

IMPLEMENTASI ALGORITMA C5.0 PADA PENILAIAN TINGKAT KEPUASAN PEMUSTAKA DALAM PENGGUNAAN DIGITAL LIBRARY (STUDI KASUS : UNISBA E-RESEARCH)

Nana Aidia Fitra Sania, Sri Lestanti, Yusniarsi Primasari

Teknik Informatika, Universitas Islam Balitar Blitar

Jalan Majapahit No.2-4 Blitar, Indonesia

frnanaa@gmail.com

ABSTRAK

Digital library merupakan pusat informasi dan pembelajaran yang penting untuk perguruan tinggi. Oleh karena itu, penting bagi perpustakaan untuk terus meningkatkan kualitas aplikasi. Untuk meningkatkan kualitas aplikasi, tentu hal itu tidak dapat dipisahkan dari pengalaman pemustaka dalam menggunakan *digital library* sebagai bahan rujukan perbaikan dan pengembangan aplikasi. Penelitian ini bertujuan untuk menguji tingkat kepuasan pemustaka dalam penggunaan layanan *digital library* menggunakan algoritma C5.0 dan teknik evaluasi menggunakan *confusion matrix*. Data diperoleh dari 251 responden melalui kuesioner yang menilai faktor-faktor seperti *effectiveness*, *efficiency*, dan *satisfaction* terhadap fitur-fitur *digital library*. Hasil analisis menunjukkan bahwa *effectiveness*, terutama dalam hal kemampuan mencari informasi dan navigasi yang intuitif, memiliki pengaruh signifikan terhadap kepuasan pengguna. Pengujian menunjukkan akurasi yang baik, dengan nilai mencapai 96,81%, serta *precision* dan *recall* yang memadai untuk label “puas” dan “tidak puas”. Hasil ini menunjukkan bahwa algoritma C5.0 yang diimplementasikan dengan RapidMiner dapat menghasilkan model yang akurat dalam menentukan kepuasan pemustaka.

Kata kunci : Implementasi, Algoritma C5.0, Tingkat Kepuasan, Pemustaka, Confusion Matrix

1. PENDAHULUAN

Kemajuan dalam teknologi informasi dan komunikasi telah mengubah berbagai aspek kehidupan, termasuk pendidikan tinggi. Salah satu perubahan signifikan adalah transisi perpustakaan konvensional menjadi perpustakaan digital. Digitalisasi perpustakaan ini mempermudah akses informasi dan mendukung kegiatan pendidikan serta penelitian [1].

Undang-Undang Republik Indonesia Nomor 43 Tahun 2007 mendefinisikan perpustakaan sebagai institusi pengelola koleksi karya tulis, cetak, dan rekam untuk memenuhi kebutuhan pendidikan, penelitian, pelestarian, informasi, dan rekreasi.

Perpustakaan berfungsi sebagai pusat informasi dan pembelajaran di semua tingkat pendidikan. Perpustakaan tidak hanya menyimpan informasi tetapi juga menyediakan sumber daya yang dapat diakses oleh masyarakat untuk berbagai keperluan. Dengan kemajuan teknologi, perpustakaan kini mengelola koleksi digital seperti *e-journal* dan *e-book* [2].

Proses pendidikan di perguruan tinggi tidak bisa dipisahkan dari aktivitas riset dan pengembangan, inovasi, dan implementasi pengetahuan. Maka, peranan pusat sumber informasi seperti perpustakaan di lembaga pendidikan tinggi memiliki signifikansi yang besar, menjadikan peran perpustakaan sebagai pusat vital dalam lingkungan universitas. UNISBA memiliki konsep *college library* yang merupakan bagian penting dari institusi pendidikan tinggi, bertujuan untuk mendukung berbagai program akademik yang disesuaikan dengan konsep Tri Dharma Perguruan Tinggi yaitu melibatkan aspek

edukasi, eksplorasi ilmiah, dan pelayanan kepada masyarakat.

Melihat fenomena digitalisasi ini, perpustakaan Universitas Islam Balitar (UNISBA) Blitar telah meluncurkan aplikasi *digital library* bernama UNISBA *E-Resourch* sejak 6 Mei 2023. Meskipun demikian, jumlah pengguna aplikasi ini masih terbatas. Data menunjukkan bahwa jumlah pengguna aplikasi masih rendah.

