

KLASIFIKASI KOMPETENSI SISWA DI SMK BASURAGA MENGUNAKAN ALGORITMA DECISION TREE

Musyaroifah¹, Martanto², Umi Hayati³

^{1,3}Teknik Informatika, ²Manajemen Informatika STMIK IKMI Cirebon
Jalan Perjuangan No10B Majasem, Kota Cirebon, Indonesia
musyarofahasshiddieqie@gmail.com

ABSTRAK

Penelitian ini difokuskan pada klasifikasi kompetensi siswa dengan metode decision tree di Sekolah Menengah Kejuruan (SMK) Basuraga. Dalam penelitian ini, permasalahan yang diangkat adalah tentang bagaimana mengenali kompetensi siswa dengan menerapkan metode decision tree. Tujuan utama adalah mengidentifikasi kompetensi siswa guna mendukung manajemen data, perancangan program pembelajaran yang lebih terfokus, dan perencanaan kegiatan ekstrakurikuler yang sesuai dengan kebutuhan siswa. Dengan menggunakan algoritma decision tree untuk mengklasifikasikan kompetensi siswa berdasarkan nilai rata-rata per mata pelajaran dari data siswa tahun ajaran 2022-2023, hasil penelitian menunjukkan akurasi sebesar 98,24%. Model ini memiliki presisi prediksi kompeten mencapai 99,30%, presisi prediksi tidak kompeten mencapai 93,10%, dan tingkat keberhasilan model untuk memprediksi siswa kompeten dan tidak kompeten masing-masing sebesar 98,60% dan 96,43%. Dengan demikian, metode klasifikasi decision tree terbukti efektif dalam mengklasifikasikan kompetensi siswa, memberikan dasar bagi pengembangan strategi manajemen pendidikan yang lebih efisien dan terfokus.

Kata kunci : *klasifikasi kompetensi, decision tree, smk basuraga*

1. PENDAHULUAN

SMK Basuraga, sebagai institusi pendidikan tingkat menengah, berkomitmen kuat untuk memastikan bahwa para siswanya memiliki keterampilan dan kompetensi sesuai dengan perkembangan terkini di dunia kerja. Manajemen kompetensi siswa menjadi elemen krusial dalam upaya meningkatkan mutu pendidikan, dan penerapan algoritma Decision Tree pada data siswa membantu mengidentifikasi faktor-faktor yang berdampak signifikan pada kompetensi, seperti nilai akademik, kehadiran, dan partisipasi dalam kegiatan ekstrakurikuler. Penelitian ini bertujuan untuk memperdalam pemahaman mengenai kompetensi siswa SMK Basuraga dan memberikan dasar yang solid untuk perencanaan pembelajaran yang lebih efektif. Namun, dalam mengidentifikasi permasalahan implementasi algoritma Decision Tree, beberapa tantangan muncul, seperti pengumpulan data yang akurat, penentuan metrik evaluasi yang tepat, pemilihan atribut yang relevan, serta skalabilitas dan efisiensi. Komunikasi yang efektif, manajemen risiko, partisipasi pemangku kepentingan, dan evaluasi terus-menerus juga menjadi faktor penting untuk memastikan keberhasilan implementasi algoritma Decision Tree di lingkungan pendidikan SMK Basuraga.

Beberapa penelitian terdahulu telah menerapkan pendekatan klasifikasi dengan menggunakan Algoritma Decision Tree. Contohnya, satu penelitian membahas penerapan Decision Tree C4.5 untuk memilih perguruan tinggi pendamping Program SMK Pusat Keunggulan. Dalam penelitian ini, pengembangan sistem pendukung keputusan dilakukan dengan metode waterfall, menghasilkan

skor yang relevan, dan diharapkan dapat berkontribusi pada peningkatan mutu pendidikan SMK di Indonesia[1]. Studi lainnya menguji Decision Tree dalam mengklasifikasikan kelulusan mahasiswa di Fakultas Ekonomi Universitas Garut, menemukan model dengan tingkat akurasi 79,53% dan nilai recall 19,23%, menyoroti pentingnya penyelesaian studi tepat waktu dan potensi data mining [2]. Sebuah penelitian juga mengevaluasi implementasi Decision Tree ID3 untuk menentukan siswa yang berhak masuk kelas unggulan di SMPN 6 Kota Kediri, mencapai tingkat kepresisian 82% dan keingatan 96%, sambil memberikan rekomendasi untuk pengembangan lebih lanjut[3]. Studi lainnya membandingkan akurasi algoritma C4.5 dan Naive Bayes dalam memprediksi kelulusan mahasiswa, menunjukkan bahwa Naive Bayes memiliki tingkat akurasi lebih tinggi, yaitu 81,58% [4]. Akhirnya, implementasi K-Nearest Neighbor, Decision Tree, dan Naive Bayes dalam mengklasifikasikan tingkat kepemimpinan siswa di sekolah menengah atas menunjukkan bahwa K-NN memiliki skor akurasi tertinggi, yaitu 95,86%, dibandingkan dengan Decision Tree dan Naive Bayes [5].

