PENERAPAN DATA MINING MENGGUNAKAN METODE ALGORITMA C4.5 DAN *NAÏVE BAYES* DALAM MENGUKUR TINGKAT KEPUASAN DI RSUD HASRI AINUN HABIBIE

Nurul Mustapa[1], Tri Pratiwi Handayani[2], Wahyudin Hasyim[3]

[1,2,3] Program Studi Sistem Informasi, Universitas Muhammadiyah Gorontalo, *Jl. Prof. Dr. Mansoer Pateda, Kab.Gorontalo , India Email :*

nurulmustapa9@gmail.com

ABSTRACT

NURUL MUSTAPA. "Implementation of Data Mining Using C4.5 Algorithm and Naïve Bayes to Measure Patient Satisfaction at Hasri Ainun Habibie Hospital" Supervised by Tri Pratiwi Handayani, S.Kom, M.Phil, M.Eng as the lead supervisor and Wahyudin Hasyim, M.Kom as the member.

This study aimed to measure patient satisfaction at Hasri Ainun Habibie Hospital. It employed the C4.5 Algorithm and Naïve Bayes methods and compared their accuracy. The study used the Slovin formula to determine a sample of 376 respondents. The author directly observed and monitored the research object using a questionnaire instrument and Likert scale. The analysis was performed using the Rapid Miner application. The results indicated that the Naïve Bayes algorithm method yielded an accuracy rate of 93.38% in measuring patient satisfaction at Hasri Ainun Habibie Hospital.

Keywords: Data Mining, C4.5 Algorithm, Naïve Bayes, Decision Tree, Patient Satisfaction

ABSTRAK

NURUL MUSTAPA. "Penerapan Data Mining Menggunakan Metode Algoritma C4.5 Dan *Naïve Bayes* dalam Mengukur Tingkat Kepuasan Pasien di RSUD Hasri Ainun Habibie". Di bombing oleh Tri Pratiwi Handayani,S.Kom,M.Phil, M.Eng sebagai ketua dan Wahyudin Hasyim,M.Kom sebagai anggota.

Penelitian ini bertujuan untuk mengukur tingkat kepuasan pasien di rumah sakit Hasri Ainun Habibie. Penelitian ini menggunakan metode Algoritma C4. 5 dan *Naïve Bayes* dengan melihat akurasi dan perbandingan. Penelitian ini menggunakan rumus slovin sampel yang diambil sebagai responden sebanyak 376. Pengamatan dan pemantauan dilakukan secara langsung oleh penulis pada objek penelitian dengan menggunakan instrument penelitian berupa kuisioner dan menggunakan skala likert. Pengujian ini dilakukan menggunakan Aplikasi *Rapid Miner*. Hasil penelitian ini menunjukkan bahwa tingkat kepuasan pasien di RSUD Hasri Ainun Habibie bahwa penggunaan metode algoritma *Naïve Bayes* menghasilkan nilai akurasi sebesar 93,38%.

Kata Kunci : Data Mining, Algoritma C4.5, *Naïve Bayes*, Pohon Keputusan dan Kepuasan Pasien

1. Latar Belakang

RSUD Hasri Ainun Habibie merupakan rumah sakit rujukan di Provinsi Gorontalo. RSUD Hasri Ainun Habibie adalah rumah sakit milik Pemerintah Provinsi Gorontalo yang berdiri sejak tanggal 3 november 2013 yang berada di Jl. Ahmad A. Wahab, Dutulanaa, Kec. Limboto, Kabupaten Gorontalo. Saat ini RSUD memiliki 548 pegawai dan jumlah pasien yang sudah terlayani dengan jumlah 11.176. Rumah sakit merupakan salah satu institusi kesehatan yang memberikan pelayanan kesehatan kepada masyarakat. Hal ini memainkan peran strategis dalam mempercepat peningkatan kesehatan masyarakat. Kepuasan pasien sangat penting pada tingkat evaluasi pelayanan rumah sakit bagi pasien. Karena kepuasan pasien merupakan indikator terpenting bagi rumah sakit.