Berdasarkan wawancara yang dilakukan kepada salah satu Pustakawan Perpustakaan UNISBA Blitar, belum ada upaya yang dilakukan untuk meningkatkan kualitas aplikasi dan belum diketahui pada aspek atau bagian mana saja yang harus dibenahi dan ditingkatkan kualitasnya untuk meningkatkan jumlah pengguna pada aplikasi UNISBA *E-Resourch*.

Penelitian ini diharapkan dapat menjadi sumber acuan yang berharga untuk melakukan upaya peningkatan kualitas dan pengembangan *digital library* pada aplikasi UNISBA *E-Resourch*. Untuk meningkatkan pengalaman pemustaka dalam menggunakan *digital library*, diperlukan evaluasi yang dapat menilai kepuasan pemustaka dalam menggunakan platform *digital library*.

Algoritma C5.0 termasuk dalam metode *machine learning* yang dimanfaatkan untuk membuat model pengelompokan dengan menggunakan struktur pohon keputusan. Menurut Kherina dkk [3] pada penelitian “Penerapan Algoritma C5.0 Untuk Memprediksi Tingkat Kepuasan Siswa Terhadap Kinerja Guru MAN Simalungun” pada tahun 2023 menyatakan bahwa algoritma C5.0 dipilih karena kemampuannya

dalam melakukan klasifikasi dengan akurasi mencapai 83,33%.

2. TINJAUAN PUSTAKA

2.1. Implementasi

Implementasi merujuk pada pelaksanaan berbagai aktivitas, langkah, eksekusi, atau penerapan sistem tertentu. Implementasi tidak hanya tentang melakukan sesuatu, melainkan sebuah proses terstruktur yang bertujuan untuk mencapai hasil yang diinginkan [4].

2.2. Data Mining

Secara sederhana, data mining adalah proses ekstraksi informasi yang penting namun tersirat dan belum diketahui sebelumnya. Selain itu, data mining terkait erat dengan beragam bidang, termasuk statistik, pembelajaran mesin, pengenalan pola, algoritma komputasi, teknologi basis data, dan komputasi berkinerja tinggi [5].

2.3. Klasifikasi

Klasifikasi adalah tahap mencari model atau fungsi yang bisa menjelaskan serta membedakan berbagai kelas data atau konsep. Hal ini bertujuan agar kita bisa memprediksi kelas suatu objek yang belum kita ketahui dengan menggunakan model tersebut [6].

2.4. Algoritma C5.0

Dalam bidang data mining, algoritma C5.0 termasuk ke dalam algoritma untuk mengumpulkan data yang umum dipakai dalam penemuan pola dataset. Tujuannya ialah untuk mengidentifikasi pola atau anomali yang mungkin ada dalam data tersebut [7].

Algoritma C5.0 adalah metode pengklasifikasian yang diterapkan ke penambangan data, utamanya pada teknik pohon keputusan. Metode tersebut merupakan versi yang lebih baik dari algoritma ID3 dan C4.5, yang dikembangkan oleh Ross Quinlan 1987 [8].

Algoritma C5.0 dikenal sebagai *decision tree* yang termasuk ke dalam algoritma pengklasifikasian. Algoritma ini melakukan klasifikasi data dalam bentuk pohon bercabang yang terdiri dari beberapa node, seperti *root node*, yaitu node di bagian paling atas dari pohon keputusan yang tidak berpanah atau *edge* yang menunjuk ke arahnya. Selain itu, terdapat juga internal node, dengan panah atau *edge* yang masuk dan keluar dari node tersebut, serta terminal node (*leaf*), yaitu node di ujung bawah pohon keputusan yang hanya memiliki panah atau *edge* yang menunjuk ke arahnya tanpa panah atau *edge* yang keluar dari node tersebut [9].

Perhitungan nilai *entropy* dapat ditemukan melalui persamaan (1):

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

Di mana $\log_2 p_i$ didapat dari persamaan (2):

$$\log(X) = \frac{\ln(X)}{\ln(2)} \quad (2)$$

Keterangan:

S : Himpunan Kasus

n : jumlah partisi S

p_i : Proporsi dari S_i terhadap S

Gain ratio menjadi fokus utama yang dinyatakan dalam persamaan (3):

$$GainRatio(A) = \frac{Gain(A)}{SplitInfo(A)} \quad (3)$$

Di mana nilai *Gain* di dapat dari persamaan (4):

$$Gain(S,A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (4)$$

Keterangan :

