



Analisis Kemampuan Penyelenggaraan Puskesmas Tahun 2024 Menggunakan Algoritma Naive Bayes Berbasis Rapid Miner

Analysis of Public Health Center Service Capability in 2024 Using the Naive Bayes Algorithm Based on Rapid Miner

Hafidz Munazzal Falahan^{1*}, Hasbi Firmansyah²

Universitas Pancasila Tegal

Email: hafidzfalahan@gmail.com^{1*}, hasbifirmansyah@upstegal.ac.id²

Article Info

Article history :

Received : 02-12-2025

Revised : 03-12-2025

Accepted : 05-12-2025

Published : 07-12-2025

Abstract

The evaluation of Public Health Center (Puskesmas) service capability plays a critical role in ensuring the quality of primary healthcare delivery in Indonesia. However, the assessment process often faces challenges related to data variability, classification accuracy, and the need for efficient analytical tools. This study aims to analyze the capability level of Puskesmas in 2024 using the Naive Bayes algorithm implemented through the RapidMiner platform. The dataset consists of key performance indicators that represent administrative, medical, managerial, and service-related components of Puskesmas operations. The Naive Bayes classifier was selected due to its simplicity, high interpretability, and effectiveness in handling categorical health-service data. The research includes data preprocessing, attribute weighting, model training, and performance evaluation. The classification results show that Naive Bayes is capable of predicting capability levels with satisfactory accuracy, supported by confusion matrix and performance metrics generated within RapidMiner. These findings indicate that data mining techniques—particularly Naive Bayes—provide a reliable approach to supporting decision-making processes in the evaluation of Puskesmas service capability. The study contributes to improving data-driven assessments and serves as a reference for future research in healthcare quality analytics.

Keywords: Public Health Center, Naive Bayes, Rapid Miner

Abstrak

Evaluasi kemampuan penyelenggaraan Puskesmas merupakan aspek penting dalam menjamin mutu pelayanan kesehatan tingkat pertama di Indonesia. Namun, proses penilaian sering menghadapi tantangan terkait keragaman data, ketepatan klasifikasi, serta kebutuhan akan alat analisis yang efisien. Penelitian ini bertujuan untuk menganalisis tingkat kemampuan penyelenggaraan Puskesmas tahun 2024 dengan menggunakan algoritma Naive Bayes yang diimplementasikan melalui platform RapidMiner. Dataset yang digunakan mencakup indikator kinerja utama yang mewakili komponen administratif, medis, manajerial, serta kualitas pelayanan Puskesmas. Algoritma Naive Bayes dipilih karena kesederhanaannya, tingkat interpretabilitas yang tinggi, serta efektivitasnya dalam mengolah data kategori pada layanan kesehatan. Tahapan penelitian meliputi praproses data, pembobotan atribut, pelatihan model, serta evaluasi kinerja. Hasil klasifikasi menunjukkan bahwa Naive Bayes mampu memprediksi tingkat kemampuan Puskesmas dengan akurasi yang memadai, didukung oleh confusion matrix dan metrik performa yang dihasilkan oleh RapidMiner. Temuan ini menunjukkan bahwa teknik data mining—khususnya Naive Bayes—dapat menjadi pendekatan yang andal dalam mendukung proses pengambilan keputusan terkait evaluasi kemampuan penyelenggaraan Puskesmas. Penelitian ini memberikan kontribusi pada pengembangan penilaian berbasis data dan menjadi rujukan untuk penelitian selanjutnya dalam analisis kualitas pelayanan

Kata Kunci: Puskesmas, NaiveBayes, RapidMiner



PENDAHULUAN

Puskesmas merupakan salah satu ujung tombak pelayanan kesehatan primer di Indonesia yang berperan penting dalam upaya promotif, preventif, kuratif, dan rehabilitatif. Sebagai institusi kesehatan dasar, Puskesmas dituntut untuk mampu menyelenggarakan pelayanan yang berkualitas, merata, dan mudah diakses oleh masyarakat. Pemerintah melalui regulasi nasional telah menetapkan berbagai indikator penilaian untuk mengukur kemampuan penyelenggaraan Puskesmas, yang mencakup aspek administrasi, manajemen, ketenagaan, upaya kesehatan, serta mutu pelayanan. Penilaian terhadap kemampuan ini sangat diperlukan untuk mengetahui tingkat kesiapan serta kualitas layanan kesehatan yang diberikan kepada masyarakat.