RSUD Hasri Ainun Habibie telah melakukan Survei Kepuasan Masyarakat (SKM) menggunakan pengukuran secara kuantitatif dan kualitatif atas pendapat masyarakat. Jika dilihat dari perkiraan jumlah penerima layanan tahun 2022, maka populasi penerima layanan pada RSUD Hasri Ainun Habibie Provinsi Gorontalo dalam kurun waktu satu tahun adalah sekitar 36.372 orang. Pada tahun 2023 rumah sakit telah melaksanakan Survei Kepusan Masyarakat (SKM) selama 1 periode atau Triwulan 1, yakni Januari hingga Maret 2023. Selanjutnya responden dipilih secara acak dari setiap jenis pelayanan besaran sampel dan populasi menggunakan tabel sampel dari Krejcie and Morgan.

2. LANDASAN TEORI

1. Kepuasan

   Kepuasan konsumen yaitu tingkat perasaan seseorang Setelah membandingkan kinerja produk yang dia rasakan dengan harapanya.(Indrasari meithiana,2019) dalam kutipan jurnal (Sesrianty et al., 2019)

2. Kualitas Pelayanan

   Kualitas pelayanan merupakan ketepatan usaha dalam memenuhi kebutuhan dan keinginan konsumen untuk menyeimbangi harapan konsumen. Tjiptono (2014:268) Dengan dimensi *Reliability* (kehandalan), *Assurance* (jaminan), *Tangibles* (tampilan/bukti fisik), *Empathy* (empati), *Responsiveness* (ketanggapan dan kepedulian)Dalam kutipan jurnal (Sofi Adinda Amri Nasution & Yayuk Yuliana, 2022)

3. Data Mining

   Data mining merupakan serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data. Data mining mulai ada sejak 1990-an sebagai cara yang benar dan tepat untuk mengambil pola dan informasi yang digunakan untuk menemukan hubungan antara data untuk melakukan pengelompokkan ke dalam satu atau lebih cluster sehingga objek-objek yang berada dalam satu cluster akan mempunyai kesamaan yang tinggi antara satu dengan lainnya. Data mining merupakan bagian dari proses penemuan pengetahuan dari basis data Knowledge Discovery in Databases (Alkhairi *et al., 2019*) dalam kutipan jurnal (Tarigan et al., 2022)

4. Klasifikasi

   Menurut Kurniawan (2018) klasifikasi ini merupakan bentuk dasar dari analisis data, sedangkan menurut (Bansal, *et al., 2017)* klasifikasi yaitu teknik yang digunakan untuk keanggotaan kelompok menurut data-data yang sudah tersedia. Dapat dikatakan juga klasifikasi ini merupakan salah satu problematika mendasar dan tugas utama pada data mining (Harry*et al., 2004* dalam Syarli, 2016).

Klasifikasi dapat digunakan untuk beragam kasus sebagai pembentukan aturan terhadap sebuah data. Dalam kutipan jurnal (Solikhah et al., 2021)

5. Rumus Slovin

Sebagai penelitian yang berdasarkan pada filsafat positivisme, digunakan untuk meneliti populasi atau sampel tertentu, metode penelitian berupa angka-angka dan analisis menggunakan statistik. Jika dilihat dari perkiraan jumlah penerima layanan tahun 2022, maka populasi penerima layanan pada RSUD Hasri Ainun Habibie sebanyak 6367 orang. Selanjutnya responden dipilih secara acak dari setiap jenis pelayanan besaran sampel dan populasi menggunakan rumus Slovin dan jumlah minimum sampel responden yang harus dikumpulkan adalah 376 responden. Adapun teknik pengambilan sampel menggunakan rumus :

$$n = \frac{N}{N.d2+1} \qquad (1)$$

Keterangan :

N = Jumlah populasi

n = Jumlah sampel

D = Batas toleransi kesalahan 10% (0,1)

6. Algoritma C4.5

Algoritma C4.5 merupakan algoritma yang dikembangkan dari algoritma ID3. C4.5 ini merupakan algoritma turunan dari algoritma ID3 dengan beragam peningkatan. Beberapa peningkatan ini diantaranya adalah, penanganan atribut-atribut numerik, missing value dan noise pada dataset, dan aturan-aturan yang dihasilkan dari model pohon yang terbentuk (Larasati *et al.,2018*) dalam kutipan jurnal (Nasrullah, 2021)

7. Skala Likert

Dalam menganalisis data yang diperoleh dari kuisioner, digunakan Skala Likert. Menurut Nazir (2014) bahwa Skala Likert adalah suatu skala psikometrik yang digunakan dalam kuisioner dan merupakan salah satu teknik yang dapat digunakan dalam evaluasi suatu program atau kebijakan perencanaan. Skala Likert digunakan untuk mengukur kesetujuan dan ketidaksetujuan seseorang terhadap sesuatu rencana program, pelaksanaan program ataupun tingkat keberhasilan suatu program.