S : Himpunan Kasus

A : Atribut

n : jumlah partisi atribut A

$|S_i|$: jumlah kasus pada partisi ke-i

$|S|$: jumlah kasus dalam S

Untuk mencari nilai *SplitInfo* digunakan persamaan (5):

$$SplitInfo(S) = - \sum_{i=1}^n \frac{|S_i|}{|S|} * \log_2 \frac{|S_i|}{|S|} \quad (5)$$

Keterangan :

S : Himpunan Kasus

A : Atribut

n : jumlah partisi atribut A

$|S_i|$: jumlah kasus pada partisi ke-i

$|S|$: jumlah kasus dalam S

2.5. Pemustaka

Pemustaka ialah mereka yang memanfaatkan layanan perpustakaan, bisa berupa individu, kelompok, komunitas, atau entitas lain. Kehadiran mereka di perpustakaan dilandaskan pada berbagai alasan, tujuan, dan harapan, dengan upaya untuk memperoleh manfaat yang diinginkan melalui pengalaman yang memuaskan [10].

2.6. Digital Library (Perpustakaan Digital)

Istilah 'perpustakaan digital' ada pada bulan Juli 1945 saat Vannevar Bush menyadari bahwa akses terhadap informasi terpublikasi terkendala oleh model cetak manual. Konsep ini disampaikannya dalam tulisannya yang berjudul "As We May Think". Konsep "perpustakaan digital" memiliki interpretasi yang bervariasi dan luas. Definisi-definisi yang berbeda berkembang seiring dengan perspektif individu terhadap konsep tersebut. Berbagai istilah seperti "perpustakaan digital," "perpustakaan virtual," "perpustakaan tanpa batas," "perpustakaan elektronik," "perpustakaan cyber," "cybrary," dan "ebrary" sering kali dianggap memiliki makna yang sama dengan istilah "perpustakaan digital" [11].

2.7. Rapidminer

RapidMiner adalah perangkat lunak yang populer untuk melakukan proses penambangan data karena menyediakan beragam algoritma untuk klasifikasi, pengelompokan, dan analisis regresi. Selain itu, antarmuka pengguna RapidMiner sangat intuitif

dengan hanya tiga tampilan yang disediakan yaitu *welcome*, *design*, dan *result* [12].

Awalnya RapidMiner disebut YALE (Yet Another Learning Environment), kemudian RapidMiner berkembang sejak tahun 2001 oleh tim yang terdiri dari Ralf Klinkenberg, Ingo Mierswa, dan Simon Fischer di Unit Kecerdasan Buatan dari University of Dortmund. Produk ini dijual menggunakan lisensi AGPL (GNU Affero General Public License) versi 3 [13].

2.8. Confusion matrix

Confusion matrix yakni sebuah struktur data berbentuk tabel yang diterapkan dalam penilaian efektifitas suatu model pengelompokan atau sistem prediksi. Dalam *confusion matrix* terdapat metrik-metrik seperti jumlah kasus yang terklasifikasi dengan benar sebagai positif (*true positive/TP*), negatif (*true negative/TN*) dan jumlah kasus yang salah diklasifikasi sebagai positif (*false positive/FP*) dan negatif (*false negative/FN*) oleh model yang digunakan.

3. METODE PENELITIAN

3.1. Teknik pengumpulan data

3.3.1. Wawancara

Wawancara digunakan untuk mengumpulkan data pengguna *digital library*, dengan narasumber Kepala Perpustakaan UNISBA Blitar. Didapatkan atribut penilaian yaitu *learnability*, *effectiveness*, *efficiency*, *memorability*, *satisfaction*.

3.3.2. Kuesioner

Penelitian ini menggunakan *platform google form* yang disebarkan kepada responden untuk mendapatkan penilaian objektif responden terhadap *digital library* UNISBA E-Resourch.

3.3.3. Studi pustaka

Studi pustaka yakni suatu tahapan bagi para peneliti melakukan eksplorasi mendalam guna menemukan teori-teori yang relevan melalui sumber-sumber seperti jurnal, karya ilmiah, skripsi, dan buku.

3.2. Jenis data

3.2.1. Data Primer

Data primer ialah informasi yang didapatkan secara langsung oleh peneliti [14]. Dalam penelitian ini, peneliti memanfaatkan data primer sebagai informasi utama yang diperoleh melalui distribusi kuisisioner menggunakan *platform Google Form*. Dalam penelitian ini, digunakan 251 responden.

