



Analisa Kenaikan Kunjungan Pasien Di Laboratotium Patologi Klinik Pada Tahun 2022 Sampai 2023 RSUD Kota Bandung Dengan Menggunakan Naive Bayes

Maulana Fansyuri¹, Lyra Kidunga², Nadya Wulandari Ega M^{3*}, Mutiara Laela S⁴, Rizkyanti Ajeng Trias M⁵, Ulfa⁶

^{1,2,3,4,5,6}Fakultas Ilmu Komputer, Progra Studi Teknik Informatika, Universitas Pamulang, Kota Tangerang Selatan, Indonesia

Email: ¹dosen02359@unpam.ac.id, ²kidungalyra@gmail.com, ^{3*}Nadyawlnndri20@gmail.com,
⁴mutiaralailaa18@gmail.com, ⁵rizkiantiajeng@gmail.com, ⁶ulfau9681@gmail.com

(* : coressponding author)

Abstrak—Penelitian ini bertujuan untuk menganalisis kenaikan kunjungan pasien laboratorium patologi klinik di RSUD Kota Bandung pada tahun 2022 hingga 2023 berdasarkan jenis pemeriksaan dengan menggunakan algoritma Naïve Bayes. Dataset yang digunakan berasal dari data laporan kunjungan pasien yang mencantumkan informasi jenis pemeriksaan yang dilakukan. Proses analisis meliputi tahapan preprocessing data, pembagian dataset, dan implementasi algoritma Naïve Bayes menggunakan aplikasi RappidMiner. Hasil penelitian menunjukkan bahwa algoritma Naïve Bayes mampu mengklasifikasikan data dengan akurasi sebesar 100%, dengan precision sebesar 100% dan rechall sebesar 100%. Penelitian ini diharapkan dapat mendukung pengelolaan dan analisis data kunjungan pasien secara lebih efektif, sehingga membantu dalam pengambilan Keputusan strategis rumah sakit.

Kata Kunci: Naïve Bayes, Laboratorium Patologi Klinik, Analisis Data, Akurasi, Precision, Recall, RappidMiner,

Abstract—This study aims to analyze the increase in patient visit to the clinical pathology laboratory at the Bandung city hospital from 2022 to 2023 based on the type of examination using the Naïve Bayes algorithm. The dataset used comes from patient visit report data that includes information on the type of examination performed. The analysis process includes data processing, dataset sharing, and implementation of the Naïve Bayes algorithm using the RappidMinerthe application, the result of the study show that the Naïve Bayes algorithm is able to classify data with an accuracy of 100%, with an accuracy of 100% and a recall of 100%. This research is expected to support the management and analysis of patient visit data more effectively, thereby helping in making strategic decisions for hospitals.

Keywords: Naïve Bayes, Clinical Pathology Laboratory, Data Analysis, Accuracy, Precision, Recall, RappidMiner

1. PENDAHULUAN

1.1 Latar Belakang

Laboratorium patologi klinik memegang peranan penting dalam proses diagnosis dan perawatan pasien di rumah sakit, karena hasil pemeriksaan laboratorium seringkali menjadi dasar utama penentuan langkah medis yang tepat. Setiap tahun, jumlah kunjungan pasien ke laboratorium terus meningkat, menghasilkan volume data yang besar dan beragam, terutama terkait jenis pemeriksaan yang dilakukan. Data yang terkumpul ini mengandung informasi berharga yang jika dikelola dengan baik, dapat memberikan wawasan mengenai kunjungan pasien dan jenis pemeriksaan yang paling sering dilakukan.

Namun, tantangan utama yang dihadapi adalah bagaimana mengelola dan menganalisis data yang begitu besar dan kompleks dengan cara yang efektif. Dalam konteks ini, penerapan metode data mining menjadi sangat relevan. Salah satu pendekatan yang sering digunakan adalah algoritma klasifikasi, seperti Naïve Bayes, yang memungkinkan analisis data berdasarkan karakteristik tertentu. Melalui penerapan algoritma ini, rumah sakit dapat memperoleh pola-pola yang dapat digunakan untuk memprediksi jenis pemeriksaan yang dibutuhkan oleh pasien di masa depan.