Berdasarkan kuisioner dengan model close-ended question dan dengan skala Likert untuk mengetahui tingkat kepuasan pasien yang terdiri dari pilihan (optional) dan skornya yaitu :

Tabel 1 Skala Likert

Pilihan Jawaban	Singkatan	Skor
Sangat Tidak Puas	STP	1
Tidak Puas	TP	2
Cukup Puas	CP	3
Puas	P	4
Sangat Puas	SP	5

Dimana kemudian menentukan penggunaan rata-rata tingkat kepuasan menurut metode Likert dalam Nazir (2014) dengan menggunakan rumus Likert dalam Nazir (2014) dengan menggunakan rumus

$$\text{Rata} - \text{rata Kepuasan} = \frac{\text{jumlah skor jawaban kuisioner}}{\text{Total Jumlah Kuisioner}}$$

Sedangkan untuk penentuan rata-rata kepuasan menggunakan teori Kaplan & Norton (2000):

Tabel 2 Skala Likert Rata-rata Kepuasan

Range Nilai	Keterangan
1-1.79	Sangat Tidak Puas
1.8-2.59	Tidak Puas
2.6-3.39	Cukup Puas
3.4-4.91	Puas
4.2-5	Sangat Puas

3. METODE PENELITIAN

Dalam melakukan penelitian ini penulis perlu menggunakan landasan teori yang digunakan sebagai bahan pendukung dalam menghasilkan tujuan penelitian, maka landasan teori yang akan dibahas yaitu Metode Algoritma C4.5 dan *Naïve Bayes*

2.1 Tahapan Metode Algoritma C4.5

Metode Algoritma C4.5 terdiri dari 6 tahapan yaitu sebagai berikut :

1. Masukan data hasil survey, menginput atribut yang sudah dipreprosesing.
2. *Entropy* adalah nilai informasi yang menyatakan ukuran ketidakpastian dari atribut suatu kumpulan objek data. Menghitung *entropy* total dengan rumus sebagai berikut

$$Entropy\ (S)$$

$$= \sum_{i-1}^{n} (-pi * log2 * pi) \qquad (2)$$

Keterangan :

S : Himpunan kasus

n : Jumlah data

pi : Proporsi pilihan

3. Kemudian menghitung *entropy* dan *gain* setiap atribut

Information *Gain* adalah ukuran efektifitas suatu atribut dalam mengklasifikasikan datadan diigunakan untuk menentukan urutan atribut dimana atribut yang memiliki nilai *Information Gain* terbesar yang dipilih dengan rumus sebagai berikut

$$gain\ (S, A)$$

$$= Entropy\ (s) - \sum_{i-1}^{n} \frac{|Si|}{|S|} * Entropy\ (Si) \qquad (3)$$

Keterangan :

S : Himpunan kasus

A : Atribut

n : Jumlah partisi atribut A

Si :Jumlah kasus pada partisi ke-i

4. Menentukan *gain* tertinggi hingga yang terendah
5. Kemudian membuat pohon keputusan, model membuat pohon keputusan yang menghasilkan pola atau aturan.

2.2 Naïve Bayes

Metode *Naïve Bayes* merupakan metode yang memanfaatkan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris *Thomas Bayes*. *Naïve Bayes* merupakan metode pengklasifikasian yang sangat sederhana dengan mengasumsikan klasifikasi atribut. Dengan metode *Naive Bayes* terlebih dahulu mencari Nilai Probabilitas dan likelihood maksimum dari setiap atribut untuk masing-masing kelas (Karthika *et al., 2015*) dalam kutipan jurnal (Hayuningtyas, 2019).