Penelitian ini memiliki tujuan untuk mendalami pemahaman tentang kompetensi siswa di SMK Basuraga dengan mengidentifikasi faktor-faktor yang memengaruhi pembentukan kompetensi mereka. Penggunaan algoritma klasifikasi seperti Decision Tree dalam penelitian ini memberikan dasar yang kokoh untuk pengambilan keputusan yang lebih akurat dalam manajemen pendidikan, meningkatkan efektivitas pengelolaan siswa di sekolah. Metode pendekatan kuantitatif melibatkan pengumpulan dan analisis data numerik, termasuk data akademik siswa,

database sekolah, dan survei. Implikasi dari penelitian ini melibatkan dukungan pengambilan keputusan yang lebih baik di sekolah, perencanaan pembelajaran yang lebih terfokus, optimalisasi alokasi sumber daya, dan pengembangan program pendidikan yang lebih relevan. Model hasil klasifikasi kompetensi siswa diharapkan dapat memberikan panduan bagi guru dan staf pendidikan dalam merancang program pembelajaran yang disesuaikan dengan kebutuhan siswa. Penerapan Decision Tree juga diharapkan dapat memfasilitasi pemahaman yang lebih baik tentang faktor-faktor yang mempengaruhi perkembangan kompetensi siswa, memberikan wawasan berharga bagi sekolah dalam merancang kebijakan pendidikan yang lebih efektif dan strategi pengembangan kurikulum yang lebih tepat sasaran. Selain memberikan manfaat pada tingkat individu, model ini diharapkan dapat berkontribusi pada perbaikan keseluruhan sistem pendidikan di SMK Basuraga, meningkatkan efisiensi evaluasi kompetensi, dan mendukung pencapaian tujuan pendidikan yang lebih luas.

2. TINJAUAN PUSTAKA

Tinjauan pustaka menjadi jendela peneliti untuk memahami dasar ilmiah yang mendukung topik penelitian. Dalam prolog ini, peneliti akan menjelajahi literatur terkait yang telah ada, merinci teori-teori utama, konsep-konsep kunci, dan metodologi yang digunakan dalam penelitian sebelumnya. Melalui langkah ini, peneliti dapat mengidentifikasi kesenjangan dalam pengetahuan, memahami perkembangan terkini, dan merinci sumbangan penelitian sebelumnya terhadap pemahaman peneliti tentang topik ini.

2.1. Metode

Dalam melaksanakan penelitian atau karya ilmiah, penting bagi peneliti untuk merujuk pada jurnal sebagai sumber panduan yang krusial. Jurnal memainkan peran yang sangat penting dalam seluruh proses penelitian, membantu peneliti menghadapi dan memahami permasalahan yang ingin diinvestigasi dengan dukungan dari berbagai sumber. Penulis menjalankan pencarian referensi jurnal melalui Google Scholar (<https://scholar.google.com>) dan menggunakan perangkat lunak Publish or Perish (<https://harzing.com/resources/publish-or-perish>). Dalam melakukan pencarian, penulis menggunakan mesin pencari Google dengan berbagai kata kunci, memanfaatkan operator boolean seperti AND, OR, dan NOT untuk memperoleh hasil pencarian yang lebih spesifik. Sebagai contoh, kata kunci "Klasifikasi" DAN "Algoritma Decision Tree" digunakan untuk mencari referensi melalui Google Scholar. Penulis juga mengadopsi strategi pencarian fokus pada topik penelitian yang diinginkan. Melalui Publish or Perish, kata kunci "Algoritma Decision Tree" menghasilkan 200 jurnal terkait, meskipun demikian, penulis tetap melakukan penyaringan lebih lanjut melalui Google

Scholar dengan kata kunci "Klasifikasi Kompetensi menggunakan algoritma Decision Tree." Selain itu, mesin pencari Google juga dimanfaatkan untuk mencari materi tambahan yang diperlukan, termasuk artikel dan referensi lain yang relevan dengan penelitian.