3.2.2. Data Sekunder

Data sekunder merujuk pada keterangan yang diambil dari sesuatu yang telah ada [14]. Dalam penelitian ini, informasi sekunder yang dimanfaatkan merupakan data pustaka yang berasal dari perpustakaan UNISBA Blitar serta informasi terkait *digital library* UNISBA E-Resourch.

3.3. Tahapan penelitian

3.3.1. Analisa masalah

Analisa ini bertujuan untuk mengidentifikasi masalah yang terkait dengan tingkat kepuasan pemustaka dalam menggunakan *digital library* di perpustakaan UNISBA Blitar serta untuk memahami kesulitan yang mungkin dihadapi oleh pemustaka dalam menggunakan dan mengakses fitur-fitur yang tersedia di situs layanan online perpustakaan

3.3.2. Studi literatur

Mempelajari dan memahami berbagai literatur yang relevan dengan topik penelitian, seperti algoritma C5.0.

3.3.3. Pengumpulan data

Pengumpulan data diterapkan dengan penggunaan kuisisioner digital melalui platform Google Form yang dipublikasikan kepada pemustaka perpustakaan UNISBA Blitar.

3.4. Pengolahan data

Setelah data terkumpul, tahap berikutnya yaitu memproses data tersebut untuk menciptakan sebuah model klasifikasi dengan algoritma C5.0 yang merujuk ke diagram alur C5.0. Adapun atribut penilaian yang digunakan untuk mengetahui faktor-faktor tingkat kepuasan pemustaka sebagai berikut:

Tabel 1. Atribut penilaian

Atribut	Pernyataan
<i>Learnability</i>	Kemudahan bagi pengguna dalam memahami suatu sistem, menyelesaikan tugas-tugas dasarnya saat pertama kali digunakan, serta upaya dan waktu yang dibutuhkan agar pengguna baru dapat berinteraksi secara efektif dan mencapai hasil terbaik secara optimal.
<i>Effectiveness</i>	Ketepatan dan kesempurnaan sistem dalam mencapai sasaran yang ditetapkan.
<i>Efficiency</i>	Keterampilan efisiensi dan jumlah sumber daya yang diperlukan oleh pengguna dalam menuntaskan aktivitasnya.
<i>Memorability</i>	Tingkat kesederhanaan dalam mengingat, saat menggunakan sistem, bahkan setelah hanya dipelajari sekali atau tidak digunakan dalam periode yang lama.
<i>Satisfaction</i>	Kepuasan pengguna dalam menggunakan produk atau sistem dalam situasi tertentu disesuaikan dengan kebutuhan mereka.

Tabel 2. Kelas yang digunakan oleh setiap atribut

Atribut	Kelas Yang Digunakan
<i>Learnability</i>	Sangat tidak setuju, Tidak setuju, Setuju, Sangat setuju
<i>Effectiveness</i>	Sangat tidak setuju, Tidak setuju, Setuju, Sangat setuju
<i>Efficiency</i>	Sangat tidak setuju, Tidak setuju, Setuju, Sangat setuju

Atribut	Kelas Yang Digunakan
Memorability	Sangat tidak setuju, Tidak setuju, Setuju, Sangat setuju
Satisfaction	Sangat tidak setuju, Tidak setuju, Setuju, Sangat setuju

Tabel 3. Kelas yang digunakan oleh setiap atribut

Kelas	Kode	Nilai
Sangat tidak setuju	STS	4-8

Kelas	Kode	Nilai
Tidak setuju	TS	9-12
Setuju	S	13-15
Sangat setuju	SS	16

Tabel 4. Ketentuan nilai pelabelan

Label	Keterangan
>64	Puas
<64	Tidak Puas

4. HASIL DAN PEMBAHASAN

4.1. Pengumpulan data

Tabel 5. Hasil kuesioner

No	Learnability	Effectiveness	Efficiency	Memorability	Satisfaction	Total
1	14	13	14	13	12	66
2	12	12	12	12	12	60
3	14	12	11	12	11	60
...	...	...	...	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...
249	12	13	15	12	15	67
250	12	12	12	12	12	60
251	12	11	11	12	12	58

4.2. Pengolahan data

Tabel 6. Transformasi hasil kuesioner

No	Learnability	Effectiveness	Efficiency	Memorability	Satisfaction	Label
1	S	S	S	S	TS	Puas
2	TS	TS	TS	TS	TS	Tidak Puas
3	S	TS	TS	TS	TS	Tidak Puas
...	...	...	...	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...
249	TS	S	S	TS	S	Puas
250	TS	TS	TS	TS	TS	Tidak Puas
251	TS	TS	TS	TS	TS	Tidak Puas

4.3. Perhitungan algoritma C5.0

- a. Hitung nilai *entropy* total

$$\text{Entropy total} = \left(-\frac{96}{251} \times \log_2 \frac{96}{251} \right) + \left(-\frac{155}{251} \times \log_2 \frac{155}{251} \right) = 0.9597$$

- b. Hitung *entropy* untuk setiap atribut pada setiap kelas datanya.

