



Artikel

PERANCANGAN APLIKASI PREDIKSI PENYAKIT JANTUNG MENGGUNAKAN METODE *NAÏVE BAYES*

Adrian Timotius¹, Indah Fenriana²

^{1,2,3} Teknik Informatika, Universitas Buddhi Dharma, Banten, Indonesia

SUBMISSION TRACK

Received: March 28, 2017
Final Revision: May 03, 2017
Available Online: May 15, 2017

KEYWORD

Algoritma, Data Mining, Penyakit Jantung, *Naïve Bayes*

KORESPONDENSI

E-mail: adrianpru1811@gmail.com
indah.f88@gmail.com

A B S T R A K

Terdapat keterbatasan informasi mengenai cara mendiagnosa penyakit di masyarakat, terutama gejala terkait dalam memprediksi penyakit jantung. Algoritma *Naïve Bayes* merupakan salah satu algoritma yang sering digunakan dalam data mining untuk memprediksi kemungkinan terjadinya penyakit jantung berdasarkan faktor resiko yang ada. Adapun masalah yang ingin diselesaikan adalah bagaimana hasil diagnosa yang diperoleh dari perancangan sistem menggunakan algoritma *Naïve Bayes* dalam mendiagnosa penyakit jantung. Teknik penelitian ini Menggunakan Algoritma *Naïve Bayes*. Sedangkan untuk Metode pengumpulan data adalah dengan dataset dan studi pustaka. Adapun hasil penelitiannya Algoritma *Naïve Bayes* dalam bidang ilmu komputer data mining, khususnya dalam pengklasifikasian penyakit jantung, terbukti dapat di implementasikan dengan baik. Pengujian yang dilakukan menggunakan metode manual dan menggunakan *software Rapid Minner* untuk menghasilkan akurasi sebagai ukuran tingkat keakuratan algoritma dalam mendiagnosa penyakit jantung. Selain itu juga memiliki tingkat akurasi yang lebih tinggi dibandingkan dengan algoritma *Regresi Linear*. metode *Naïve Bayes* memiliki akurasi sebesar 86.26%, sedangkan algoritma *Regresi Linear* hanya memiliki akurasi sebesar 83%.

Pendahuluan

Penyakit jantung merupakan suatu kondisi medis yang terjadi ketika terdapat gangguan pada jantung atau pembuluh darah di sekitarnya. Ini mencakup beberapa masalah seperti penyakit jantung koroner, kegagalan jantung, gangguan irama jantung, dan gangguan katup jantung. Beberapa faktor yang dapat menjadi penyebab penyakit jantung termasuk kebiasaan makan yang tidak sehat, kurangnya aktivitas fisik, kebiasaan merokok, tekanan darah tinggi, kadar kolesterol tinggi, diabetes, obesitas, dan faktor genetik. Meskipun penyakit jantung tetap menjadi tantangan global, penelitian terus dilakukan untuk meningkatkan pemahaman tentang penyakit ini, mengembangkan metode deteksi yang lebih baik, dan mencari terapi yang lebih efektif. Upaya yang berkelanjutan dalam deteksi dini, pencegahan, dan pengobatan diharapkan dapat secara signifikan mengurangi tingkat kematian akibat penyakit jantung di masa depan.

Oleh karena itu metode yang dapat digunakan untuk menganalisis data kesehatan dan memprediksi kemungkinan terjadinya penyakit jantung adalah data mining. Algoritma *Naive Bayes* merupakan salah satu algoritma yang sering digunakan dalam data mining untuk memprediksi kemungkinan terjadinya penyakit jantung berdasarkan faktor resiko yang ada.

Jantung merupakan organ vital yang berfungsi sebagai pompa untuk memompa darah ke seluruh tubuh. Ketika jantung mengalami gangguan, penanganan medis diperlukan. Penyakit jantung saat ini tetap menjadi penyebab kematian utama di dunia, terutama pada kelompok usia orang dewasa dan lansia. Diagnosa penyakit jantung bertujuan untuk menentukan jenis penyakit dan melakukan pemeriksaan serta analisis gejala yang terjadi.