Dengan menggunakan metode ini, rumah sakit tidak hanya dapat meningkatkan efisiensi dalam proses diagnosis dan pengobatan, namun juga dapat merencanakan kebutuhan sumber daya yang lebih tepat dan terukur. Selain itu, pemanfaatan analisis data ini dapat membantu manajemen



rumah sakit dalam hal membuat keputusan strategis yang lebih baik, seperti dalam hal pengalokasian tenaga medis, pembelian perawatan, serta perencanaan kebijakan pelayanan laboratorium yang lebih efektif. Penelitian ini difokuskan pada analisis data kunjungan pasien RSUD Kota Bandung pada tahun 2022 hingga 2023, dengan tujuan untuk memanfaatkan algoritma Naïve Bayes dalam menghasilkan wawasan yang akurat mengenai pola-pola kunjungan dan jenis pemeriksaan yang dilakukan oleh pasien.

1.2 Tujuan Penelitian

1. Menganalisis pola kunjungan pasien.
2. Menerapkan algoritma Naïve Bayes dalam klasifikasi data.
3. Memprediksi tren kunjungan pasien.

2. LANDASAN TEORI

2.1 Data Mining dan Analisis

Data mining adalah proses eksplorasi dan penggalian pola, tren, atau informasi penting dari dataset yang besar dan kompleks (Han & Kamber, 2011). Dalam konteks pelayanan Kesehatan, data mining digunakan untuk menganalisis data kunjungan pasien, mengidentifikasi pola kenaikan kunjungan, serta memberikan wawasan strategis bagi pengelolaan rumah sakit (Setiawan, 2019). Analisis dengan data mining dapat mencakup identifikasi faktor-faktor yang mempengaruhi kenaikan kunjungan pasien, memprediksi tren di masa depan, dan mendukung pengambilan Keputusan berbasis data.

2.2 Algoritma Naïve Bayes

Algoritma Naïve Bayes adalah salah satu metode analisis dalam data mining yang berfungsi sebagai alat klasifikasi probabilistic. Metode ini berdasarkan pada Teorema Bayes dengan asumsi bahwa semua fitur bersifat independent. Walaupun asumsi independensi sering kali tidak sepenuhnya terpenuhi, algoritma ini tetap dikenal karena kesederhanaan, kecepatan, dan keefektifannya, terutama dalam pengolahan dataset besar dan kompleks (Sumijan, 2021). Dalam penelitian ini, Naïve Bayes digunakan untuk menganalisis kenaikan kunjungan pasien laboratorium patologi klinik pada tahun 2022 hingga 2023 di RSUD Kota Bandung. Melalui analisis ini, dapat dihasilkan klasifikasi yang membantu rumahsakit memahami tren kunjungan dan meningkatkan efisiensi pengelolaan data pasien.

3. METODE PENELITIAN

3.1 Pengklasifikasian Naïve Bayes

Pengklasifikasi Naïve Bayes (NBC) diambil dari teorema Bayes dengan asumsi independensi diantara variabel prediktor. Model Naïve Bayes mudah dibuat dan sangat berguna untuk prediksi pada kumpulan data yang lebih besar. Selain sederhana dan mudah dipahami, Pengklasifikasi Naïve Bayes memiliki kinerja yang dapat lebih unggul daripada klasifikasi yang sangat canggih. Teorema Bayes memiliki rumus sebagai berikut :

$$P(B|A) = \frac{P(A|B) \cdot P(A)}{P(B)}$$

Persamaan probabilitas Bayes dijabarkan menjadi :

P : Probabilitas

A : Kejadian yang ingin kita cari probabilitasnya setelah kejadian lain diketahui atau biasa disebut hipotesis.

B : Kejadian yang baru saja terjadi atau informasi yang baru diberikan.

P(B|A) : Probabilitas kejadian A terjadi jika B sudah terjadi (probabilitas bersyarat).



$P(B|A)$: Probabilitas kejadian B terjadi jika A sudah terjadi.

$P(A)$: Probabilitas awal sebelum kejadian A,

$P(B)$: Probabilitas awal sebelum kejadian B.