Namun demikian, proses evaluasi kemampuan penyelenggaraan Puskesmas sering kali menghadapi kendala, terutama terkait dengan keragaman data, jumlah indikator yang besar, serta kebutuhan akan metode analisis yang cepat dan akurat. Pendekatan konvensional dalam pengolahan data kesehatan sering membutuhkan waktu yang lama dan rentan terhadap bias manusia. Di era digital, pemanfaatan teknik data mining menjadi alternatif yang efektif untuk membantu proses klasifikasi dan pengambilan keputusan berbasis data. Salah satu metode data mining yang banyak digunakan adalah algoritma Naive Bayes, yang dikenal memiliki kemampuan klasifikasi yang baik, sederhana, serta efektif dalam menangani data kategori.

RapidMiner merupakan perangkat lunak data mining yang menyediakan berbagai fitur untuk pengolahan data, pelatihan model, dan evaluasi performa algoritma secara praktis. Penggunaan RapidMiner dalam penelitian memungkinkan proses analisis menjadi lebih terstruktur, terstandarisasi, dan mudah direplikasi oleh peneliti lainnya. Dengan demikian, penerapan algoritma Naive Bayes berbasis RapidMiner dalam menganalisis kemampuan penyelenggaraan Puskesmas dapat memberikan solusi yang efisien dalam memetakan kategori kemampuan Puskesmas berdasarkan data kinerja tahun 2024.

Penelitian ini bertujuan untuk menganalisis tingkat kemampuan penyelenggaraan Puskesmas tahun 2024 menggunakan algoritma Naive Bayes pada platform RapidMiner. Penelitian ini juga memberikan gambaran mengenai performa model klasifikasi berdasarkan metrik seperti akurasi, precision, recall, dan analisis confusion matrix. Hasil penelitian diharapkan dapat memberikan kontribusi dalam mendukung proses evaluasi berbasis data serta menjadi acuan bagi pihak terkait, seperti Dinas Kesehatan dan pengelola Puskesmas, dalam meningkatkan kualitas pelayanan kesehatan primer.

METODE PENELITIAN

Metode penelitian ini dirancang untuk menganalisis tingkat kemampuan penyelenggaraan Puskesmas tahun 2024 menggunakan algoritma Naive Bayes dengan bantuan perangkat lunak RapidMiner. Proses penelitian mencakup beberapa tahapan utama, yaitu pengumpulan data, praproses data, perancangan model, pengujian model, serta evaluasi hasil. Adapun tahapan metode penelitian dijelaskan sebagai berikut:

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari:

1. Laporan capaian kinerja Puskesmas tahun 2024
2. Data indikator penilaian kemampuan penyelenggaraan Puskesmas



3. Dokumen evaluasi internal atau eksternal yang relevan

Pengujian dilakukan dengan teknik validasi silang (cross-validation) untuk memperoleh hasil yang lebih akurat dan mengurangi overfitting. Pada tahap ini, model diuji menggunakan data yang berbeda dari data pelatihan.

Output yang dihasilkan meliputi:

1. Confusion Matrix
2. Accuracy
3. Precision
4. Recall
5. F-Measure

Secara keseluruhan, alur penelitian dapat digambarkan sebagai berikut:

1. Identifikasi masalah
2. Pengumpulan data
3. Praproses data
4. Pemodelan Naive Bayes pada RapidMiner
5. Pengujian dan evaluasi
6. Interpretasi hasil
7. Penyusunan laporan penelitian

Penelitian ini termasuk dalam kategori penelitian kuantitatif dengan pendekatan data mining khususnya metode klasifikasi. Algoritma Naive Bayes digunakan untuk memprediksi kategori kemampuan Puskesmas berdasarkan variabel atau indikator kinerja yang relevan. Penelitian ini menggunakan desain eksperimental di mana model dilatih dan diuji menggunakan dataset tahun 2024.