Berdasarkan data riset kesehatan dasar, penyakit kardiovaskular seperti penyakit jantung terus mengalami peningkatan dan menjadi penyebab kematian tertinggi di Indonesia, terutama pada usia-usia produktif.

jumlah kematian akibat penyakit jantung secara global mencapai 18,6 juta orang setiap tahunnya. Angka kematian tersebut diperkirakan akan terus meningkat menjadi 20,5 juta orang pada tahun 2020 Bahkan penyakit jantung ini menjadi beban biaya terbesar. Berdasarkan data BPJS Kesehatan pada 2021 pembiayaan kesehatan terbesar ada pada penyakit jantung sebesar Rp.7,7 triliun. [1]. Penelitian terus dilakukan untuk meningkatkan pemahaman tentang penyakit ini, Oleh karena itu metode yang dapat digunakan untuk menganalisis data kesehatan dan memprediksi kemungkinan terjadinya penyakit jantung adalah data mining. Algoritma *Naive Bayes* merupakan salah satu algoritma yang sering digunakan dalam data mining untuk memprediksi kemungkinan terjadinya penyakit jantung berdasarkan faktor resiko yang ada. untuk membuat suatu program komputer untuk mendiagnosa penyakit jantung menggunakan data set yang peneliti dapat dari kaggle yaitu *hearth failure prediction* dengan menerapkan Algoritma *Naive Bayes* dalam mengetahui tingkat akurasi yang dihasilkan algoritma tersebut, diharapkan dapat memudahkan pengguna untuk mendiagnosa penyakit jantung secara efektif dan memperoleh hasil diagnosa penyakit jantung yang akurat dengan Algoritma *Naive Bayes*.

I. METODE

1.1 Data

Data adalah kumpulan informasi atau fakta yang dikumpulkan dari hasil pengamatan atau pencarian dari berbagai sumber. Secara umum, data terdiri dari elemen dasar seperti simbol, angka, kata-kata, atau gambar. Namun, data dalam bentuknya yang awalnya belum diproses belum memiliki banyak makna yang jelas. Oleh karena itu, diperlukan tahap pengolahan tambahan menggunakan suatu model agar data dapat menghasilkan informasi yang berarti dan kesimpulan yang berguna, serta dapat digunakan sebagai dasar dalam proses pengambilan keputusan [2]

Data adalah informasi yang digunakan untuk menggambarkan suatu kejadian (*event*) dalam bentuk deskripsi kenyataan. Data terdiri dari fakta dan angka, meskipun belum tentu semuanya memiliki signifikansi yang sama bagi penerima data [3].

1.2 Data mining

Kata *Mining* merupakan kiasan dari Bahasa Inggris, mine. Jika *mine* berarti menambang sumber daya yang tersembunyi di dalam tanah, maka Data Mining merupakan penggalian makna yang tersembunyi dari kumpulan data yang sangat besar. [4] Karena itu Data Mining sebenarnya memiliki akar yang panjang dari bidang ilmu seperti kecerdasan buatan (*artificial intelligent*), *machine learning*, statistik dan basis Data. Data mining merupakan sebuah proses ekstraksi data dari berbagai sumber informasi yang memiliki nilai. [5] Tujuan utama dari data mining adalah untuk mengungkap pola-pola yang tersembunyi dalam data tersebut. Pola-pola ini dapat ditemukan dalam beragam jenis basis data, termasuk basis data relasional, data *warehouse*, data transaksi, dan data berorientasi objek. [6] Penerapan teknik data mining memberikan keuntungan signifikan bagi para pebisnis karena mampu memberikan dukungan dalam proses pengambilan keputusan yang lebih cepat dan akurat [7].

1.3 Algoritma Regresi linear

Regresi linier adalah teknik analisis statistik yang digunakan untuk memprediksi hubungan antara dua atau lebih variabel. Variabel ini dapat dibagi menjadi dua jenis: variabel terikat (*dependen*; Y) dan variabel bebas (*independen*; X). Regresi linier sederhana merujuk pada model di mana hanya ada satu variabel bebas, sementara regresi

linier berganda melibatkan lebih dari satu variabel bebas [8].

1.4 Naïve bayes

Algoritma *Naïve Bayes* adalah bahwa metode ini membutuhkan jumlah data pelatihan yang relatif kecil untuk mengestimasi parameter yang diperlukan dalam proses klasifikasi. Metode ini mengasumsikan adanya keterkaitan antara variabel, sehingga hanya variasi dalam variabel yang diperlukan untuk melakukan klasifikasi [9].