Dataset

Dataset yang digunakan dalam penelitian ini mencakup data kunjungan pasien laboratorium Patologi klinik di RSUD Kota Bandung. Atribut dataset meliputi :

1. Total Pasien Keseluruhan
2. Jenis pemeriksaan (misalnya, hematologi, kimia klinik, mikrobiologi)
3. Jumlah pasien setiap pemeriksaan
4. Presentase jumlah pemeriksaan pertahun

Aplikasi / Software

Pada penelitian kali ini, kami menggunakan aplikasi RapidMiner sebagai perangkat lunak untuk mengimplementasikan algoritma Naïve Bayes. RapidMiner merupakan aplikasi berbasis data mining yang banyak digunakan untuk analisis data dan penelitian dari berbagai bidang, termasuk implementasi algoritma Naïve Bayes. Dalam penelitian, RapidMiner mempermudah proses implementasi Naïve Bayes dengan antarmuka drag-and-drop, yang memungkinkan peneliti untuk mengimpor data, mempersiapkan atribut, melatih model, dan mengevaluasi hasil dengan cepat. Dengan fitur seperti evaluasi performa, confusion matrix, dan visualisasi data, RapidMiner mendukung proses analisis yang mendalam, membantu peneliti mengeksplorasi akurasi, precision, dan recall dari model Naïve Bayes. Aplikasi ini juga berguna untuk studi prediksi, pengelompokan data, atau identifikasi pola dalam dataset besar dengan efisiensi tinggi tanpa memerlukan keahlian pemrograman yang mendalam.

3.2 Tahapan Penelitian

Penelitian ini terdiri dari beberapa tahapan yang dilakukan untuk mencapai tujuan analisis kenaikan kunjungan pasien di laboratorium patologi klinik RSUD Kota Bandung pada tahun 2022 hingga 2023. Adapun tahapan penelitian adalah sebagai berikut :

1. Pengumpulan Data

Data yang digunakan dalam penelitian ini diperoleh dari laporan kunjungan pasien di laboratorium patologi klinik RSUD Kota Bandung (opendata.bandung.go.id). Dataset ini mencakup informasi mengenai kode dan nama provinsi, kode dan nama kabupaten/kota, jenis pemeriksaan, jumlah pasien per pemeriksaan, dan tahun pemeriksaan. Dataset inilah yang kemudian akan menghasilkan model prediksi kenaikan atau penurunan kunjungan pasien.

Dataset awal terdiri dari 64 baris dengan 8 atribut yang digambarkan pada gambar dibawah. Namun kami hanya menggunakan data dari tahun 2022 sampai 2023 untuk penelitian ini

id	kode_provinsi	nama_provinsi	bps_kode_kabupaten_kot	bps_nama_kabupaten_kot	jenis_pemeriksaan	jumlah_pasien	tahun
1	32	JAWA BARAT	3273	KOTA BANDUNG	KIMIA KLINIK	46372 ORANG	2015
2	32	JAWA BARAT	3273	KOTA BANDUNG	GULA DARAH	34571 ORANG	2015
3	32	JAWA BARAT	3273	KOTA BANDUNG	HEMATOLOGI	227125 ORANG	2015
4	32	JAWA BARAT	3273	KOTA BANDUNG	SEROLOGI	3591 ORANG	2015
5	32	JAWA BARAT	3273	KOTA BANDUNG	BAKTERIOLOGI	1889 ORANG	2015
6	32	JAWA BARAT	3273	KOTA BANDUNG	URINE	12409 ORANG	2015
7	32	JAWA BARAT	3273	KOTA BANDUNG	TINJAU (FESES)	1125 ORANG	2015
8	32	JAWA BARAT	3273	KOTA BANDUNG	KIMIA KLINIK	56791 ORANG	2016
9	32	JAWA BARAT	3273	KOTA BANDUNG	GULA DARAH	40044 ORANG	2016
10	32	JAWA BARAT	3273	KOTA BANDUNG	HEMATOLOGI	303157 ORANG	2016
11	32	JAWA BARAT	3273	KOTA BANDUNG	SEROLOGI	3185 ORANG	2016
12	32	JAWA BARAT	3273	KOTA BANDUNG	BAKTERIOLOGI	1738 ORANG	2016
13	32	JAWA BARAT	3273	KOTA BANDUNG	URINE	13507 ORANG	2016
14	32	JAWA BARAT	3273	KOTA BANDUNG	TINJAU (FESES)	880 ORANG	2016
15	32	JAWA BARAT	3273	KOTA BANDUNG	KIMIA KLINIK	55392 ORANG	2017
16	32	JAWA BARAT	3273	KOTA BANDUNG	GULA DARAH	41226 ORANG	2017
17	32	JAWA BARAT	3273	KOTA BANDUNG	HEMATOLOGI	258700 ORANG	2017
18	32	JAWA BARAT	3273	KOTA BANDUNG	SEROLOGI	3346 ORANG	2017
19	32	JAWA BARAT	3273	KOTA BANDUNG	BAKTERIOLOGI	1364 ORANG	2017
20	32	JAWA BARAT	3273	KOTA BANDUNG	URINE	13774 ORANG	2017
21	32	JAWA BARAT	3273	KOTA BANDUNG	TINJAU (FESES)	950 ORANG	2017