HASIL DAN PEMBAHASAN

Deskripsi Singkat Data dan Tujuan Klasifikasi

Data kemampuan penyelenggaraan Puskesmas tahun 2024 yang digunakan dalam penelitian ini berjumlah 881 instance data. Setiap baris data merepresentasikan satu unit observasi (misalnya satu Puskesmas atau satu hasil penilaian) yang berisi sejumlah indikator:

1. sumber daya (tenaga kesehatan, fasilitas),
2. jenis layanan yang diselenggarakan,
3. capaian program,
4. aspek administratif dan manajerial, dan lain-lain.



Label kelas (class) dibagi menjadi dua kategori utama:

1. Rawat Inap – Puskesmas yang mampu menyelenggarakan layanan rawat inap.
2. Non Rawat Inap – Puskesmas yang belum atau tidak menyelenggarakan layanan rawat inap.

Pada output RapidMiner, muncul satu kategori tambahan “Non Rawat Inap” (dengan spasi ganda). Ini bukan kelas konsep baru, tetapi kesalahan penulisan label di data Excel yang terbaca sebagai kelas terpisah oleh RapidMiner. Di jurnal, hal ini penting disebut sebagai catatan kualitas data.

Tujuan pemodelan adalah memprediksi apakah suatu Puskesmas termasuk Rawat Inap atau Non Rawat Inap berdasarkan indikator-indikator kemampuan penyelenggaraannya.

Hasil Confusion Matrix

Berdasarkan proses pemodelan Naive Bayes di RapidMiner, diperoleh **confusion matrix** sebagai berikut:

	true Rawat Inap	true Non Rawat Inap	true Non Rawat Inap	class precision
pred. Rawat Inap	367	6	0	98.39%
pred. Non Rawat Inap	4	503	0	99.21%
pred. Non Rawat inap	0	0	1	100.00%
class recall	98.92%	98.82%	100.00%	

Gambar 1.Tampilan Akurasi Rapid Miner

“Non Rawat Inap” adalah label yang sama dengan “Non Rawat Inap” namun tertulis ganda spasi sehingga terbaca sebagai kelas berbeda.

Artinya:

1. Dari seluruh data Rawat Inap, sebanyak 367 instance diklasifikasikan dengan benar sebagai Rawat Inap, dan hanya 6 instance yang salah tebak menjadi Non Rawat Inap.
2. Dari data Non Rawat Inap, sebanyak 503 instance terklasifikasi dengan benar, dan 4 instance salah diprediksi sebagai Rawat Inap.
3. Terdapat 1 instance dengan label “Non Rawat Inap”, dan model memprediksinya 100% benar pada kelas yang sama.

Jika dijumlahkan, terdapat:

- a. $(367 + 503 + 1) = 871$ instance yang tepat diklasifikasikan.
- b. $(6 + 4 + 0) = 10$ instance yang keliru.



classification_error: 1.14%

	true Rawat Inap	true Non Rawat Inap	true Non Rawat Inap	class precision
pred. Rawat Inap	367	6	0	98.39%
pred. Non Rawat Inap	4	503	0	99.21%
pred. Non Rawat Inap	0	0	1	100.00%
class recall	98.92%	98.82%	100.00%	

Akurasi dan Metrik Kinerja Model

RapidMiner menghasilkan ringkasan kinerja sebagai berikut:

1. Accuracy = 98,86%
2. Classification error = 1,14%

Untuk masing-masing kelas:

1. Kelas Rawat Inap
 - a. *Class precision* = 98,39%
 - b. *Class recall* = 98,92%
2. Kelas Non Rawat Inap
 - a. *Class precision* = 99,21%
 - b. *Class recall* = 98,82%
3. Kelas Non Rawat Inap (satu data)
 - a. *Class precision* = 100%
 - b. *Class recall* = 100%

Interpretasi singkat:

1. Accuracy 98,86% berarti sekitar 98–99 dari 100 data dapat diklasifikasikan dengan benar oleh model Naive Bayes.
2. Precision yang tinggi menunjukkan bahwa ketika model memutuskan “Rawat Inap”, hampir semua memang benar Rawat Inap; demikian pula untuk Non Rawat Inap.
3. Recall yang tinggi berarti hampir semua Puskesmas Rawat Inap dan Non Rawat Inap berhasil “tertangkap” oleh model, hanya sebagian kecil yang terlewat atau salah klasifikasi.