Rumus Algoritma naïve bayes :

$$P(H|E) = \frac{P(E|H) \times P(H)}{P(E)}$$

$$P(E)$$

Dalam aturan Bayes, terdapat beberapa konsep yang penting, antara lain:

1. $P(H|E)$: Probabilitas kondisional akhir suatu hipotesis H terjadi jika diberikan bukti E.
2. $P(E|H)$: Probabilitas bukti E terjadi jika hipotesis H terjadi.
3. $P(H)$: Probabilitas awal hipotesis H terjadi tanpa mempertimbangkan bukti apapun.
4. $P(E)$: Probabilitas awal bukti E terjadi tanpa mempertimbangkan hipotesis/bukti lainnya.

Aturan *Bayes* memberikan cara untuk memperkirakan probabilitas akhir hipotesis berdasarkan bukti yang diamati. Probabilitas awal (*prior*) H, atau $P(H)$, adalah probabilitas hipotesis sebelum bukti diamati..

- Probabilitas akhir H atau $P(H|E)$ adalah probabilitas hipotesis setelah bukti diamati.

Dalam contoh ini, data yang akan diuji adalah:

Fbs (kadar gula darah) = salah dan exang (induksi angina) = tidak.

Dari rumus di atas, data yang diambil adalah:

1. Jumlah data = 918.
2. Jumlah kemunculan positif = 508 => $p(\text{positif}) = 508/918 = 0,55$.
3. Jumlah kemunculan negatif = 410 => $p(\text{negatif}) = 410/918 = 0,44$.

Tahap awal pada proses dalam memperhitungkan naïve bayes ialah dengan cara pengambilan data pelatihan dari dataset heart failure prediction yang sebelumnya sudah ter input pada program. Data akan di bagi menjadi 11 atribut tipe data dan 1 untuk hasil. Berikut atribut yang di gunakan antara lain ialah:

age, sex, chest pain, restingBP, chol, fastingBS, exercise angina, , ST-Slope, dan Heart Disease dimana data akan di bagi menjadi 2 parameter yaitu negatif dan positif yang menentukan dimana masing – masing parameter atribut nantinya. Tiap parameter atribut di peroleh di dataset sebelumnya kemudia di hitung sehingga bisa memperoleh probabilitas parameter hasil yang di inginkan.

Berdasarkan hasil tinjauan pustaka yang penulis dapatkan untuk mendiagnosa penyakit jantung menggunakan data set yang peneliti dapet dari *kaggle* yaitu *heart failure prediction* dengan menerapkan *Algoritma naïve bayes* dalam mengetahui tingkat akurasi yang dihasilkan algoritma tersebut, diharapkan dapat memudahkan pengguna untuk mendiagnosa penyakit jantung secara efektif dan memperoleh hasil diagnosa penyakit jantung yang akurat dengan *Algoritma naïve bayes*.

1.5 Penyakit jantung

Jantung berfungsi sebagai organ dalam tubuh yang bertugas memompa darah, dan ia bekerja sejak janin dalam kandungan sampai kita lahir. Jantung menggunakan ototnya untuk memberikan rangsangan dan mendorong darah ke seluruh tubuh.[10] Otot jantung terdiri dari jaringan saraf untuk memberikan rangsangan berdenyut secara

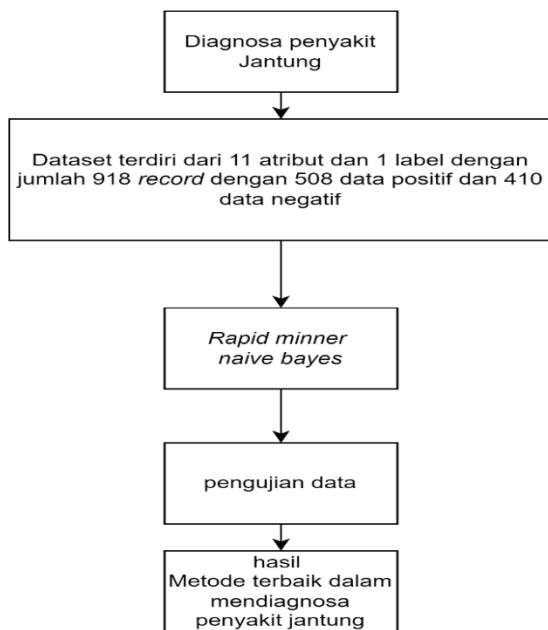
teratur dan otomatis kepada jantung. Di setiap denyutnya, jantung akan mengalirkan darah yang mengandung oksigen serta zat-zat yang kita konsumsi ke seluruh badan , termasuk juga ke arteri koroner, dan mengalirkan darah yang mengandung sedikit oksigen kedalam paru-paru untuk memperoleh oksigen. [11] Sakit jantung sering kali menghasilkan kerusakan pada sel-sel jantung yang mengganggu aliran darah ke seluruh tubuh karena kurangnya pasokan oksigen yang dibawa oleh pembuluh darah di dalam jantung. Hal ini juga bisa disebabkan oleh kejang pada otot jantung yang mengakibatkan gangguan dalam pemompaan darah. Penyakit jantung terdiri dari enam jenis yang berbeda, termasuk Gagal Jantung Akut, Jantung Koroner, Jantung Hipertensi, Gagal Jantung Kronik, Jantung Katup, dan Jantung Perikarditif. Proses diagnosis penyakit jantung melibatkan penginputan identitas pasien dan kondisi yang dialami pasien, termasuk gejala dan faktor risiko, sebagai kasus baru. Selanjutnya, tingkat kesamaan atau similaritas dihitung berdasarkan faktor usia, jenis kelamin, gejala, dan faktor risiko dengan kasus-kasus sebelumnya yang tersimpan dalam basis data kasus. Penilaian ini kemudian dikalikan dengan tingkat keyakinan. Setiap usia, jenis kelamin, gejala, dan faktor risiko memiliki bobot dengan nilai tertentu yang ditentukan berdasarkan jenis penyakit yang diderita oleh pasien.