2.2. Hasil

Penelitian "Penerapan Algoritma C4.5 untuk Prediksi Kelulusan Siswa SMK Wahidin" membuktikan keefektifan algoritma C4.5 dalam meramalkan kelulusan siswa dengan akurasi 82,19% hingga 93,15% menggunakan berbagai teknik sampling. Metode decision tree digunakan, dan hasil penelitian memberikan kontribusi penting untuk pengembangan sistem prediksi kelulusan siswa di SMK dengan potensi menjadi dasar untuk sistem yang lebih unggul di masa depan[6]

Penelitian kedua, berjudul "Prediksi Mahasiswa Berpotensi Non Aktif Menggunakan Data Mining dengan Decision Tree dan Algoritma C4.5," menggunakan metode data mining untuk mengklasifikasikan calon mahasiswa berprestasi di STMIK Denpasar. Algoritma C4.5 dipilih berdasarkan keberhasilannya dalam penelitian sebelumnya, seperti memprediksi masa studi mahasiswa dan penerimaan calon pegawai baru. Studi ini juga memanfaatkan pohon keputusan untuk mengolah data mahasiswa aktif dan non-aktif berdasarkan jadwal kuliah, nilai absensi, nilai gagal, dan status pembayaran uang kuliah sesuai peraturan kampus[7]

Penelitian ketiga oleh Imelda Ruwae Lutunani dan Adi Nugroho membahas prediksi kelulusan mahasiswa tepat waktu di FTI UKSW menggunakan metode decision tree. Mereka mempertimbangkan faktor-faktor seperti mata kuliah, kerja praktek, dan seminar tugas akhir sebagai penentu kelulusan. Penelitian menggunakan data mahasiswa dari 2014 hingga 2020 dengan berbagai atribut. Tujuannya adalah memberikan solusi untuk meningkatkan tingkat kelulusan mahasiswa FTI UKSW[8].

Penelitian ini membahas penerapan algoritma decision tree dalam mengklasifikasi kompetensi siswa di SMK Multicomp Depok. Dilakukan oleh Hilman Rifa'i dan tim di STMIK IKMI Cirebon, penelitian ini bertujuan memprediksi kompetensi siswa berdasarkan data nilai rapor semester. Menggunakan RapidMiner, hasilnya menunjukkan efektivitas metode decision tree dan algoritma C4.5 dalam mengukur kemampuan siswa, memberikan pemahaman sesuai kebutuhan pendidikan di Indonesia, dan dapat digunakan untuk prediksi kompetensi siswa guna evaluasi guru dan pemantauan siswa [9]

Penelitian ini menggunakan algoritma decision tree ID3 untuk mengidentifikasi faktor kunci kelulusan siswa Politeknik Statistika STIS melalui analisis data SPMB tahun 2022. Fokusnya pada manajemen bimbingan belajar untuk memahami faktor kontribusi kelulusan siswa. Dengan menggabungkan dataset, penelitian ini efektif menerapkan algoritma decision

tree ID3 dengan evaluasi kinerja tinggi, memberikan wawasan berharga bagi manajemen bimbingan belajar.[10]

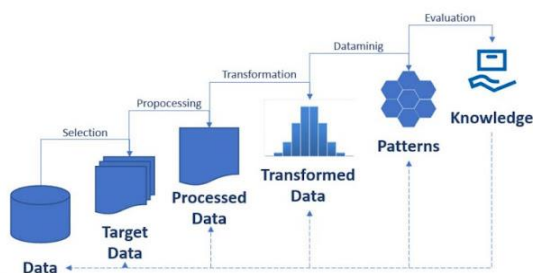
2.3. Tabel

Tabel 1. Ringkasan

Judul	Author	Masalah	Metode
Penerapan Model Prediksi Menggunakan Algoritma C.45 Untuk Prediksi Kelulusan Siswa SMK Wahidin	Farid Ali Ma'ruf ¹ , Aria Pratama ² , Iin Sholihin ³ , Ade Rizki Rinaldi ⁴ , Faturrohman ⁵	Variasi akurasi prediksi kelulusan siswa dengan teknik sampling yang berbeda.	C4.5
Evaluasi Metode Data Mining Decesion Tree C4.5 Dan Naïve Bayes Dalam Memprediksi Pemilihan Jurusan Siswa (Studi Kasus Sma Master Indonesia)	Khaerudin	Kekurangan terkait penggunaan atribut yang relevan.	C4.5 dan Naïve Bayes
Analisis Prediksi Mahasiswa Terhadap Kelulusan Tepat Waktu Menggunakan Metode Data Mining Decision Tree (Studi Kasus: FTI UKSW)	Imelda Ruwae Lutunani, Adi Nugroho	Memprediksi kelulusan mahasiswa tepat waktu di Fakultas Teknologi Informasi Universitas Kristen Satya Wacana (FTI UKSW)	C4.5
Implementasi Algoritma Decision Tree Dalam Klasifikasi Kompetensi Siswa	Hilman Rifa'i, Ryan Hamonangan, Dian Ade Kurnia Kaslani, Mulyawan	Prediksi kelulusan mahasiswa tepat waktu	C4.5
Penerapan Data Mining untuk Memprediksi Tingkat Kelulusan Mahasiswa Menggunakan Algoritma Decision Tree C4.5	Uci Suriani	Kurangnya variasi data yang digunakan	C4,5

