede, and H. Tuwaidan, "Kajian Pengaturan perlindungan Hukum Terhadap Tenaga Kerja Indonesia Diluar Negeri," *LEX Crim.*, vol. 13, no. 4, 2025.

[2] R. Nursyahfitri, A. N. Maharadja, R. A. Farissa, and Y. Umaidah, "Klasifikasi Penentuan Jenis Obat Menggunakan Algoritma Decision Tree," *JIP (Jurnal Inform. Polinema)*, vol. 7, no. 3, pp. 53–60, 2021, doi: 10.33795/jip.v7i3.629.

[3] Ahmada Ilham Huzaini and Zaehol fatah, "Implementasi Decision Tree Dalam Prediksi Penyakit Stroke Menggunakan Rapidminer," *J. Mhs. Tek. Inform.*, vol. 5, no. 1, pp. 187–192, 2026, doi: 10.35473/jamastika.v5i1.4646.

[4] S. Sheila and A. A. Diphayana, "Analisis Performa Algoritma C.45 dan Klasifikasi Decision-Tree dalam Memprediksi Penyakit Diabetes," *SINTESIA J. Sist. dan Teknol. Inf. Indones.*, vol. 04, no. 1, pp. 1–9, 2024.

[5] Fitra Lia Shofiah and Zaehol Fatah, "Klasifikasi Data Penyakit Demam Berdarah Dengue (DBD) Melalui Algoritma Decision Tree Dengan RapidMiner," *J. Mhs. Tek. Inform.*, vol. 5, no. 1, pp. 34–40, 2026, doi: 10.35473/jamastika.v5i1.4571.

[6] N. Wayan, W. Putu, G. Surya, C. Nugraha, E. Hartono, and I. W. Dharma, "penerapan data mining untuk klasifikasi penjualan barang terlaris menggunakan metode Decesion Tree," *J. Teknol. Inf. dan Komput.*, vol. 8, no. 3, pp. 268–279, 2022, [Online]. Available: https://jurnal.undhirabali.ac.id/index.php/jutik/article/view/2081

[7] R. Ningsih, M. D. Basunanda, and W. Erawati, "Implementasi Algoritma C4.5 Untuk Klasifikasi Tingkat Kenaikan Jabatan Pada PT Siprama Cakrawala Depok," *J. Infortech*, vol. 7, no. 1, pp. 6–13, 2025, doi: 10.31294/infortech.v7i1.25469.

[8] P. R. Rosmila, "Komparasi Algoritma Naive Bayes dan Decision Tree untuk Menganalisis Kepuasan Masyarakat," *Inf. Syst. J.*, vol. 8, no. 2, pp. 109–118, 2025, [Online]. Available: doi:https://doi.org/10.24076/infosjournal.2025v8i02.2364

[9] F. Y. Biyanto and V. Frendiana, "Perancangan Website Sistem Manajemen 'Peduli PMI' sebagai Wadah Pengaduan dan Perlindungan untuk Pekerja Migran Indonesia (PMI) di Malaysia," in *SNESTIK Seminar Nasional Teknik Elektro, Sistem Informasi, dan Teknik Informatika*, Surabaya: Fakultas Teknik Elektro dan Teknologi Informasi - Institut Teknologi Adhi Tama Surabaya, 2024, pp. 570–577. doi: 10.31284/p.snestik.2024.5916.

[10] Pemerintah Indonesia, "Undang-Undang Republik Indonesia NO. 18 2017 Perlindungan Pekerja Migran Indonesia," *Pemerintah Indones.*, 2017.

[11] A. D. M. Saputra and H. Firmansyah, "Implementasi Algoritma Decision Tree untuk Memprediksi Kualitas Udara dan Polusi dengan RapidMiner," *sudo J. Tek. Inform.*, vol. 4, no. 3, pp. 212–219, 2025, doi: https://doi.org/10.56211/sudo.v4i3.1145.

[12] M. A. Fais, M. R. A. Lubis, A. Aulia, and I. Syafitri, "Implementasi Algoritma Decision Tree untuk Klasifikasi Serangan Jantung," *J. Sist. Inf. dan Ilmu Komput.*, vol. 1, no. 4, pp. 207–212, 2023, doi: 10.59581/jusiik-widyakarya.v1i4.1895.

[13] C. Nisa and M. I. Rosyadi, "Prediksi Kesehatan Mental Mahasiswa Universitas Yudharta Pasuruan Dengan Klasifikasi Decision Tree," *J. Inform. dan Tek. Elektro Terap.*, vol. 13, no. 3S1, pp. 925–933, 2025, doi: http://dx.doi.org/10.23960/jitet.v13i3S1.7925.

[14] U. Suropati and D. Saputra, "Implementasi Algoritma Decision Tree C4.5 untuk Klasifikasi Pengangkatan Karyawan Tetap (Studi Kasus) di PT Intinusa Teknik Sejahtera)," *J. Manag. Bussines*, vol. 7, no. 5, pp. 427–433, 2025, doi: https://doi.org/10.31539/68x3rd85.

[15] S. Wahyuni and Z. Fatah, "Analisis Klasifikasi Status Perokok Menggunakan Algoritma Decision Tree Berbasis RapidMiner," *J. Mhs. Tek. Inform.*, vol. 5, no. 1, pp. 312–318, 2026, doi: 10.35473/jamastika.v5i1.4639.

[16] R. N. Kholisha and Z. Fatah, "Penerapan Algoritma Decision Tree untuk Klasifikasi Diagnosa Pasien," *J. Mhs. Tek. Inform.*, vol. 5, no. 1, pp. 98–103, 2026, doi: 10.35473/jamastika.v5i1.4596.

[17] S. R. Putri, "Implementation of the decision tree method and C4.5 algorithm for teacher classification in obtaining web-based position promotion at SMK Al-Ihya Selajambe Kuningan," *J. Info Sains Inform. dan Sains*, vol. 14, no. 01, pp. 783–794, 2024, doi: 10.54209/infosains.v14i01.