2. Preprocessing Data

Tahap preprocessing dilakukan untuk menyiapkan data sebelum dianalisis. Proses ini meliputi pembersihan data (data cleaning) untuk menghilangkan data yang tidak relevan atau hilang (missing values), serta konversi data ke dalam format yang dapat diproses oleh algoritma Naïve Bayes. Kami memangkas data yang tidak kami gunakan serta menambahkan kolom total pasien dan presentase jumlah pasien.

3. Pembagian Dataset

Dataset yang telah diproses kemudian dibagi menjadi dua bagian : data pelatihan (training data) dan data uji (testing data). Pembagian ini dilakukan dengan menggunakan Teknik pembagian 50% untuk data pelatihan dan 50% untuk data uji.

id	jenis_pemeriksaan	jumlah_pasien	total_pasien	persentase_2022	id	jenis_pemeriksaan	jumlah_pasien	total_pasien	persentase_2023
1	KIMIA KLINIK	31506	545680	5.773713532	1	KIMIA KLINIK	32907	492803	6.677516168
2	GULA DARAH	37635	545680	6.896899282	2	GULA DARAH	35218	492803	7.146466235
3	HEMATOLOGI	448113	545680	82.12010702	3	HEMATOLOGI	402254	492803	81.62572062
4	SEROLOGI	20394	545680	3.737355227	4	SEROLOGI	12914	492803	2.620519761
5	BAKTERIOLOGI	184	545680	0.033719396	5	BAKTERIOLOGI	287	492803	0.058238282
6	URINE	7186	545680	1.316889019	6	URINE	8387	492803	1.701897107
7	TINJA (FESES)	662	545680	0.121316523	7	TINJA (FESES)	836	492803	0.169641824

4. Implementasi Algoritma Naïve Bayes

Algoritma Naïve Bayes diimplementasikan menggunakan aplikasi RapidMiner. Model yang dihasilkan kemudian diuji menggunakan data uji untuk mengukur kinerja model berdasarkan metrik akurasi, precision, dan recall. Hasil model yang dihasilkan dievaluasi dengan menggunakan metrik-metrik tersebut dan dibandingkan dengan nilai yang diharapkan, guna menentukan seberapa efektif algoritma Naïve Bayes dalam mengklasifikasikan jenis pemeriksaan dan menganalisis kenaikan kunjungan pasien. Untuk menganalisa kenaikan dan penurunan kunjungan pasien kami menggunakan rumus **Persentase 2023 – persentase 2022**.

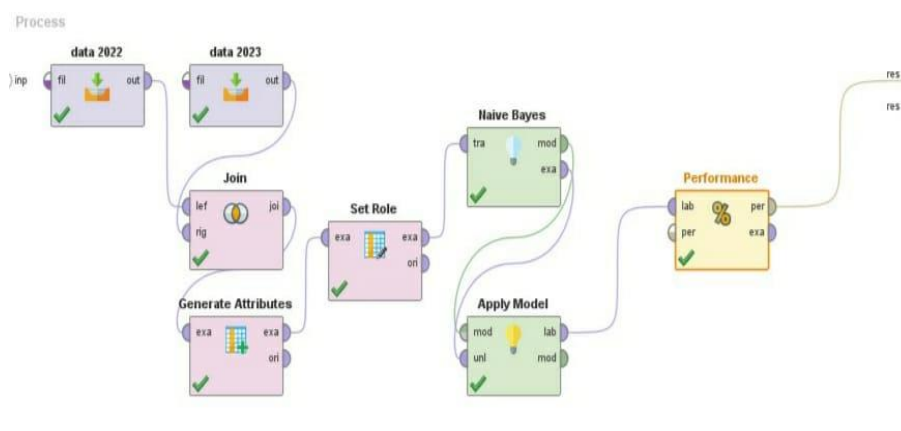
Berikut adalah perhitungan kenaikan dan penurunan kunjungan pasien :