Di bawah ini merupakan perhitungan *entropy* pada atribut *learnability*:

$$\text{Entropy SS} = \left(-\frac{30}{31} \times \log_2 \frac{30}{31} \right) + \left(-\frac{1}{31} \times \log_2 \frac{1}{31} \right) = 0.2055$$

$$\text{Entropy S} = \left(-\frac{19}{130} \times \log_2 \frac{19}{130} \right) + \left(-\frac{111}{130} \times \log_2 \frac{111}{130} \right) = 0.9372$$

$$\text{Entropy TS} = \left(-\frac{13}{134} \times \log_2 \frac{13}{134} \right) + \left(-\frac{121}{134} \times \log_2 \frac{121}{134} \right) = 0.4594$$

- c. Hitung *gain* pada setiap atribut.

Di bawah ini merupakan perhitungan *gain* pada atribut *learnability*:

$$\text{Gain} = (0.9597) - \left(\left(\frac{31}{251} \right) \times 0.2055 \right) - \left(\left(\frac{82}{251} \right) \times 0.9372 \right) - \left(\left(\frac{134}{251} \right) \times 0.4594 \right) = 0.3828$$

- d. Hitung *split info* pada setiap atribut.

Di bawah ini merupakan perhitungan *split info* pada atribut *learnability*:

$$\text{Split Info} = \left(-\frac{31}{251} \times \log_2 \frac{31}{251} \right) + \left(-\frac{82}{251} \times \log_2 \frac{82}{251} \right) + \left(-\frac{134}{251} \times \log_2 \frac{134}{251} \right) + \left(-\frac{4}{251} \times \log_2 \frac{4}{251} \right) = 1.4784$$

- e. Hitung *gain ratio* pada setiap atribut.

Di bawah ini merupakan perhitungan *gain ratio* pada atribut *learnability*:

$$\text{Gain Ratio} = \frac{0.3828}{1.4784} = 0.2589$$

Tabel 7. Hasil perhitungan *entropy*, *gain*, *split info*, dan *gain ratio*

		Jumlah	Puas	Tidak Puas	Entropy	Gain	Split Info	Gain Ratio
Total		251	96	155	0,9597			
Learnability						0,3828	1,4784	0,2589
	SS	31	30	1	0,2055			
	S	82	53	29	0,9372			
	TS	134	13	121	0,4594			
	STS	4	0	4	0			
Effectiveness						0,4737	1,4108	0,3358
	SS	21	20	1	0,2761			
	S	93	69	24	0,8238			
	TS	133	7	126	0,2974			
	STS	4	0	4	0			
Efficiency						0,4479	1,4064	0,3185
	SS	23	23	0	0			
	S	83	61	22	0,8342			
	TS	141	12	129	0,4199			
	STS	4	0	4	0			
Memorability						0,3777	1,3795	0,2737
	SS	14	14	0	0			
	S	80	62	18	0,7691			
	TS	148	20	128	0,5713			
	STS	9	0	9	0			
Satisfaction						0,3563	1,3956	0,2553
	SS	18	18	0	0			
	S	75	56	19	0,8165			
	TS	150	22	128	0,6014			
	STS	8	0	8	0			

Setelah nilai *gain ratio* tertinggi ditemukan pada salah satu atribut, atribut tersebut dijadikan simpul akar utama. Dalam perhitungan di atas, yang menjadi akar utama adalah atribut *effectiveness*. Kemudian,

terdapat empat cabang yang muncul, yaitu SS, S, TS, dan STS dimana STS menghasilkan 4 orang tidak puas.