3. METODE PENELITIAN

Studi ini menerapkan metode penelitian kuantitatif dengan fokus pada implementasi metode klasifikasi pada data siswa SMK Basuraga. Proses ini mencakup pengumpulan data, pemrosesan data, pelatihan model klasifikasi, evaluasi model, dan analisis hasil. Selain itu, penelitian ini memanfaatkan metode analisis statistik untuk menguji signifikansi dari hasil yang diperoleh. Langkah-langkah metode penelitian terdiri dari:



Gambar 1. Knowledge discovery in databases

Dari gambar 1 diatas dapat di uraikan tahapan-tahapan penelitian adalah sebagai berikut :

- Tahap pertama adalah data selection dimana pada tahap ini dilakukan pengumpulan data yang berupa nilai raport siswa di smk basuraga tahun ajaran 2022/2023.
- Tahap kedua adalah data preprocessing yang bertujuan untuk pembersihan data yang duplikat atau yang tidak valid.
- Tahap ketiga adalah Tranformasi data dimana pada tahap ini dilakukan pengelompokkan variable dan pelabelan atribut untuk

meningkatkan kualitas data untuk analisis klasifikasi.

- Tahap keempat adalah data mining. Pada tahap ini dilakukan pemodelan algoritma decision tree menggunakan rapid miner.
- Tahap terakhir adalah Evaluasi dimana Peneliti perlu menggunakan metrik evaluasi seperti akurasi, presisi, recall, F1-score.

3.1. Sumber Data

Peneliti telah mengumpulkan data yang bersumber dari data nilai akademik siswa yang mencakup nilai-nilai siswa dalam berbagai mata pelajaran pada rapor tiap peserta didik.

3.2. Populasi dan Sampel

Populasi yang menjadi fokus utama adalah seluruh siswa yang terdaftar di SMK Basuraga pada tahun ajaran 2022/2023 sebanyak 171 siswa. Sampel yang digunakan merupakan sampel jenuh atau semua populasi akan dijadikan sampel dalam penelitian ini dengan jumlah 171 siswa.

3.3. Teknik Pengambilan Data

Pengumpulan data primer melibatkan penggunaan data yang sudah ada, seperti data akademik siswa. Data ini diperoleh dari arsip sekolah dan catatan siswa yang telah ada pada tanggal 1 Oktober s/d 7 Oktober 2023.

3.4. Teknik Analisis Data

Decision Tree merupakan teknik klasifikasi yang merancang model berdasarkan informasi yang

diperoleh dari entropy. Saat menggunakan sampel data, akan dilakukan sebuah perhitungan entropy dan gain. Entropy merupakan indikator statistik yang mencerminkan tingkat ketidakpastian dalam data. Sementara itu, gain digunakan sebagai metrik untuk menilai seberapa efektif suatu atribut dalam memisahkan contoh pelatihan sesuai dengan kelas target. Proses ini melibatkan perhitungan nilai entropy dan gain untuk memutuskan atribut mana yang optimal dalam pembagian data pelatihan. Berikut adalah rumus perhitungan entropy dan gain :

a. *Entropy*

entropy dapat menggunakan rumus berikut.

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

Keterangan :

S = Himpunan kasus

n = Jumlah partisi A

pi = proporsi Si terhadap S

b. *Gain*

gain dapat menggunakan rumus berikut.

$$\text{Gain (S,A)} = \text{Entropy (S)} - \sum_{i=1}^n \text{Entropy (Si)}$$

Keterangan :

S = Himpunan kasus

A = Jumlah Partisi A

|Si| = Proporsi Si terhadap S

|S| = Jumlah kasus dalam S

4. HASIL DAN PEMBAHASAN

4.1. Data Nilai Raport

Data nilai siswa tahun ajaran 2022-2023 diperoleh peneliti dari sekolah melalui surat permohonan. Data tersebut berupa nilai rata-rata siswa per mata pelajaran dan nilai rata-rata siswa per kelas, yang terdapat pada tabel 4.1.