2.3 Tahapan Metode *Naïve Bayes*

Tahapan dari proses metode *Naïve Bayes* adalah

1. Menghitung nilai kasus disetiap kelas (label) yang ada "P(c)".
2. Menghitung nilai jumlah kasus di setiap kelas "P(x|c)".
3. Kalikan semua variabel/kelas P(x|c) x P(c).
4. Bandingkan hasil/kelas
5. Tampilkan output hasil klasifikasi
6. *Counfusion matrix*

Rumus Algoritma *Naïve Bayes* berdasarkan probabilitas yaitu :

$$P(C|X) = \frac{P(X|C).P(c)}{P(x)} \qquad (4)$$

Keterangan :

x : Data dengan class yang belum diketahui

c : Sebuah class spesifik

P(C|X): Probabilitas class berdasar kondisi (*posterior probability*)

P(c): Probabilitas class yang dicari (*prior probability*)

P(x|c): Probabilitas berdasarkan kondisi pada class

P(x) : Probabilitas suatu input dari keseluruhan data (*predictor prior probability*)

2.4 *Confusion Matrix*

Dalam tahapan ini mengevaluasi keakuratan hasil yang dicapai oleh model yang digunakan. Menggunakan *confusion matrix* yang telah disediakan dalam *framework RapidMiner*.

Pengukuran terhadap kinerja suatu sistem klasifikasi merupakan hal yang penting. Kinerja sistem klasifikasi menggambarkan seberapa baik sistem dalam mengklasifikasikan data. Berdasarkan nilai *True Negative* (TN), *False Positive* (FP), *False Negative* (FN), dan *True Positive* (TP) dapat diperoleh nilai akurasi, *presisi* dan *recall*.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \; x\; 100\% \quad (5)$$

$$Presisi = \frac{\mathrm{TP}}{FP + TP} \; x\; 100\% \quad (6)$$

$$Recall = \frac{\mathrm{TP}}{FN + TP} \; x\; 100\% \quad (7)$$

2.5 Kriteria

Menurut jurnal (Irawan & Sitanggang, 2020) Adapun kriteria kualitas pelayanan adalah kegiatan atau manfaat yang dapat diberikan oleh suatu pihak kepada pihak lain yang pada dasarnya tidak berwujud dan tidak pula berakibat pemilihan sesuatu. Faktor-faktor yang digunakan untuk mengukur kualitas pelayanan yaitu :

1. Bukti Fisik yaitu aspek yang terlihat secara fisik, misal peralatan dan personil
2. Kehandalan yaitu kemampuan untuk memiliki performa yang bisa diandalkan dan akurat
3. Daya Tanggap yaitu kemampuan untuk merespon keinginan atau kebutuhan akan bantuan dari pelanggan, serta pelayanan yang cepat
4. Jaminan yaitu kemampuan para personel untuk menimbulkan rasa percaya dan aman kepada pelanggan
5. Empati yaitu kemampuan personel untuk peduli dan

memperhatikan setiap pelanggan

2.6 Aplikasi RapidMiner

RapidMiner merupakan perangkat lunak yang dibuat oleh Dr.Markus Hofman dari Institute of Technologi Blanchardstown dan Ralf Klinkenberg dari rapid-i.com dengan tampilan GUI *(Graphical User Interface)* sehingga memudahkan pengguna dalam menggunakan perangkat lunak ini. Perangkat lunak ini bersifat open source dan dibuat dengan menggunakan program Java di bawah lisensi GNU *Public Licence* dan *RapidMiner* dapat dijalankan di sistem operasi manapun. Dengan menggunakan *RapidMiner*, tidak dibutuhkan kemampuan koding khusus, karena semua fasilitas sudah disediakan.

3.1 Pengumpulan Data

Pada penelitian ini data yang digunakan didapatkan dari RSUD Hasri Ainun Habibie terutama pada paiesn rawat inap di RSUD Hasri Ainun Habibie. Penelitian ini bertujuan untuk mengetahui kepuasan pasien terhadap pelayanan yang diberikan di RSUD Ainun Habibie dengan menggunakan metode Algoritma C4.5 dan *Naïve Bayes*.