Tabel 8. Simpul akar *effectiveness* kelas data "ss"

No	Learnability	Effectiveness	Efficiency	Memorability	Satisfaction	Label
1	SS	SS	SS	SS	SS	Puas
2	S	SS	S	SS	SS	Puas
3	SS	SS	SS	S	S	Puas
4	S	SS	S	S	S	Puas
5	SS	SS	SS	S	TS	Puas
6	TS	SS	TS	TS	TS	Tidak Puas
7	SS	SS	SS	SS	SS	Puas
8	S	SS	S	S	S	Puas
9	SS	SS	SS	SS	SS	Puas
10	SS	SS	SS	S	SS	Puas
...	...	...	...	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...
15	S	SS	S	S	S	Puas
16	SS	SS	SS	S	TS	Puas
17	SS	SS	SS	SS	SS	Puas
18	SS	SS	SS	SS	SS	Puas
19	SS	SS	S	SS	SS	Puas
20	SS	SS	SS	S	S	Puas
21	S	SS	SS	S	S	Puas

Hitung nilai *entropy* total pada simpul akar *effectiveness* kelas data "SS":

$$\text{Entropy total} = \left(-\frac{20}{21} \times \log_2 \frac{20}{21}\right) + \left(-\frac{1}{21} \times \log_2 \frac{1}{21}\right) = 0,2761$$

Hitung *entropy* pada atribut *learnability*:

$$\text{Entropy TS} = \left(-\frac{2}{3} \times \log_2 \frac{2}{3}\right) + \left(-\frac{1}{3} \times \log_2 \frac{1}{3}\right) = 0,9182$$

Setelah itu, hitung *gain*, *split info*, dan *gain ratio* pada atribut *learnability*:

$$Gain = (0,2761) - \left(\left(\frac{3}{21}\right) * 0.9182\right) = 0.1450$$

$$Gain Ratio = \frac{0.1450}{1.3223} = 0.1096$$

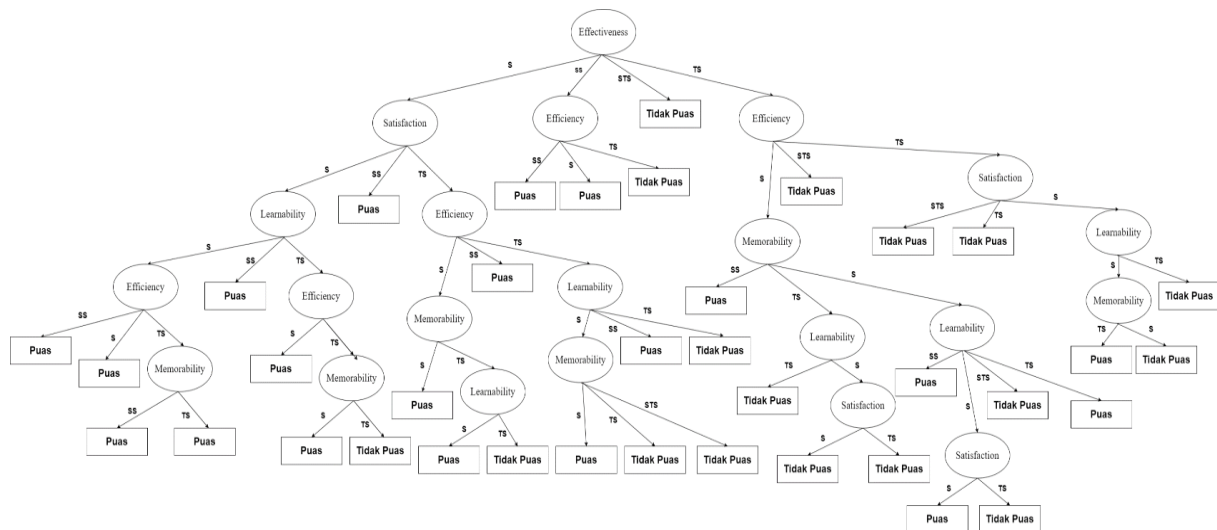
$$Split Info = \left(-\frac{13}{21} \times \log_2 \frac{13}{21}\right) + \left(-\frac{5}{21} \times \log_2 \frac{5}{21}\right) + \left(-\frac{3}{21} \times \log_2 \frac{3}{21}\right) = 1.3223$$

Tabel 9. Hasil perhitungan *entropy*, *gain*, *split info*, dan *gain ratio*

		Jumlah	Puas	Tidak Puas	Entropy	Gain	Split Info	Gain Ratio
Total		21	20	1	0,2761			
Learnability						0,1450	1,3223	0,1096
	SS	13	13	0	0			
	S	5	5	0	0			
	TS	3	2	1	0,9182			
Efficiency						0,9182	4,3357	0,2117
	SS	13	13	0	0			
	S	7	7	0	0			
	TS	1	0	1	0			
Memorability						0,1809	1,3421	0,1348
	SS	8	8	0	0			
	S	11	11	0	0			
	TS	2	1	1	1			
Satisfaction						0,1216	1,49370	0,0814
	SS	10	10	0	0			
	S	7	7	0	0			
	TS	4	3	1	0,8112			