Tabel 2. Data siswa

No	Nama	Nisn	P ai	B .I	.	Mul ok	Rata- Rata
1	Bayu Permana	0073675798	84	83	.	83	83,125
2	Heny Nurjana	0086968790	86	83	.	83	82,125
3	Keyla Adinda	0077353577	79	81	.	78	77,75
4	Kharisma	0078004412	85	83	.	83	83,25
5	Laelatul Rohmah	3074463254	85	81	.	83	82,875
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
171	Yayan	0063698627	75	74	75	75	75

4.2. Data Preprocessing

Preprocessing data adalah tahap awal dalam proses klasifikasi kompetensi. Tahap ini dilakukan untuk mempersiapkan data agar dapat diolah dengan optimal oleh algoritma decision tree.

Tabel 3. Data preprocessing

Nama	1	2	Rata-Rata	Kkm	Hasil
Bayu.P	83,1	87,5	85,3	70	K
Heny. N	83,1	84,9	84	70	K
Keyla. A	77,7	77,6	85	70	K
Kharisma	83,2	86,8	85	70	K
Laelatul.A	82,8	84,5	83,6	70	K
Mohammad.A.K	82	82,4	82,2	70	K
.	.	.	.	.	.
Yayan	69,2	69,2	75	75	TK

4.3. Data Transformation

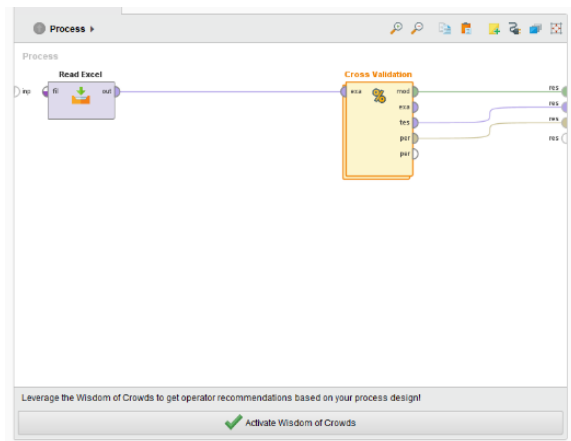
Transformasi data bertujuan untuk mempersiapkan data agar dapat diuji. Tahap pertama adalah mengunggah data set dan melakukan drag dan drop. Tahap selanjutnya adalah mengubah tipe data sesuai dengan kebutuhan algoritma dimana atribut 'Nama Peserta Didik' menjadi 'Id' dan atribut 'hasil' menjadi 'label'. Data tersebut dapat dilihat pada gambar 2.

Row No.	NAMA PESERTA DIDIK	HASIL	No	NISN	KETERANGAN	Smt 1	Smt 2	RATA-RATA	KKM
1	BAYU PERMANA	KOMPETEN	1	73675798	X RPL	83,125	87,583	85,344	70
2	HENY NURJANA	KOMPETEN	2	86968790	X RPL	83,125	84,538	84,051	70
3	KEYLA ADINDA	KOMPETEN	3	77353577	X RPL	77,750	77,588	77,169	70
4	KHARISMA	KOMPETEN	4	78004412	X RPL	83,250	86,813	85,031	70
5	LAELATUL R.	KOMPETEN	5	3074463254	X RPL	82,875	84,500	83,688	70
6	MOHAMMAD A.	KOMPETEN	6	3079721375	X RPL	82,063	82,438	82,251	70
7	MUHAMMAD A.	KOMPETEN	7	85216417	X RPL	82,500	84,813	83,656	70
8	NAPISATUL A.	KOMPETEN	8	72516870	X RPL	83,375	86,125	84,750	70
9	HUR KHATUN.	KOMPETEN	9	77707888	X RPL	82,750	85,063	84,406	70
10	HUR KHATUN.	KOMPETEN	10	91936120	X RPL	82,500	85,063	84,281	70
11	RATU AYUN.	KOMPETEN	11	137108036	X RPL	81,813	84,875	83,344	70
12	ROPATUL A.	KOMPETEN	12	307220434	X RPL	82,938	84,313	83,626	70
13	SILVIA R.	KOMPETEN	13	85278399	X RPL	82,813	85,500	84,156	70
14	DELLY	KOMPETEN	14	70873248	X RPL	83,375	85,875	84,625	70