Dalam melakukan pengumpulan data agar kegiatan penelitian tersebut menjadi sistematis dan di permudah dalam penelitian ini diantaranya adalah

a. Wawancara

Wawancara adalah dengan melakukan tanya jawab terhadap staf perencanaan yang mengelolah data kepuasan masyarakat. Proses wawancara dilakukan dengan mengajukan pertanyaan, kemudian mencatat jawaban-jawaban dan meminta penjelasan mengenai hal-hal terkait dengan penelitian.

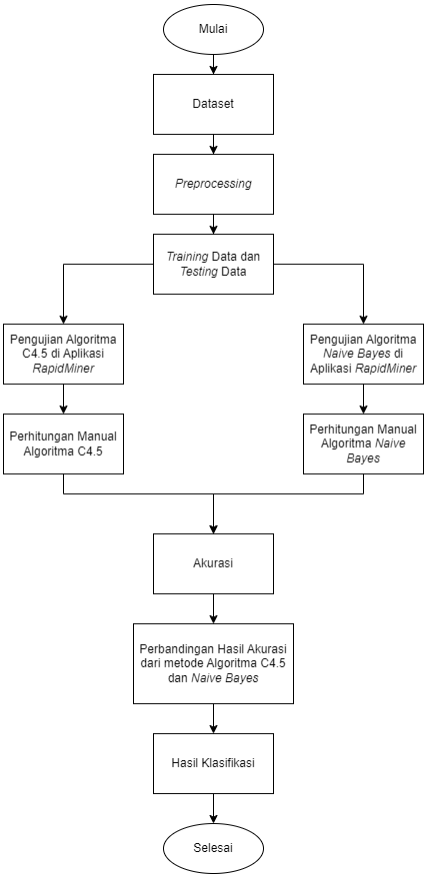
b. Kuisioner

Kuisioner berisikan pertanyaan dan jawaban terhadap tingkat kepuasan pasien. Narasumber yaitu keluarga dari pasien memilih jawaban sesuai dengan tanggapan dan pendapatnya dengan cara memberikan *check list* / centang pada kolom yang tersedia.

3.2 Tahapan Penelitian

Tahapan yang dilakukan dalam pelaksanaan penelitian ini adalah sebagai berikut:

Gambar 1. Tahapan Penelitian



4. HASIL DAN PEMBAHASAN

A. Hasil Pengumpulan Data

Pada hasil pengumpulan data yaitu dalam bentuk kuisioner dan telah mengumpulkan sebanyak 376 data kuisioner. Tahap selanjutnya yaitu tahap *pre-processing* data, yaitu dengan cara menghapus data yang tidak relevan. Data yang dihapus adalah nama dan jenis kelamin. Kemudian dilakukan tahapan data *training* yang digunakan untuk melatih model pembelajaran *mechine learning* dimana untuk memahami pola, tren dan hubungan dalam data. Data *testing* ini merupakan data set yang digunakan untuk melihat akurasi atau melihat performanya. Kemudian data dibagi menjadi 80% untuk data *training* dan 20% untuk data *testing*. Setelah itu dilakukan perhitungan menggunakan manual di excel menggunakan metode Algoritma C4.5 dan *Naïve Bayes* dan dilakukan tahapan pengujian ke aplikasi *rapidminer*. Tahap selanjutnya yaitu melihat nilai akurasi tertinggi pada 2 algoritma tersebut. Untuk mengevaluasi penelitian klasifikasi untuk hasil pengujiannya menggunakan *cross validation* untuk mendapatkan hasil *accuracy*. Dan terakhir tahapan klasifikasi yaitu menentukan hasil klasifikasi yang bertujuan untuk mengetahui algoritma mana yang menghasilkan akurasi tertinggi.

B. Pre-processing

Pre-processing merupakan analisis data dimana data mentah diubah, dibersihkan, dan dipersiapkan agar cocok untuk analisis lebih lanjut. Maka sebelum melakukan tahap pengolahan data, harus dilakukan tahap *pre-processing*. Tahap *pre-processing* ini sangat penting karena kualitas data tidak terolah dapat menghasilakn hasil analisis yang tidak akurat atau model yang tidak efektif.

A.