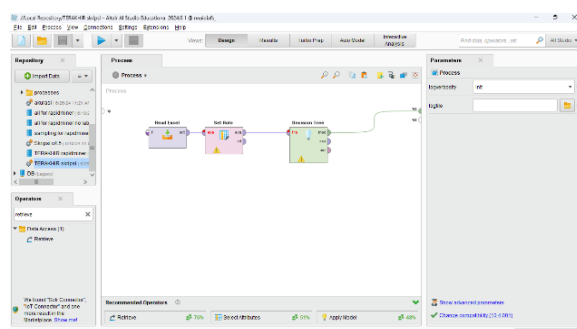
Perhitungan tersebut dilakukan sampai mencapai keputusan akhir yang terbentuk pada node terakhir, sehingga akan diperoleh hasil pohon keputusan seperti

yang ditunjukkan pada gambar 4.1 dan gambar 4.2 dengan simpul akar *effectiveness*.



Gambar 1. Hasil pohon keputusan

4.4. RapidMiner



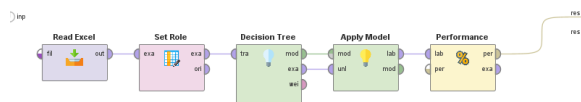
Gambar 2. Skema pembentukan pohon keputusan

Lakukan proses pemodelan untuk membuat pohon keputusan dengan mencari operator *read excel*, *set role*, dan *decision tree*. Hubungkan ketiga operator sehingga terlihat seperti pada gambar 2. Pastikan parameter pada operator *decision tree* adalah *gain_ratio* karena *gain_ratio* merupakan parameter untuk membuat pohon keputusan dengan algoritma C5.0. Setelah selesai, *runing* semua operator yang telah terhubung sehingga dapat memperlihatkan hasil pohon keputusannya.



Gambar 3. Hasil pohon keputusan menggunakan rapidminer

4.5. Confusion Matrix



Gambar 4. Skema uji akurasi

Dalam uji akurasi confusion matrix, terdapat 5 operator yang digunakan yaitu *read excel*, *set role*, *decision tree*, *apply model*, dan *performance*. Operator *performance* dapat menghitung berbagai metrik evaluasi seperti akurasi, presisi, *recall*, dan F1-score. Metrik ini membantu mengevaluasi seberapa baik model berkinerja dalam melakukan klasifikasi dan memahami kekuatan dan kelemahan model tersebut.

accuracy: 96,81%			
	True Pos	True Total Pos	class precision
pred Pos	93	5	94,99%
pred Total Pos	3	150	98,04%
class recall	96,88%	95,77%	

Gambar 5. Hasil uji akurasi

Akurasi yang dihasilkan adalah 96,81%, yang berarti model berhasil memprediksi dengan benar sebesar 96,81% dari total data yang dievaluasi. Dan memiliki dua kelas yaitu "Puas" dan "Tidak Puas". Berdasarkan hasil prediksi model, terdapat 93 data yang diprediksi sebagai "Puas" dengan benar atau disebut dengan TP (*true positive*), 150 data yang diprediksi sebagai "Tidak Puas" dengan benar atau yang disebut juga dengan TN (*true negative*). Namun, terdapat 3 data yang salah diprediksi sebagai "Tidak Puas" padahal seharusnya "Puas" atau yang disebut juga dengan FN (*false negative*), dan 5 data yang salah diprediksi sebagai "Puas" padahal seharusnya "Tidak Puas" atau yang disebut juga dengan FP (*false positive*). Dengan demikian, hasil evaluasi performa model ini menunjukkan akurasi yang cukup tinggi, di mana model mampu memprediksi dengan benar sebagian besar kategori "Puas" dan "Tidak Puas".