Analisis Detail Per Kelas

1. Kelas Rawat Inap

Dari confusion matrix:

- a. Benar Rawat Inap → Rawat Inap: 367 data
- b. Salah Rawat Inap → Non Rawat Inap: 6 data



Dengan recall sebesar **98,92%**, hampir seluruh Puskesmas yang secara nyata mampu menyelenggarakan rawat inap **berhasil dikenali** oleh model.

Kesalahan 6 data dapat disebabkan oleh:

- indikator fasilitas atau tenaga kesehatan yang hampir serupa dengan Puskesmas Non Rawat Inap (misalnya tempat tidur sedikit, fasilitas terbatas),
- atau adanya ketidakkonsistenan pengisian data.

Secara praktis, kemampuan model mengenali Puskesmas Rawat Inap ini penting untuk:

- pemetaan distribusi layanan rawat inap,
- perencanaan rujukan,
- penentuan prioritas pengembangan fasilitas.

2. Kelas Non Rawat Inap

Untuk kelas Non Rawat Inap:

- Benar Non Rawat Inap → Non Rawat Inap: 503 data
- Salah Non Rawat Inap → Rawat Inap: 4 data

Recall sebesar **98,82%** menunjukkan bahwa hampir semua Puskesmas Non Rawat Inap terdeteksi dengan tepat.

Kesalahan 4 data yang diprediksi sebagai Rawat Inap kemungkinan terjadi pada Puskesmas yang memiliki karakteristik “hampir Rawat Inap”, misalnya:

- sarana dan prasarana sudah cukup lengkap,
- jumlah tenaga kesehatan tinggi,
- tetapi secara administrasi belum dikukuhkan sebagai Puskesmas Rawat Inap.

Hal ini menarik karena dari sisi kebijakan, 4 Puskesmas tersebut bisa menjadi kandidat pengembangan ke arah layanan rawat inap di masa depan.

3. Kelas “Non Rawat Inap” (Kesalahan Label)

Adanya satu label “Non Rawat Inap” dengan spasi ganda merupakan contoh kesalahan penulisan (typo) atau inkonsistensi pengkodean di dataset. Walaupun model dapat mengklasifikasikannya dengan benar, di sisi metodologi hal ini perlu dicatat sebagai:

- potensi sumber bias,
- indikator bahwa data cleaning perlu lebih ketat (penyeragaman nama kelas).

Di bagian pembahasan jurnal, kamu bisa menekankan bahwa konsistensi label sangat penting agar algoritma tidak membaca satu kategori menjadi beberapa kelas berbeda.

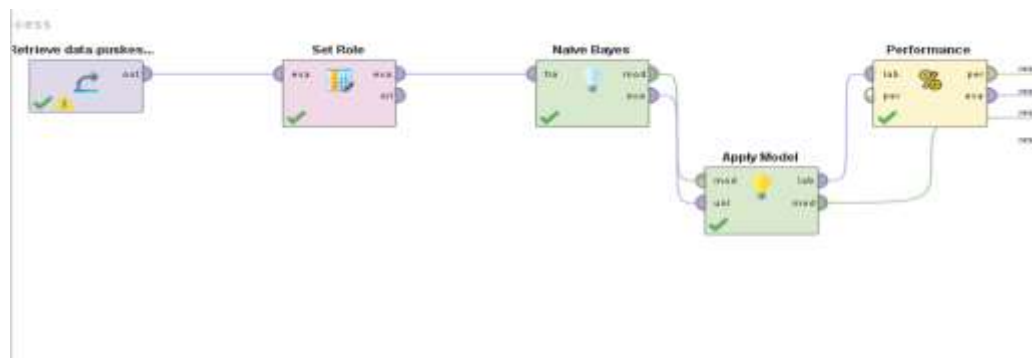


Attribute	Parameter	Rawat Inap	Non Rawat Inap	Non Rawat Inap
tahun_data	mean	2024	2024	2024
tahun_data	standard deviation	0.001	0.001	0.001
kabupaten_kota	value=Kab Cilacap	0.062	0.029	0.001
kabupaten_kota	value=Kab Banyumas	0.043	0.045	0.962
kabupaten_kota	value=Kab Purwalingga	0.030	0.022	0.001
kabupaten_kota	value=Kab Banjarnegara	0.040	0.039	0.001
kabupaten_kota	value=Kab Kebumen	0.030	0.047	0.001
kabupaten_kota	value=Kab Purwokerto	0.032	0.029	0.001
kabupaten_kota	value=Kab Wonosobo	0.022	0.031	0.001
kabupaten_kota	value=Kab Magelang	0.022	0.041	0.001
kabupaten_kota	value=Kab Boyali	0.032	0.026	0.001
kabupaten_kota	value=Kab Klaten	0.040	0.037	0.001
kabupaten_kota	value=Kab Sukoharjo	0.027	0.004	0.001
kabupaten_kota	value=Kab Wonorejo	0.013	0.057	0.001