Penyakit jantung merupakan istilah yang digunakan secara umum untuk menggambarkan berbagai gangguan yang mempengaruhi fungsi jantung[12]. Perlu dibedakan antara penyakit jantung dan penyakit kardiovaskular, yang mengacu pada kondisi di mana pembuluh darah mengalami penyempitan atau penyumbatan yang dapat menyebabkan serangan jantung, nyeri dada, atau stroke. Salah satu jenis penyakit

kardiovaskular yang umum adalah penyakit jantung koroner.

Penyakit jantung menjadi salah satu penyebab kematian di seluruh dunia, karena itu, penting untuk melakukan diagnosis dini yang akurat dan menjadi fokus penelitian medis. Proses diagnosis penyakit jantung kompleks karena terdapat hubungan kompleks antara berbagai faktor.

1.6 kerangka pemikiran

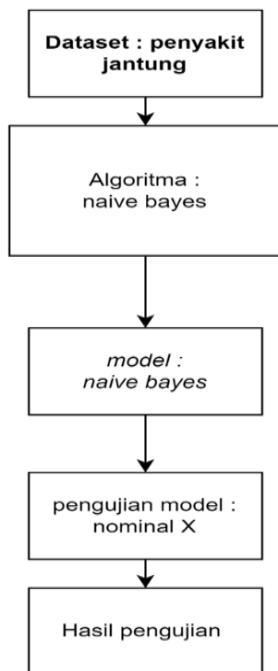


Gambar 1. kerangka pemikiran

II. PERANCANGAN

2.1 Tahap Persiapan

Tahap persiapan dan seleksi data yang telah dikumpulkan dan diubah menjadi bentuk yang dapat diolah dalam model yang ditentukan selanjutnya [13].



Gambar 2. Alur Tahap Persiapan

Tahapan pemahaman data dimulai dengan pengumpulan data awal dan hasil kegiatan

-Tampilan Dataset

Berikut ini adalah tampilan dataset “*hearth failure prediction*” sebelum dilakukan proses yang *Data Preparation* yang selanjutnya:

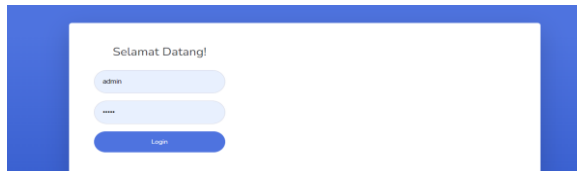
Penelitian ini menggunakan metode pembelajaran mesin dalam prediksi penyakit jantung yaitu *naïve bayes* dan Dataset yang telah diperoleh dibagi sesuai dengan kriteria kemudian dilakukan pengujian menggunakan metode *naïve bayes* tersebut untuk

dalam rangka untuk membiasakan diri dengan data untuk mengidentifikasi masalah data dan mendeteksi subset menarik untuk membentuk hipotesis untuk informasi yang tersembunyi[14].

2.2 Desain Aplikasi

Aplikasi direncanakan dibangun dalam bentuk berbasis web dengan menggunakan Visual Studio Code, serta menggunakan bahasa pemrograman PHP, C#, dan HTML.