Gambar 2. Data transformation

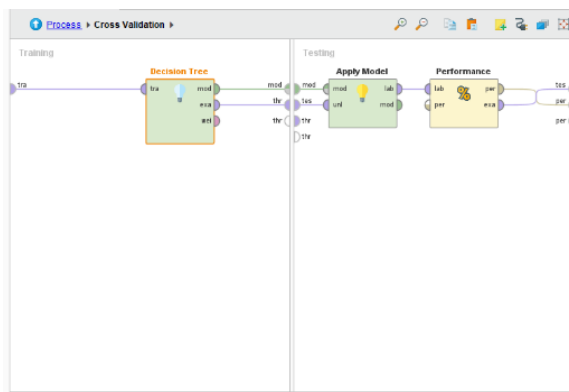
4.4. Data Mining

Setelah data ditransformasi, langkah selanjutnya adalah membangun model algoritma decision tree. Langkah ini dilakukan dengan mengambil data set dan memindahkannya ke dalam desain model. Setelah itu, diambil tools cross validation untuk melakukan klasifikasi data. Tahapan tersebut akan diuraikan sebagai berikut :



Gambar 3. Cross validation

Dari gambar 3 di atas, dapat dijelaskan bahwa dalam mengembangkan model algoritma Decision Tree, diperlukan penerapan operator cross-validation untuk melakukan klasifikasi data.



Gambar 4. Design algoritma decision tree

Gambar 4 menunjukkan bahwa untuk melakukan pelatihan dan pengujian model algoritma Decision Tree, perlu diterapkan operator decision tree, apply model, dan performance. Operator decision tree digunakan untuk melatih model Decision Tree. Apply model digunakan untuk menerapkan model Decision Tree yang telah dilatih ke dalam data pengujian. Performance digunakan untuk mengukur tingkat akurasi model Decision Tree pada data pengujian.

4.5. Pembahasan

Penelitian ini mengukur kompetensi nilai data siswa terhadap suatu mata pelajaran dengan menggunakan metode klasifikasi. Data nilai siswa diperoleh dari sistem informasi pendidikan. Algoritma decision tree digunakan untuk mengklasifikasikan nilai data siswa ke dalam dua kategori, yaitu kompeten dan tidak kompeten.

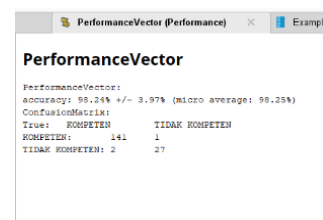
Akurasi model dihitung dengan menggunakan metrik yang tersedia di Rapid Miner. Nilai confusion matrix dapat dilihat pada menu PerformanceVector di Description. Confusion matrix membagi data aktual ke dalam empat kategori, yaitu:

- TP: Data aktual yang benar-benar diklasifikasikan sebagai kompeten.
- TN: Data aktual yang benar-benar diklasifikasikan sebagai tidak kompeten.
- FP: Data aktual yang diklasifikasikan sebagai kompeten, tetapi sebenarnya tidak kompeten.
- FN: Data aktual yang diklasifikasikan sebagai tidak kompeten, tetapi sebenarnya kompeten.

Tabel 4. Confussion matrix

	True kompeten	True tidak kompeten
Pred. kompeten	TP	FP
Pred tidak kompeten	FN	TN

Peneliti menguji model klasifikasi dengan menggunakan data uji nilai siswa SMK Basuraga sebanyak 171 data. Dari hasil uji tersebut, diketahui bahwa model klasifikasi berhasil memprediksi 141 data kompeten dengan benar dan 27 data tidak kompeten dengan benar. Namun, model klasifikasi juga memprediksi 1 data kompeten sebagai tidak kompeten dan 2 data tidak kompeten sebagai kompeten. Hal ini dapat dilihat pada gambar 4.4.



Gambar 5. Confussion matrix

Berdasarkan tabel 4 dan gambar 5, maka rumus Confusion matrix yang menggambarkan nilai accuracy adalah:

$$\begin{aligned}
 \text{Akurasi} &= \frac{TP + TN}{TP + TN + FP + FN} * 100\% \\
 &= \frac{141 + 27}{141 + 27 + 1 + 2} * 100\% \\
 &= \frac{168}{171} * 100\% \\
 &= 98,24\%
 \end{aligned} \tag{1}$$

Nilai accuracy model klasifikasi adalah 0.9824, atau sekitar 98,24%. Artinya, model klasifikasi tersebut memprediksi dengan benar 98,24% dari data uji. Kemudian, peneliti akan menghitung nilai presisi kompeten. Presisi kompeten adalah ukuran seberapa baik model klasifikasi memprediksi siswa yang kompeten sebagai kompeten. Rumus untuk menghitung presisi kompeten adalah sebagai berikut:

$$\begin{aligned}
 \text{Precision} &= \frac{TP}{TP + FP} * 100\% \\
 &= \frac{141}{142} * 100\% \\
 &= 99,30\%
 \end{aligned} \tag{2}$$