Gambar 2.Data Puskesmas tahun2024

Name	Type	Missing	Statistics
✓ kemampuan_penyelempger...	Polynomial	0	Count: Non Rawat Inap (1), Non Rawat Inap (509), Rawat Inap (509), Rawat Inap (507)
✓ prediction(kemampuan_pe...	Polynomial	0	Count: Non Rawat Inap (1), Non Rawat Inap (507), Rawat Inap (507), Rawat Inap (507)
✓ confidence(Rawat Inap)	Real	0	Min: 0.000, Max: 1.000, Average: 0.422
✓ confidence(Non Rawat Inap)	Real	0	Min: 0.000, Max: 1.000, Average: 0.576
✓ confidence(Non Rawat Inap)	Real	0	Min: 0.000, Max: 1.000, Average: 0.002
✓ tahun_data	Integer	0	Min: 2024, Max: 2024, Range: 2024
✓ kabupaten_kota	Polynomial	0	Count: Kota Magelang (5), Kab.Banyumas (40), Kab.Brebes (40), Kab.Brebes (40)

Gambar 3.Indikator Kemampuan Puskesmas

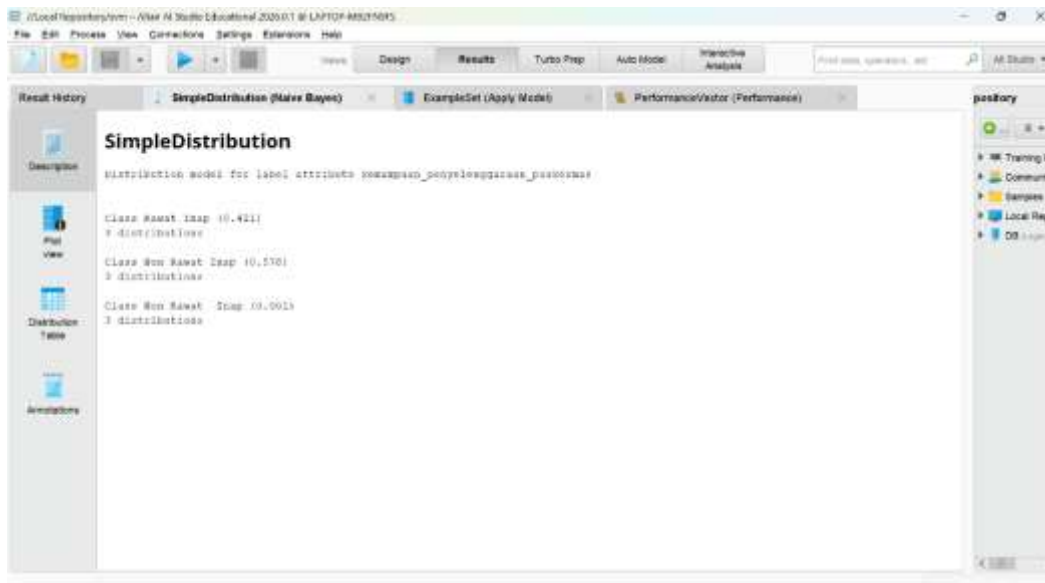


Gambar 4.Operator NaiveBayes

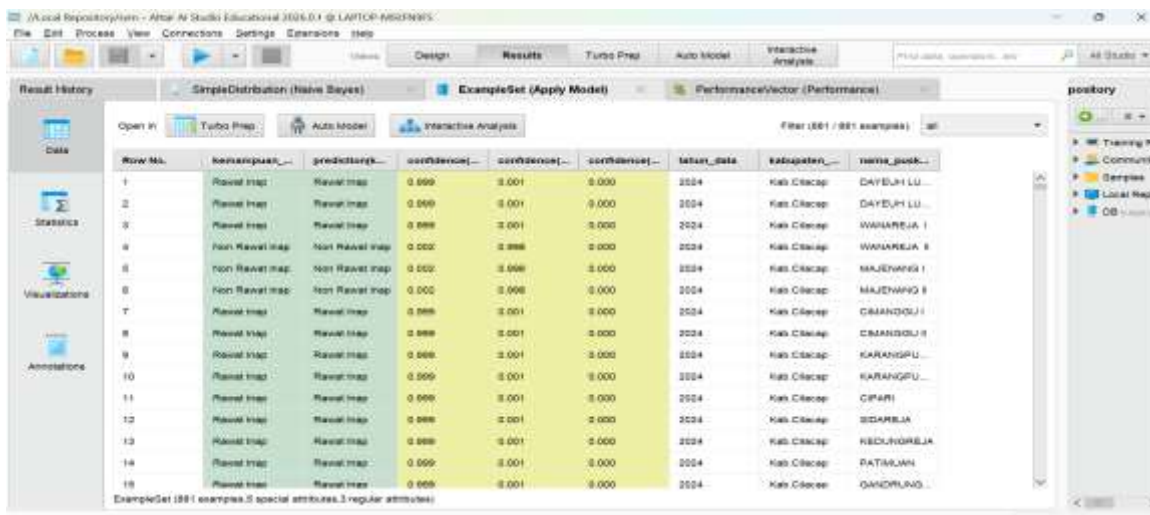
Implikasi Hasil terhadap Kemampuan Penyelenggaraan Puskesmas

Dari kinerja model yang sangat tinggi (akurasi 98,86%), dapat diinterpretasikan bahwa:

1. **Indikator kemampuan penyelenggaraan Puskesmas sangat informatif**
Pola perbedaan antara Puskesmas Rawat Inap dan Non Rawat Inap ternyata terbaca jelas oleh algoritma Naive Bayes. Artinya, indikator seperti:
 - a. ketersediaan tempat tidur,
 - b. kelengkapan sarana rawat inap,
 - c. jumlah dan kualifikasi tenaga kesehatan,
 - d. kesiapan layanan 24 jam,
 - e. capaian program,
2. **Naive Bayes layak digunakan sebagai alat bantu pengambilan keputusan**
Dengan akurasi hampir 99%, model ini dapat dimanfaatkan oleh:
 - a. Dinas Kesehatan untuk memetakan kemampuan Puskesmas secara cepat,