- Kimia Klinik : $6.677516168 - 5.773713532 = 0.904$
- Gula Darah : $7.146466235 - 6.896899282 = 0.250$
- Hematologi : $81.62572062 - 82.12010702 = -0.494$
- Serologi : $2.62572062 - 3.737355227 = -1.117$
- Bakteriologi: $0.058238282 - 0.033719396 = 0.025$
- Urine : $1.701897107 - 1.316889019 = 0.385$
- Tinja (Feses) : $0.169641824 - 0.121316523 = 0.048$

4. ANALISA DAN PEMBAHASAN

Implementasi data mining dilakukan oleh aplikasi RapidMiner. Dataset untuk penelitian kali ini diperoleh dari data kunjungan pasien Laboratorium Patologi Klinik RSUD Kota Bandung yang tersedia pada website opendata.bandung.go.id. Atribut yang terdapat pada kumpulan data kunjungan pasien antara lain id, Kode Provinsi, Nama Provinsi, Kode Kabupaten/Kota, Nama Kabupaten/Kota, Jenis Pemeriksaan, Jumlah Pasien, Status, dan Tahun. Namun kami memodifikasi dataset tersebut agar sesuai dengan kebutuhan penelitian ini. Sehingga dataset yang digunakan beratribut, Id, Jenis Pemeriksaan, Jumlah Pasien, Total Keseluruhan Pasien, dan Presentase Jumlah Pasien. Jenis pemeriksaan ada tujuh yaitu Kimia Klinik, Gula Darah, Hematologi, Serologi, Bakteriologi, Urine, dan Tinja(Feses).

Dibawah ini merupakan desain solusi RapidMiner untuk analisa data kunjungan pasien menggunakan Klasifikasi Naïve Bayes.



Tujuan utama dari penelitian ini adalah untuk menganalisa kenaikan presentase kunjungan pasien laboratorium patologi dari tahun 2022 sampai 2023. Desain RapidMiner diatas menggambarkan alur proses analisis yang dilakukan di RapidMiner. Proses diawali dengan mengimpor dataset tahun 2022 dan dataset tahun 2023 dengan operator read excel, yang kemudian digabungkan dengan menggunakan operator Join. Setelah itu, atribut tambahan seperti perhitungan selisih presentase dan label kenaikan dan penurunan dihasilkan melalui proses Generate Atributes, dan peran atribut ditentukan menggunakan Set Roles. Data yang telah dipersiapkan kemudian digunakan untuk melatih model menggunakan algoritma Naïve Bayes. Model ini kemudian diterapkan pada data uji, dan hasilnya dievaluasi menggunakan matriks Performance.

Hasil prediksi model Naïve Bayes dengan menggunakan perangkat lunak RapidMiner ditunjukkan dengan gambar dibawah.

RowNo.	label_perub...	prediction[la...	confidence[...	confidence[...	id	jenis_pemer...	jumlah_pasi...	total_pasien	persentase...	persentase...	selisih
1	Naik	Naik	1.000	0.000	1	KIMIA KLINIK	31506	545680	5.774	6.678	0.904
2	Naik	Naik	1.000	0.000	2	GULA DARAH	37635	545680	6.897	7.146	0.250
3	Turun	Turun	0	1	3	HEMATOLOGI	448113	545680	82.120	81.626	-0.494
4	Turun	Turun	0.192	0.808	4	SEROLOGI	20394	545680	3.737	2.621	-1.117
5	Naik	Naik	1.000	0.000	5	BAKTERIOLOGI...	184	545680	0.034	0.058	0.025
6	Naik	Naik	1.000	0.000	6	URINE	7186	545680	1.317	1.702	0.385
7	Naik	Naik	1.000	0.000	7	TINJA (FESES)	662	545680	0.121	0.170	0.048

Gambar diatas menampilkan hasil prediksi model Naïve Bayes yang diterapkan pada data uji. Tabel ini menunjukkan label aktual, prediksi model, dan tingkat kepercayaan(Confidence) untuk masing masing prediksi. Model memberikan hasil prediksi dengan confidence yang tinggi bagi sebagian besr data.