B. Hasil Klasifikasi Algoritma C4.5

accuracy: 70.00% +/- 13.01% (micro average: 70.00%)

	true Sangat Puas	true Puas	true Tidak Puas	true Sangat Tidak Puas	true Cukup Puas	class precision
pred. Sangat Puas	316	80	0	0	0	79.80%
pred. Puas	122	146	1	3	2	53.28%
pred. Tidak Puas	0	0	19	2	7	67.86%
pred. Sangat Tidak Puas	0	0	0	33	0	100.00%
pred. Cukup Puas	2	4	0	2	11	57.89%
class recall	71.82%	63.48%	95.00%	82.50%	55.00%	

Gambar 2 Hasil Klasifikasi Algoritma C4.5

Untuk hasil klasifikasi Algoritma C4.5 menggunakan data testing 20% atau 75 data. Dan menggunakan operator *Cross Validation* dengan *Number Fold* 10 kali pengujian, jadi total data testing menjadi 750 data testing.

Tabel 3 Metriks Evaluasi Algoritma C4.5

Akurasi	***Class Precision***	***Class Recall***
70.00% ± 13.01%	Sangat Puas: 79.80% Puas: 53.28% Tidak Puas: 67.86% Sangat Tidak Puas: 100.00% Cukup Puas: 57.89%	Sangat Puas: 71.82% Puas: 63.48% Tidak Puas: 95.00% Sangat Tidak Puas: 82.50% Cukup Puas: 55.00%

Algoritma C4.5 menunjukkan presisi yang bervariasi untuk setiap kelas, dengan presisi yang tinggi untuk "Sangat Tidak Puas" dan "Sangat Puas," namun presisi yang lebih rendah untuk kelas "Puas" dan "Cukup Puas."

Recall untuk kelas "Tidak Puas" dan "Sangat Tidak Puas" tinggi, menunjukkan bahwa hampir semua instansi dalam kelas-kelas ini berhasil diidentifikasi dengan benar.

Recall untuk kelas "Sangat Puas," "Puas," dan "Cukup Puas" menunjukkan bahwa ada banyak instansi yang terlewatkan, menunjukkan kebutuhan untuk perbaikan dalam identifikasi instansi ini.

Secara keseluruhan, algoritma C4.5 bekerja cukup baik namun tidak seefektif algoritma Naïve Bayes, dengan akurasi 70.00% dan variasi kinerja yang besar di antara kelas-kelas.

C. Hasil Klasifikasi Algoritma *Naïve Bayes*

accuracy: 86.40% +/- 0.84% (micro average: 86.40%)

	true Sangat Puas	true Puas	true Tidak Puas	true Sangat Tidak Puas	true Cukup Puas	class precision
pred. Sangat Puas	349	0	0	0	0	100.00%
pred. Puas	92	230	0	0	0	71.43%
pred. Tidak Puas	0	0	20	0	10	66.67%
pred. Sangat Tidak Puas	0	0	0	40	0	100.00%
pred. Cukup Puas	0	0	0	0	10	100.00%
class recall	79.09%	100.00%	100.00%	100.00%	50.00%	

Gambar 3 Hasil Klasifikasi Algoritma *Naïve Bayes*

Untuk hasil klasifikasi Algoritma *Naïve Bayes* menggunakan data testing 20% atau 75 data. Dan menggunakan operator *Cross Validation* dengan *Number Fold* 10 kali pengujian, jadi total data testing menjadi 750 data testing.

Tabel 4 Metriks Evaluasi Algoritma *Naive Bayes*

Akurasi	***Class Precision***	***Class Recall***
86.40% ± 0.84%	Sangat Puas: 100.00% Puas: 71.43% Tidak Puas: 66.67% Sangat Tidak Puas: 100.00% Cukup Puas: 100.00%	Sangat Puas: 79.09% Puas: 100.00% Tidak Puas: 100.00% Sangat Tidak Puas: 100.00% Cukup Puas: 50.00%

Algoritma Naïve Bayes menunjukkan presisi yang kuat untuk sebagian besar kelas, terutama "Sangat Puas," "Sangat Tidak Puas," dan "Cukup Puas," yang menunjukkan sedikit atau tidak ada false positives.