 - b. pengambil kebijakan untuk menilai kebutuhan peningkatan sarana,
 - c. perencanaan rujukan dan penyebaran layanan.
3. **Identifikasi Puskesmas borderline**
Data yang salah klasifikasi (6 + 4 instance) justru dapat menjadi **objek kajian menarik**, karena:
 - a. Puskesmas Rawat Inap yang “mirip” Non Rawat Inap dapat didalami apakah fasilitasnya kurang,
 - b. Puskesmas Non Rawat Inap yang “terbaca” seperti Rawat Inap bisa menjadi calon lokasi pengembangan layanan rawat inap.



Gambar 5.SimpleDistribution



Gambar 6.ExampleSet

Ringkasan Pembahasan

Hasil analisis menggunakan algoritma Naive Bayes berbasis RapidMiner menunjukkan bahwa model mampu mengklasifikasikan kemampuan penyelenggaraan Puskesmas tahun 2024 dengan tingkat akurasi yang sangat tinggi, yaitu 98,86%. Tingginya akurasi ini menunjukkan bahwa indikator-indikator yang digunakan dalam penilaian Puskesmas memiliki pola yang jelas dan konsisten sehingga dapat dikenali dengan baik oleh model probabilistik. Confusion matrix menggambarkan bahwa dari seluruh data Rawat Inap, sebanyak 367 instance berhasil diprediksi dengan benar, sedangkan 6 instance mengalami kesalahan klasifikasi. Demikian pula pada kelas Non Rawat Inap, sebanyak 503 instance diklasifikasikan dengan benar dan hanya 4 instance yang salah prediksi. Terdapat satu kategori tambahan “Non Rawat Inap” yang muncul akibat inkonsistensi penulisan label pada data asli, namun model tetap memprediksinya secara akurat. Secara keseluruhan, hanya 10 data yang keliru, menunjukkan stabilitas model yang sangat baik.