Nilai presisi kompeten yang didapat adalah 0,9930 atau 99,30%. Artinya, model klasifikasi

tersebut berhasil memprediksi 99,30% dari data kompeten dengan benar. Selain menghitung nilai presisi kompeten, peneliti juga menghitung nilai presisi tidak kompeten. Presisi tidak kompeten adalah ukuran seberapa akurat model klasifikasi memprediksi siswa yang tidak kompeten. Rumus presisi tidak kompeten adalah sebagai berikut:

$$\begin{aligned} \text{Precision} &= \frac{TN}{TN + FN} * 100\% \\ &= \frac{27}{29} * 100\% \\ &= 93,10\% \end{aligned} \quad (3)$$

Presisi tidak kompeten sebesar 0,9310 menunjukkan bahwa model klasifikasi memprediksi siswa yang tidak kompeten dengan akurasi 93,10%. Artinya, model klasifikasi berhasil memprediksi 93,10% dari data tidak kompeten dengan benar. Selanjutnya, peneliti juga menggunakan perhitungan recall true kompeten untuk mengukur kinerja model klasifikasi untuk sampel true kompeten. Nilai recall yang tinggi menunjukkan bahwa model mampu mendeteksi dengan baik sampel true kompeten. Rumus menghitung recall true kompeten adalah sebagai berikut:

$$\begin{aligned} \text{Recall} &= \frac{TP}{TP + FN} * 100\% \\ &= \frac{141}{143} * 100\% \\ &= 98,60\% \end{aligned} \quad (4)$$

Nilai recall true kompeten sebesar 0,9860 menunjukkan bahwa model mampu mendeteksi 98,60% dari sampel true kompeten yang sebenarnya. Artinya, model mampu mendeteksi dengan baik siswa yang benar-benar kompeten. Kemudian, peneliti menghitung nilai recall true tidak kompeten. Recall true tidak kompeten menggambarkan seberapa baik model dapat mendeteksi siswa yang benar-benar tidak kompeten. Rumus menghitung recall true tidak kompeten adalah sebagai berikut:

$$\begin{aligned} \text{Recall} &= \frac{TN}{TN + FP} * 100\% \\ &= \frac{27}{28} * 100\% \\ &= 96,43\% \end{aligned} \quad (5)$$

Hasil menunjukkan bahwa model dapat mendeteksi 95% dari siswa yang benar-benar tidak kompeten. Hal ini dapat dilihat dari nilai recall true tidak kompeten yang sebesar 0,95. Recall true tidak kompeten adalah ukuran seberapa baik model dapat mendeteksi siswa yang benar-benar tidak kompeten.

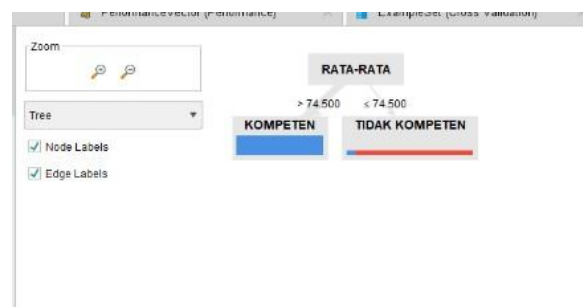
Setelah menyelesaikan proses perhitungan matriks kebingungan, diperoleh hasil akurasi sebesar 98,24%. Presisi prediksi untuk kategori kompeten mencapai 99,30%, sementara presisi prediksi untuk kategori tidak kompeten adalah 93,10%. Rincian prediksi kompeten yang sesuai dengan hasil sebanyak 141 data, dan terdapat 1 data yang diprediksi tidak

kompeten tetapi hasilnya kompeten. Sementara itu, terdapat 2 data yang diprediksi tidak kompeten tetapi hasilnya kompeten, dan 27 data yang diprediksi tidak kompeten dengan hasil tidak kompeten. Dengan demikian, tingkat keberhasilan untuk kategori kompeten adalah 98,60%, sedangkan untuk kategori tidak kompeten adalah 96,43%. Penjelasan ini dapat ditemukan pada Gambar 6.

	true KOMPETEN	true TIDAK KOMPETEN	class precision
pred. KOMPETEN	141	1	99.30%
pred. TIDAK KOMPETEN	2	27	93.10%
class recall	98.60%	96.43%	

Gambar 6. Hasil akurasi

Setelah melakukan perhitungan menggunakan Confusion Matrix, peneliti juga menghasilkan pohon keputusan melalui RapidMiner sebagai alat pengelola data mining. Gambaran pohon keputusan dapat dilihat pada gambar 7.