Recall sempurna (100%) untuk "Puas," "Tidak Puas," dan "Sangat Tidak Puas," yang berarti semua instansi dalam kelas-kelas ini berhasil diidentifikasi dengan benar.

Recall untuk "Sangat Puas" dan "Cukup Puas" menunjukkan ruang untuk perbaikan, karena beberapa instansi terlewatkan.

Secara keseluruhan, algoritma Naïve Bayes bekerja dengan baik, dengan akurasi 86.40%, namun penyetelan lebih lanjut mungkin diperlukan untuk meningkatkan recall untuk beberapa kelas.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil analisis data menggunakan algoritma C4.5 dan *Naïve Bayes* untuk mengukur tingkat kepuasan pasien di RSUD Hasri Ainun Habibie, dapat disimpulkan bahwa algoritma *Naïve Bayes* menunjukkan kinerja yang lebih baik dibandingkan dengan algoritma C4.5. Hal ini dapat dilihat dari akurasi keseluruhan, presisi, dan recall yang lebih tinggi pada *Naïve Bayes* dibandingkan dengan C4.5.

- Akurasi Keseluruhan

  Algoritma *Naïve Bayes* mencapai akurasi keseluruhan sebesar 86.40% dengan margin kesalahan ± 0.84%, yang menunjukkan bahwa algoritma ini stabil dan andal dalam mengklasifikasikan tingkat kepuasan pasien. Sebaliknya, algoritma C4.5 hanya mencapai akurasi 70.00% dengan margin kesalahan ± 13.01%, yang menunjukkan ketidakstabilan dan kinerja yang lebih rendah.

- Presisi dan Recall Berdasarkan Kelas

  Naïve Bayes menunjukkan kinerja yang sangat baik dalam hal presisi dan recall untuk sebagian besar kelas kepuasan. Misalnya, pada kelas "Sangat Puas," *Naïve Bayes* memiliki presisi sempurna sebesar 100.00% dan recall 79.09%, sedangkan C4.5 memiliki presisi 79.80% dan recall 71.82%. Untuk kelas "Puas," *Naïve Bayes* juga unggul dengan presisi 71.43% dan recall 100.00%, dibandingkan dengan presisi 53.28% dan recall 63.48% pada C4.5. Pada kelas "Tidak Puas," kedua algoritma menunjukkan kinerja yang baik, namun *Naïve Bayes* kembali unggul dengan recall sempurna sebesar 100.00%, memastikan bahwa semua instansi "Tidak Puas" teridentifikasi dengan benar. Untuk kelas "Sangat Tidak Puas," *Naïve Bayes* menunjukkan presisi dan recall sempurna 100.00%, sedangkan C4.5 memiliki recall 82.50%.

- Kelemahan dan Rekomendasi

  Kedua algoritma menunjukkan kelemahan dalam mengklasifikasikan kelas "Cukup Puas." *Naïve Bayes* menunjukkan presisi sempurna 100.00% namun recall rendah 50.00%, sementara C4.5 memiliki presisi 57.89% dan recall 55.00%. Oleh karena itu, perbaikan diperlukan untuk meningkatkan kinerja pada kelas ini, baik melalui penyesuaian parameter atau peningkatan jumlah data pelatihan.

- Implikasi untuk Manajemen Rumah Sakit

  Hasil ini menunjukkan bahwa algoritma *Naïve Bayes* lebih disarankan untuk digunakan dalam mengklasifikasikan kepuasan pasien di RSUD Hasri Ainun Habibie. Implementasi algoritma ini dapat membantu manajemen rumah sakit dalam mengambil keputusan yang lebih akurat dan efektif berdasarkan data kepuasan pasien yang dihasilkan. Namun, perhatian khusus perlu diberikan pada perbaikan klasifikasi kelas "Cukup Puas" untuk memastikan bahwa semua tingkat kepuasan pasien teridentifikasi dengan benar dan dapat ditindaklanjuti dengan tepat.

Dari hasil penelitian ini, ada beberapa saran yang terkait dengan penelitian mengukur tingkat kepuasan pasien, yakni sebagai berikut:

- Pengembangan Metode Analisis:

  Untuk mencapai model klasifikasi yang lebih baik dengan cakupan metode analisis yang luas dan dapat diterapkan pada berbagai bidang ilmu dan studi kasus, peneliti selanjutnya dapat mengembangkan

dan menguji metode analisis lainnya. Kombinasi atau hibridisasi antara algoritma yang berbeda juga bisa dipertimbangkan untuk meningkatkan kinerja klasifikasi.