Tingginya nilai precision dan recall pada kedua kelas utama mengindikasikan bahwa variabel-variabel dalam dataset memiliki kemampuan diskriminatif yang kuat. Hal ini



mencerminkan bahwa Puskesmas Rawat Inap dan Non Rawat Inap memiliki perbedaan signifikan pada aspek-aspek seperti ketersediaan tempat tidur, jumlah tenaga kesehatan, kelengkapan sarana, kesiapan pelayanan 24 jam, serta capaian program. Perbedaan-perbedaan ini menghasilkan struktur data yang mudah separable sehingga algoritma Naive Bayes dapat mengidentifikasi probabilitas posterior dengan tingkat kesalahan yang sangat kecil. Bahkan asumsi independence antar atribut yang secara teori jarang terpenuhi pada data kesehatan nyata tidak menghambat performa model, yang menunjukkan bahwa pola hubungan antar indikator tetap dapat dipelajari dengan baik oleh algoritma.

Kesalahan klasifikasi yang terjadi justru memberikan insight penting. Data Rawat Inap yang terprediksi sebagai Non Rawat Inap kemungkinan berasal dari Puskesmas yang berada dalam kondisi borderline, misalnya fasilitas rawat inap minim, kapasitas tenaga kesehatan terbatas, atau sedang dalam masa transisi operasional. Sebaliknya, data Non Rawat Inap yang teridentifikasi sebagai Rawat Inap dapat menjadi indikasi bahwa Puskesmas tersebut memiliki kemiripan karakteristik dengan kategori Rawat Inap, sehingga dapat dipertimbangkan sebagai kandidat pengembangan layanan. Dengan demikian, kesalahan klasifikasi bukanlah kelemahan model, melainkan dapat dimanfaatkan sebagai bahan evaluasi strategis bagi Dinas Kesehatan.

Selain itu, munculnya label ganda “Non Rawat Inap” dengan spasi berbeda mencerminkan perlunya peningkatan kualitas data administratif. Konsistensi label kelas sangat penting untuk menghindari fragmentasi kelas pada model klasifikasi. Walaupun satu label tersebut tidak berpengaruh signifikan terhadap performa, hal ini menjadi catatan penting bahwa proses data cleaning harus dilakukan secara ketat sebelum pengolahan data.

Secara keseluruhan, temuan penelitian ini memperlihatkan bahwa Naive Bayes merupakan metode yang sangat efektif untuk memetakan kemampuan penyelenggaraan Puskesmas. Performanya yang tinggi menunjukkan bahwa indikator penilaian yang digunakan Kementerian Kesehatan telah cukup kuat dalam membedakan kapasitas pelayanan Puskesmas. Model ini dapat digunakan sebagai alat bantu pengambilan keputusan dalam monitoring kinerja, pengembangan layanan rawat inap, penentuan prioritas peningkatan sarana, serta identifikasi Puskesmas yang membutuhkan intervensi manajerial dan teknis. Dengan demikian, integrasi teknik data mining dalam evaluasi Puskesmas dapat memberikan manfaat strategis dalam upaya peningkatan mutu pelayanan kesehatan primer di Indonesia.

KESIMPULAN

Penelitian ini bertujuan untuk menganalisis kemampuan penyelenggaraan Puskesmas tahun 2024 menggunakan algoritma Naive Bayes berbasis RapidMiner. Berdasarkan hasil pemodelan, diperoleh bahwa algoritma Naive Bayes mampu mengklasifikasikan jenis layanan Puskesmas (Rawat Inap dan Non Rawat Inap) dengan tingkat akurasi yang sangat tinggi, yaitu 98,86%, dengan tingkat kesalahan hanya 1,14%. Confusion matrix menunjukkan bahwa sebagian besar data berhasil diklasifikasikan secara benar, sementara kesalahan klasifikasi umumnya terjadi pada Puskesmas yang memiliki karakteristik borderline atau menunjukkan kemiripan indikator dengan kategori lainnya.

Hasil ini membuktikan bahwa indikator-indikator penilaian Puskesmas memiliki kekuatan diskriminatif yang baik dan sangat relevan dalam membedakan kemampuan penyelenggaraan



layanan. Selain itu, performa Naive Bayes yang konsisten mengindikasikan bahwa metode ini sangat cocok digunakan dalam analisis data kesehatan yang bersifat kategorikal, sekaligus menegaskan bahwa RapidMiner merupakan platform yang efektif dan mudah digunakan untuk mendukung proses analisis berbasis data mining.