2.2.1 TAMPILAN LOGIN



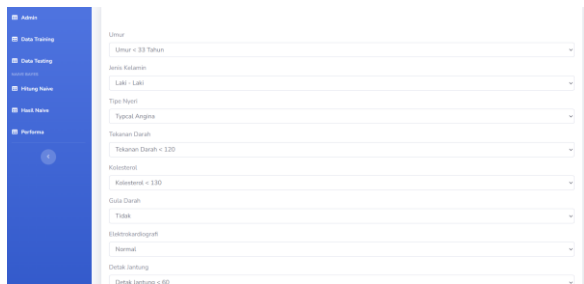
Gambar 3. Tampilan login

2.2.2 TAMPILAN MENU



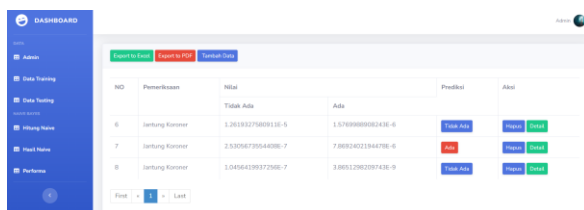
Gambar 4. Tampilan menu

2.2.3 TAMPILAN HALAMAN INPUT NAÏVE BAYES



Gambar 5. Tampilan input naïve bayes

2.2.4 TAMPILAN HASIL DATA YANG SUDAH DI INPUT



Gambar 6. Tampilan hasil data yang di input

III. PEMBAHASAN

3.1 Pengumpulan data

Pada penelitian kali ini, penulis menggunakan dataset “*hearth failure prediction*” Dataset yang akan di teliti didapat dari sebuah situs penyedia data kaggle [15]

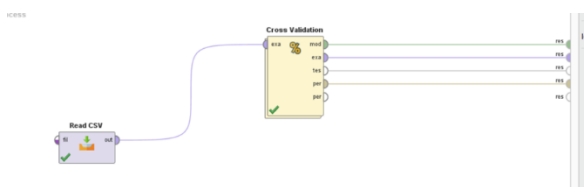
pada dataset tersebut memiliki 918 Record, dengan 12 atribut yang digunakan penelitian ini antara lain:

Tabel 1. Atribut dari dataset

No	Atribut
1	Umur
2	Jenis kelamin
3	<i>Chestpain</i>
4	<i>Resting blood presure</i>
5	<i>cholesterol</i>
6	<i>Fasting blood sugar</i>
7	<i>Resting ECG</i>
8	<i>Max HR</i>
9	Exercise angina
10	Old peak
11	St – slope
12	Hasil

3.2 Modeling

Berikut ini adalah model yang digunakan untuk mengelolah dataset yang menggunakan algoritma naïve bayes:

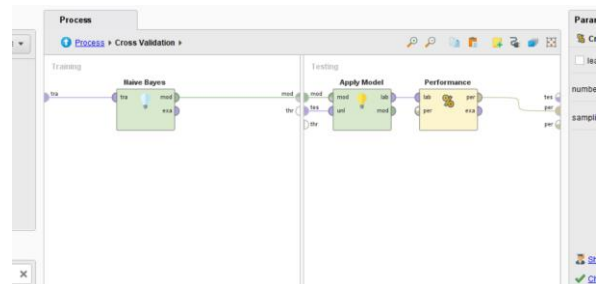


Gambar 7. Model pengolahan Algoritma *Niave bayes* pada *Rapid Minner* Studio

Fungsi tiap operator

1. Read cvs : operator ini berfungsi untuk memasukan dataset dalam bentuk file csv

2. *Cross validation* : salah satu teknik penting dalam pengembangan model dan evaluasi kinerja model dalam machine learning dan statistik. Tujuan utamanya adalah untuk mengukur sejauh mana model statistik atau algoritma machine learning dapat digeneralisasi ke data yang belum pernah dilihat sebelumnya



Gambar 8. Proses Pengolahan Model Algoritma *naïve bayes* pada *rapidminner* studio

Fungsi tiap operator

1. *naïve bayes* : mengklasifikasikan data ke dalam kategori atau kelas yang sesuai berdasarkan atribut-atribut yang ada
2. *apply model* : proses penggunaan model yang telah Anda kembangkan atau latih. Ini adalah langkah di mana Anda mengambil model yang telah diperoleh dari pelatihan dan menggunakan model tersebut untuk melakukan prediksi atau analisis pada data yang baru atau belum terlihat sebelumnya.
3. *Perfomence* : proses penting untuk mengukur seberapa baik model Anda bekerja pada data yang telah diolah.