- Penggunaan Aplikasi atau Perangkat Lunak yang Berbeda:
  Penulis menyarankan agar penelitian selanjutnya dilakukan dengan menggunakan aplikasi atau perangkat lunak yang berbeda dari penelitian sebelumnya sebagai pembanding untuk mendapatkan hasil yang lebih baik. Penggunaan berbagai alat analisis dapat memberikan perspektif yang lebih luas dan memungkinkan identifikasi metode yang paling efektif untuk kasus spesifik ini.
- Perbaikan Klasifikasi Kelas "Cukup Puas":
  Penelitian lebih lanjut perlu memperhatikan perbaikan dalam mengklasifikasikan kelas "Cukup Puas" dengan menyesuaikan parameter algoritma atau meningkatkan jumlah data pelatihan. Hal ini penting untuk memastikan bahwa semua tingkat kepuasan pasien dapat teridentifikasi dengan benar dan dapat ditindaklanjuti dengan tepat.
- Peningkatan Kualitas Data:
  Disarankan untuk mengumpulkan lebih banyak data yang lebih beragam untuk meningkatkan keakuratan model. Kualitas dan kuantitas data yang lebih baik dapat membantu dalam menghasilkan model yang lebih akurat dan andal.
- Evaluasi dan Validasi Model:
  Peneliti selanjutnya sebaiknya melakukan evaluasi dan validasi model secara lebih mendalam dengan menggunakan teknik cross-validation atau metode validasi lainnya untuk memastikan bahwa model yang dikembangkan memiliki kinerja yang baik dan dapat diandalkan di berbagai kondisi.

DAFTAR PUSTAKA

Hayuningtyas, R. Y. (2019). Penerapan Algoritma Naïve Bayes untuk Rekomendasi Pakaian Wanita. *Jurnal Informatika*, *6*(1), 18–22. https://doi.org/10.31294/ji.v6i1.4685

Irawan, B., & Sitanggang, E. D. (2020). *Pelayanan Rumah Sakit Berdasarkan Metode Service Quality (SERVQUAL) 2 Amik Medan Business Polytechnic*. *3*(1).

Nasrullah, A. H. (2021). Implementasi Algoritma Decision Tree Untuk Klasifikasi Produk Laris. *Jurnal Ilmiah Ilmu Komputer*, *7*(2), 45–51. https://doi.org/10.35329/jiik.v7i2.203

Sesrianty, V., Machmud, R., & Yeni, F. (2019). Analisa Kepuasan Pasien Terhadap Mutu Pelayanan Keperawatan. *Jurnal Kesehatan Perintis (Perintis's Health Journal)*, *6*(2), 116–126. https://doi.org/10.33653/jkp.v6i2.317

Sofi Adinda Amri Nasution, & Yayuk Yuliana. (2022). Pengaruh Citra Merek, Kualitas Layanan, Dan Harga Terhadap Kepuasan Pengguna Jasa Pengiriman J&T Express Di Drop

Point Krakatau Medan. *Jurnal Ilmu Manajemen Dan Kewirausahaan (Jimk)*, *2*(2), 185–194. https://doi.org/10.32696/jimk.v2i1.1458

Solikhah, F., Febianah, M., Kamil, A. L., Arifin, W. A., & Shelly Janu Setyaning Tyas. (2021). Analisis Perbandingan Algoritma Naive Bayes Dan C.45 Dalam Klasifikasi Data Mining Untuk Memprediksi Kelulusan. *Tematik*, *8*(1), 96–103. https://doi.org/10.38204/tematik.v8i1.576

Tarigan, P. M. S., Hardinata, J. T., Qurniawan, H., Safii, M., & Winanjaya, R. (2022). Implementasi Data Mining Menggunakan Algoritma Apriori Dalam Menentukan Persediaan Barang. *Jurnal Janitra Informatika dan Sistem Informasi*, *2*(1), 9–19. https://doi.org/10.25008/janitra.v2i1.142