Secara keseluruhan, penelitian ini menegaskan bahwa integrasi teknik data mining ke dalam sistem evaluasi Puskesmas dapat meningkatkan efisiensi, ketepatan, serta kecepatan dalam proses pengambilan keputusan. Model klasifikasi ini berpotensi digunakan sebagai alat bantu monitoring kinerja Puskesmas, identifikasi layanan yang perlu ditingkatkan, serta dasar perencanaan pengembangan layanan kesehatan primer di masa depan.

SARAN

1. Perbaiki Kualitas Data

Diperlukan penyeragaman penulisan label kelas dan konsistensi pengisian indikator Puskesmas untuk menghindari fragmentasi kelas dan memastikan data yang dianalisis benar-benar bersih.

2. Pengembangan Penelitian Lanjutan

Penelitian selanjutnya dapat menggunakan algoritma lain seperti Random Forest, SVM, atau Gradient Boosting untuk dibandingkan dengan Naive Bayes sehingga dapat diketahui metode terbaik untuk dataset kesehatan Puskesmas.

3. Integrasi ke Sistem Monitoring

Model klasifikasi dapat diintegrasikan ke dalam dashboard monitoring kinerja Puskesmas sehingga Dinas Kesehatan dapat melakukan evaluasi secara real-time atau periodik.

4. Analisis Variabel Lebih Mendalam

Perlu dilakukan analisis lebih lanjut terhadap Puskesmas yang salah klasifikasi untuk mengetahui faktor-faktor penyebabnya dan menyediakan rekomendasi kebijakan yang tepat.

5. Penggunaan Dataset Lebih Luas

Model dapat diuji dengan dataset dari tahun berbeda atau wilayah yang lebih luas untuk mengetahui tingkat generalisasi model.

DAFTAR PUSTAKA

A. Pratama dan S. Hidayat, "Analisis Prediksi Pelayanan Kesehatan Menggunakan Metode Naive Bayes," *Jurnal Sistem Informasi dan Sains Data (JSISD)*, vol. 5, no. 1, pp. 45–53, 2021.

Badan Penelitian dan Pengembangan Kesehatan, "Laporan Nasional Riset Kesehatan Dasar (Riskesdas)," Jakarta: Kementerian Kesehatan RI, 2018.

Dinas Kesehatan Provinsi Jawa Tengah, "Profil Kesehatan Provinsi Jawa Tengah Tahun 2023," Semarang: Dinkes Jateng, 2024.

H. Santoso dan R. Widodo, "Penerapan Data Mining Menggunakan Algoritma Naive Bayes untuk Klasifikasi Dataset Kesehatan," *Jurnal Teknologi Informasi dan Komputer (JTIK)*, vol. 8, no. 2, pp. 112–120, 2022.

Kementerian Kesehatan Republik Indonesia, "Pedoman Penyelenggaraan Puskesmas," Direktorat Jenderal Pelayanan Kesehatan, Jakarta, 2020.



Kementerian Kesehatan Republik Indonesia, “Peraturan Menteri Kesehatan Republik Indonesia Nomor 43 Tahun 2019 tentang Puskesmas,” Jakarta: Kemenkes RI, 2019.

Kementerian Kesehatan Republik Indonesia, “Profil Kesehatan Indonesia 2023,” Jakarta: Kementerian Kesehatan RI, 2024.

L. Marliana dan T. W. Nugroho, “Implementasi Algoritma Naive Bayes untuk Klasifikasi Data Pelayanan Kesehatan,” *Jurnal Ilmiah Teknologi Informasi dan Kesehatan*, vol. 4, no. 3, pp. 187–195, 2020.

Pusat Data dan Informasi Kementerian Kesehatan (PUSDATIN), “Statistik Puskesmas Indonesia 2024,” Jakarta: Kemenkes RI, 2024.

S. R. Anggraini dan M. H. Putra, “Pemanfaatan RapidMiner dalam Analisis Data Kesehatan Puskesmas,” *Jurnal Informatika Nusantara*, vol. 7, no. 1, pp. 55–64, 2022