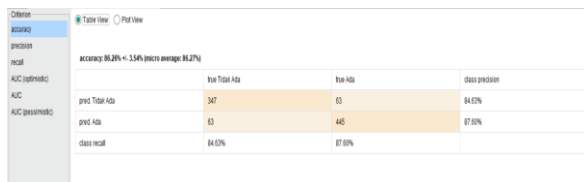
3.3 Evaluasi

Tahapan ini untuk melakukan validasi dan pengukuran keakuratan hasil pengolahan dengan Algoritma pada Software RapidMiner kemudian melakukan Analisis terhadap hasil yang diperoleh dari Penelitian ini [16].

3.3.1 Fase Permodelan

Metode yang di gunakan untuk melakukan komparasi adalah naïve bayes untuk memprediksi penyakit jantung, dalam hal ini yang di gunakan dataset yang terdiri dari 12 atribut, metode yang akan di uji merupakan metode sebelumnya dan metode naïve bayes dan hasil yang akan di sajikan adalah *accuracy* dan *auc*

3.3.2 Naïve bayes



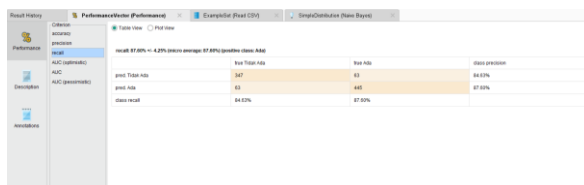
accuracy: 86.25% (+ 1.54% (info average 86.27%))				
	True Target Ada	True Ada	class precision	
AUC (optimistic)	347	63	84.63%	
AUC	pred True Ada	63	84.63%	
AUC (pessimistic)	pred Ada	445	87.68%	
class recall	84.63%	87.68%		

Gambar 9. hasil perhitungan *accuracy* memakai algoritma *naïve bayes* dengan *rapid minner studio*



precision: 87.26% (+ 1.88% (info average 87.88%) (positive class Ada))				
	True Target Ada	True Ada	class precision	
AUC (optimistic)	347	63	84.63%	
AUC	pred True Ada	63	84.63%	
AUC (pessimistic)	pred Ada	445	87.68%	
class recall	84.63%	87.68%		

Gambar 10. Hasil perhitungan *precision* menggunakan algoritma *naïve bayes* dengan *rapid miner studio*



recall: 87.68% (+ 2.21% (info average 87.89%) (positive class Ada))				
	True Target Ada	True Ada	class precision	
AUC (optimistic)	347	63	84.63%	
AUC	pred True Ada	63	84.63%	
AUC (pessimistic)	pred Ada	445	87.68%	
class recall	84.63%	87.68%		

Gambar 11. Hasil perhitungan *recall* menggunakan algoritma *naïve bayes* dengan *rapid miner studio*

3.4 pengujian *confusion matrix*

Confusion matrix, juga dikenal sebagai tabel kesalahan, adalah alat evaluasi yang

umum digunakan dalam pemrosesan data dan pembelajaran mesin untuk menggambarkan performa model klasifikasi. Ini memberikan gambaran tentang seberapa baik model dapat memprediksi kelas target yang benar dan sejauh mana model tersebut mengalami kesalahan.[17]

Confusion matrix berbentuk matriks persegi yang terdiri dari empat entri utama: true positive (TP), true negative (TN), false positive (FP), dan false negative (FN). Mari kita jelaskan masing-masing entri ini:

1. *True Positive (TP)*: Ini adalah merupakan contoh *positif* serta benar diprediksi dengan benar oleh model. Artinya, model dengan benar mengidentifikasi contoh sebagai positif saat memang benar-benar positif.
2. *True Negative (TN)*: Ini adalah merupakan contoh *negatif* serta benar diprediksi dengan benar oleh model. Artinya, model dengan benar mengidentifikasi contoh sebagai negatif saat memang benar-benar negatif.
3. *False Positive (FP)*: Ini adalah jumlah contoh negatif dengan salah diprediksi sebagai *positif* serta model. Artinya, model secara keliru mengidentifikasi contoh sebagai positif saat seharusnya negatif. False positive juga dikenal sebagai kesalahan jenis I.
4. *False Negative (FN)*: Ini adalah jumlah contoh positif yang salah diprediksi sebagai negatif oleh model. Artinya, model secara keliru mengidentifikasi contoh sebagai negatif saat seharusnya positif. False negative juga dikenal sebagai kesalahan jenis II.