Walaupun terdapat satu prediksi yang lebih rendah pada kelas Naik. Namun secara keseluruhan, hasil predksi mencerminkan performa model yang sangat baik.

Hasil pengujian dari kedua contoh data uji menunjukkan bahwa ada lima jenis pemeriksaan yang persentase kunjungannya naik dari tahun sebelumnya yaitu Kimia Klinik, Gula Darah, Bakteriologi, Urine, dan Tinja(Feses), serta dua jenis pemeriksaan yang persentase kunjungannya turun dari tahun sebelumnya yaitu Hematologi dan Serologi. Hal ini sesuai dengan data asli dimana selisih persentase pemeriksaan Kimia Klinik, Gila Darah, Bakteriologi, Urine, dan Tinja(Feses) hasilnya positif(naik), sedangkan persentase pemeriksaan Hematologi dan Serologi hasilnya negatif(turun).



Selanjutnya, perhitungan performa klasifikasi dilakukan dengan operator Performance yang menampilkan hasil performa model dalam bentuk Confusion Matrix yang ditunjukkan pada gambar di bawah ini.

accuracy: 100.00%

	true Naik	true Turun	class precision
pred. Naik	5	0	100.00%
pred. Turun	0	2	100.00%
class recall	100.00%	100.00%	

Gambar di atas memperlihatkan hasil performa model Naïve Bayes yang digunakan dalam analisis. Performa klasifikasi menunjukkan kinerja yang baik dengan akurasi sebesar 100%, dengan seluruh data uji pada kelas Turun dan Naik berhasil diklasifikasikan dengan benar. Confusion Matrix menunjukkan bahwa tidak ada kesalahan prediksi yang menghasilkan Precision sebesar 100%, dan Recall sebesar 100% untuk kedua kelas. Hasil ini menunjukkan bahwa algoritma Naïve Bayes dapat mengklasifikasikan data kunjungan pasien dengan cukup baik, serta memprediksi kenaikan dan penurunan jenis pemeriksaan yang dilakukan pada tahun 2022 dan 2023.

5. KESIMPULAN

Penelitian ini menunjukkan bahwa algoritma Naïve Bayes merupakan metode yang efektif untuk menganalisis kenaikan kunjungan pasien di laboratorium patologi klinik RSUD Kota Bandung pada tahun 2022 hingga 2023. Dengan menggunakan aplikasi RapidMiner, penelitian ini berhasil mengklasifikasikan data kunjungan pasien berdasarkan jenis pemeriksaan dengan tingkat akurasi, precision, dan recall yang masing-masing mencapai 100%. Hasil analisis menunjukkan adanya kenaikan kunjungan pasien pada lima jenis pemeriksaan, yaitu Kimia Klinik, Gula Darah, Bakteriologi, Urine, dan Tinja (Feses), sementara dua jenis pemeriksaan lainnya, yaitu Hematologi dan Serologi, mengalami penurunan. Proses analisis yang melibatkan tahapan seperti preprocessing data, pembagian dataset, dan penerapan algoritma Naïve Bayes berjalan dengan baik, menghasilkan model yang mampu memberikan prediksi yang akurat serta wawasan strategis bagi rumah sakit. Dengan demikian, penelitian ini berkontribusi pada pengelolaan data pasien secara lebih efisien, mendukung pengambilan Keputusan strategis berbasis data, serta memberikan dasar untuk perencanaan layanan Kesehatan yang lebih baik di masa depan.

REFERENCES

- Han, J. &. (2011). *Data Mining: Concepts and Techniques*. Elsevier.
- Pratama, Y. (2020). Implementasi Data Mining dalam Sektor Kesehatan. *Jurnal Teknologi Informasi dan Komputer*, 7(3), 45-53.
- Setiawan, D. (2019). Aplikasi Data Mining untuk Analisis Data Pasien Rumah Sakit. *Jurnal Sistem Informasi Indonesia*, 5(4), 88-96.
- Sumijan, M. (2021). Algoritma Naïve Bayes untuk Prediksi Tren Data Kesehatan. *Jurnal Informatika dan Sistem Informasi*, 8(2), 22-30.