Gambar 7. Pohon keputusan

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil mengembangkan model klasifikasi menggunakan algoritma decision tree untuk memprediksi kompetensi siswa SMK Basuraga. Model ini mencapai tingkat akurasi sebesar 98,24%, yang menunjukkan bahwa model mampu memprediksi kompetensi siswa dengan akurasi yang tinggi. Selain itu, model ini juga menunjukkan nilai presisi yang tinggi, yaitu 99,30% untuk prediksi kompeten dan 93,10% untuk prediksi tidak kompeten. Nilai recall untuk kategori kompeten mencapai 98,60%, sementara recall untuk kategori tidak kompeten mencapai 96,43%. Hasil ini menunjukkan bahwa model mampu mendeteksi baik siswa yang kompeten maupun tidak kompeten dengan baik. Dengan demikian, implementasi metode klasifikasi menggunakan algoritma decision tree pada data siswa SMK Basuraga memberikan hasil yang sangat baik dalam memprediksi kompetensi siswa, dengan tingkat akurasi dan presisi yang tinggi.

DAFTAR PUSTAKA

- [1] M. R. Pasha, R. R. Hidayat, and M. I. Abas, "Implementasi Decision Tree C4 . 5 dalam Memilih Perguruan Tinggi Pendamping Program SMK Pusat Keunggulan," vol. 3, no. 6, pp. 1129–1139, 2023, doi: 10.30865/klik.v3i6.852.
- [2] F. F. Roji and D. Ramdani, "The Study of Decision Tree Algorithm For Classification of Student Graduation (Case Study: Faculty of Economics, University of Garut)," *RISTEC Res. Inf. Syst. Technol.*, vol. 2, no. 1, pp. 62–72, 2021, doi: 10.31980/ristec.v2i1.1215.
- [3] A. Tani, R. Helilintar, and A. B. Setiawan, "Sistem Penentuan Kelas Unggulan Menggunakan Metode Decision Tree Id3 di SMPN 6 Kota Kediri," vol. 7, pp. 1053–1060, 2023.
- [4] M. R. Qisthiano, P. A. Prayesy, and I. Ruswita, "Penerapan Algoritma Decision Tree dalam Klasifikasi Data Prediksi Kelulusan Mahasiswa," *G-Tech J. Teknol. Terap.*, vol. 7, no. 1, pp. 21–28, 2023, doi: 10.33379/gtech.v7i1.1850.
- [5] D. Sayhidin, G. Haris, and C. Juliane, "Implementasi Data Mining Tingkat Kepemimpinan Siswa dengan K-Nearest Neighbor, Decision Tree, dan Naïve Bayes," *J. Media Inform. Budidarma*, vol. 7, no. 1, pp. 199–206, 2023, doi: 10.30865/mib.v7i1.5351.
- [6] F. Ali Ma, A. Pratama, I. Sholihin, and A. Rizki Rinaldi, "Penerapan Model Prediksi Menggunakan Algoritma C.45 Untuk Prediksi Kelulusan Siswa SMK Wahidin," *J. Data Sci. Inform.*, vol. 1, no. 1, pp. 16–20, 2021.
- [7] N. Y. L. Gaol, "Prediksi Mahasiswa Berpotensi Non Aktif Menggunakan Data Mining dalam Decision Tree dan Algoritma C4.5," *J. Inf. Teknol.*, vol. 2, pp. 23–29, 2020, doi: 10.37034/jidt.v2i1.22.
- [8] I. R. Lutunani and A. Nugroho, "Analisis Prediksi Mahasiswa Terhadap Kelulusan Tepat Waktu Menggunakan Metode Data Mining Decision Tree (Studi Kasus: FTI UKSW)," *J. JTik (Jurnal Teknol. Inf. dan Komunikasi)*, vol. 7, no. 2, pp. 313–321, 2023, doi: 10.35870/jtik.v7i2.781.
- [9] A. I. Rizmayanti, N. Hidayati, F. S. Nugraha, and W. Gata, "Penerapan Data Mining Untuk Memprediksi Kompetensi Siswa Menggunakan Metode Decission Tree (Studi Kasus Smk Multicomp Depok)," *Swabumi*, vol. 9, no. 1, pp. 9–18, 2021, doi: 10.31294/swabumi.v9i1.8363.
- [10] Rovidatul, Y. Yunus, and G. W. Nurcahyo, "Perbandingan algoritma c4.5 dan naive bayes dalam prediksi kelulusan mahasiswa," *J. CoSciTech (Computer Sci. Inf. Technol.)*, vol. 4, no. 1, pp. 193–199, 2023, doi: 10.37859/coscitech.v4i1.4755.