Dengan menggunakan entri-entri di atas, *confusion matrix* memberikan gambaran visual tentang kinerja model klasifikasi.

Berikut adalah contoh sederhana dari confusion matrix:

Dengan menggunakan *confusion matrix*, kita dapat menghitung beberapa metrik *evaluasi* model klasifikasi, antara lain:

Tabel 2. *matrix evaluasi model*

	Prediksi positif	Prediksi negatif
Positif	TP	FP
Negatif	FN	TN

1. **Akurasi (Accuracy):** Menunjukkan seberapa baik model melakukan prediksi secara keseluruhan dengan membagi jumlah prediksi yang benar (TP dan TN) dengan jumlah total contoh. Akurasi menggambarkan persentase keseluruhan prediksi yang benar.
2. **Presisi (Precision):** Menggambarkan seberapa baik model dalam mengidentifikasi contoh yang benar-benar positif. Presisi dihitung dengan membagi TP dengan jumlah prediksi positif (TP dan FP). Presisi menggambarkan seberapa jarang model memberikan hasil positif palsu.
3. **Recall (Sensitivity atau True Positive Rate):** Menunjukkan seberapa baik model dapat menemukan semua contoh *positif*. *Recall* dihitung dengan membagi TP dengan jumlah contoh *positif* sebenarnya (TP dan FN). *Recall* menggambarkan seberapa jarang model memberikan hasil negatif palsu.
4. **Parafrese: *F1-score*** menggabungkan presisi dan *recall* menjadi satu skor tunggal yang menggambarkan keseimbangan antara keduanya. *Skor F1-score* merupakan rata-rata harmonik dari presisi dan recall. Skor ini bermanfaat ketika terdapat ketidakseimbangan antara kelas positif dan negatif dalam dataset.

Confusion matrix sangat penting dalam memahami performa model klasifikasi karena memberikan informasi terperinci tentang prediksi yang benar dan kesalahan yang dibuat oleh model. Dengan memperhatikan metrik evaluasi yang dihasilkan dari *confusion matrix*, pengambilan keputusan yang tepat dapat dilakukan terkait penggunaan model tersebut dalam suatu konteks aplikasi.

3.5 Fase Permodelan

Metode yang di gunakan untuk melakukan komparasi adalah naïve bayes untuk memprediksi penyakit jantung, dalam hal ini yang di gunakan dataset yang terdiri dari 12 atribut, metode yang akan di uji merupakan metode sebelumnya dan metode naïve bayes dan hasil yang akan di sajikan adalah *accuracy dan auc*.

3.6 Analisis Hasil Komparasi

Model yang di peroleh dari *Naïve bayes* dan *logistik regresion* di uji dengan Cross Validation, terlihat bahwa algoritma naïve bayes memiliki accuracy yang tinggi

Tabel 3. hasil *cross validation*

	Algoritma naïve bayes	Algoritma logistik regresion
Accuracy	86.26%	83%
Auc	0.920	0.88

Hasil nilai dari accuracy algoritma naïve bayes memiliki nilai 86,26% dibandingkan dengan algoritma logistik yang hanya memiliki nilai 83% berdasarkan hasil di atas bahwa algoritma naïve bayes memiliki klasifikasi yang sangat baik karna memiliki nilai antara 0.920-0.025.

IV. Kesimpulan

Berdasarkan uji coba dan evaluasi yang telah di lakukan maka dapat diambil kesimpulan sebagai berikut:

1. Algoritma naïve bayes dalam bidang ilmu komputer data mining, khususnya dalam pengklasifikasian penyakit jantung, terbukti dapat diimplementasikan dengan baik. Pengujian yang dilakukan menggunakan metode manual dan menggunakan software RapidMiner menghasilkan akurasi sebagai ukuran tingkat keakuratan algoritma dalam mendiagnosa penyakit jantung.

2. algoritma naïve bayes memiliki kinerja yang lebih unggul dalam prediksi penyakit jantung.dibandingkan metode penelitian sebelumnya dengan menggunakan algoritma linear regresion