5. KESIMPULAN DAN SARAN

Berdasarkan penelitian ini, data dari 251 responden pengguna *digital library* dianalisis menggunakan algoritma C5.0 dengan Rapidminer untuk menilai tingkat kepuasan pemustaka. Hasilnya menunjukkan bahwa tingkat kepuasan dipengaruhi oleh faktor-faktor seperti *effectiveness*, *efficiency*,

satisfaction, *learnability*, dan *memorability*. Faktor *effectiveness*, terutama dalam hal kemudahan navigasi dan pencarian informasi, merupakan yang paling signifikan. Sementara *efficiency*, *learnability*, *memorability*, dan *satisfaction* juga berperan penting. Pengujian akurasi algoritma C5.0 menunjukkan hasil yang tinggi dengan akurasi 96,81%, *precision* 94,90% untuk label puas, dan *precision* 98,04% untuk label tidak puas, serta *recall* yang seimbang pada kedua label. Hasil ini menegaskan bahwa algoritma C5.0 efektif dalam menilai kepuasan pengguna *digital library*. Berdasarkan hasil penelitian, disarankan agar penelitian selanjutnya menggunakan dataset yang lebih besar dan beragam dari berbagai perpustakaan digital untuk meningkatkan generalisasi dan keandalan model.

DAFTAR PUSTAKA

- [1] G. Crane, "The Perseus Digital Library and the future of libraries," *Int. J. Digit. Libr.*, vol. 24, no. 2, pp. 117–128, 2023, doi: 10.1007/s00799-022-00333-2.
- [2] W. Widiyanto, "Perpustakaan digital dalam pendidikan tinggi: Analisis bibliometrik," *J. Sastra Indones.*, vol. 12, no. 3, pp. 234–242, 2023, doi: 10.15294/jsi.v12i3.74695.
- [3] Kherina Surya Ningsih and I. Zufria, "Penerapan Algoritma C5.0 Untuk Memprediksi Tingkat Kepuasan Siswa Terhadap Kinerja Guru MAN Simalungun," *J. Fasilkom*, vol. 13, no. 3, pp. 406–415, 2023, doi: 10.37859/jf.v13i3.6347.
- [4] I. Anggraeni, "Pengertian Implementasi dan PendapaT Ahli," *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 16–36, 2019.
- [5] M. A. Muslim *et al.*, "Data Mining Algoritma C4.5," pp. 7823–7830, 2019.
- [6] L. N. Rani, "Klasifikasi Nasabah Menggunakan Algoritma C4.5 Sebagai Dasar Pemberian Kredit," *INOVTEK Polbeng - Seri Inform.*, vol. 1, no. 2, p. 126, 2016, doi: 10.35314/isi.v1i2.131.
- [7] H. Siska, S. Aji, and S. Eko, "Implementasi Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma C4.5 (Studi Kasus: Universitas Dehasen Bengkulu)," *J. Media Infotama*, vol. 11, no. 2, pp. 130–138, 2020.
- [8] D. J. Purba, M. R. Lubis, and Z. A. Siregar, "Analisis Kepuasan Pelanggan Menggunakan Algoritma C5.0," *J. Informatics Manag. Inf. Technol.*, vol. 2, no. 2, pp. 115–121, 2022, doi: 10.54367/jtiust.v7i1.1983.
- [9] N. Benediktus and R. S. Oetama, "Algoritma Klasifikasi Decision Tree C5.0 untuk Memprediksi Performa Akademik Siswa Natanael," *Ultim. J. Tek. Inform.*, vol. 12, no. 1, pp. 14–19, 2020.
- [10] E. Puspa, "Analisis Kepuasan Pemustaka Terhadap Pelayanan Perpustakaan Pusat Penelitian Dan Pengembangan Perikanan Budidaya," *J. Pari*, vol. 2, no. 2, p. 113, 2017,

- doi: 10.15578/jp.v2i2.3256.
- [11] R. Sayekti and Mardianto, *Buku Peprustakaan Digital Lengkap*. 2019.
- [12] H. Hijrah, "Analisis Perbandingan Aplikasi Data Mining Dalam Memprediksi Kualitas Kinerja Karyawan Menggunakan Metode Algoritma C4.5," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 9, no. 2, pp. 1655–1665, 2022, doi: 10.35957/jatisi.v9i2.1992.
- [13] W. R. Ningsih, "Penerapan Data Mining Dalam Menentukan Masa Studi Mahasiswa Menggunakan Algoritma C4.5," *Range Manag. Agrofor.*, vol. 4, no. 1, pp. 1–15, 2020, doi: 10.1016/j.fcr.2017.06.020.
- [14] I. Nurdin and S. Hartati, *Metodologi Penelitian Sosial*. 2019.