REFERENSI

- [1] Kemkes.go.id, 21 september 2022. [Online]. Available: <https://sehatnegeriku.kemkes.go.id/baca/umum/20210927/5638626/penyakit-jantung-koroner-didominasi-masyarakat-kota/>.
- [2] C. Aprianti, M. Faishal, and Y. Umaidah, "Penerapan Data Mining Untuk Klasifikasi Penjualan Baju Muslim Dimasa Pandemi Covid-19 Menggunakan Metode Algoritma C4.5 (Studi Kasus: Garaya Collection)," *J. Ilm. Wahana Pendidik.*, vol. 8, no. 1, pp. 104–112, 2021, doi: 10.5281/zenodo.5816231.
- [3] T. Dongoran, D. Septriani, and Y. J. Batkunde, "Perancangan Aplikasi Penjualan Dan Produksi Ud.Sehati Gas," *Ilm. Sist. Inf. dan Tek. Inform. "JISTI,"* vol. 3, pp. 48–54, 2020.
- [4] R. D. Safitri, A. Susanto, R. Rino, and L. W. Kusuma, "Rancang Bangun Aplikasi Absensi Sekolah Minggu Dengan Pengenalan Wajah Menggunakan Principal Component Analysis (Pca) Pada Gereja Gbi Modernland," *Algor*, vol. 2, no. 2, pp. 31–40, 2021, doi: 10.31253/algor.v2i2.567.
- [5] M. R. L. Iin Parlina, Agus Perdana Windarto, Anjar Wanto, "Memanfaatkan Algoritma K-Means Dalam Menentukan Pegawai Yang Layak Mengikuti Asessment Center," *Memanfaatkan Algoritm. K-Means Dalam Menentukan Pegawai Yang Layak Mengikuti Asessment Cent. Untuk Clust. Progr. Sdp*, vol. 3, no. 1, pp. 87–93, 2018.
- [6] A. Veronica Agustin and A. Voutama, "Implementasi Data Mining Klasifikasi Penyakit Diabetes Pada Perempuan Menggunakan Naïve Bayes," *JATI (Jurnal Mhs. Tek. Inform.,* vol. 7, no. 2, pp. 1002–1007, 2023, doi: 10.36040/jati.v7i2.6808.
- [7] J. Nasir, "PENERAPAN DATA MINING CLUSTERING DALAM MENGELOMPOKAN BUKU DENGAN METODE K-MEANS," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 11, no. 2, pp. 690–703, Oct. 2020, doi: 10.24176/simet.v11i2.5482.
- [8] A. Wijayadhi, M. Makmun Effendi, and S. Budi Rahardjo, "Prediksi Penyakit Jantung Dengan Algoritma Regresi Linier," *Bull. Inf. Technol.*, vol. 4, no. 1, pp. 15–28, 2023, doi: 10.47065/bit.v4i1.463.
- [9] Rayuwati, Husna Gemasih, and Irma Nizar, "IMPLEMENTASI ALGORITMA NAIVE BAYES UNTUK MEMREDIKSI TINGKAT PENYEBARAN COVID," *Jural Ris. Rumpun Ilmu Tek.*, vol. 1, no. 1, pp. 38–46, 2022, doi: 10.55606/jurritek.v1i1.127.
- [10] L. F. Tampubolon, A. Ginting, and F. E. Saragi Turnip, "Gambaran Faktor yang Mempengaruhi Kejadian Penyakit Jantung Koroner (PJK) di Pusat Jantung Terpadu (PJT)," *J. Ilm. Permas J. Ilm. STIKES Kendal*, vol. 13, no. 3, pp. 1043–1052, 2023, doi: 10.32583/pskm.v13i3.1077.
- [11] S. D. Lestari, I. S. Wahyuningsih, and A. I. Amal, "Gambaran aktivitas fisik dan tingkat kenyamanan pasien dengan penyakit jantung koroner," *J. Ilm. Sultan Agung*, vol. 2, no. 1, pp. 575–582, 2023, [Online]. Available: <https://jurnal.unissula.ac.id/index.php/JIMU/article/view/31327>.
- [12] A. Putranto, N. L. Azizah, and I. R. I. Astutik, "Sistem Prediksi Penyakit Jantung Berbasis Web Menggunakan Metode Svm Dan Framework," *J. Penerapan Sist. Inf. (Komputer Manajemen)*, vol. 4, no. 2, pp. 442–452, 2023, [Online]. Available: <https://archive.ics.uci.edu/ml/datasets/heart+disease>.
- [13] M. Hakimah and R. R. Muhimah, "Klasifikasi Penderita Penyakit Jantung Menggunakan Metode Naive Bayes dengan Chi-Square untuk Pemilihan Atribut," *Semin. Nas. Tek. Elektro, Sist. Inf. dan Tek. Inform.*, no. 1, pp. 257–262, 2021.
- [14] A. Maulana, F. Fauziah, and R. T. Komalasari, "Penerapan Algoritma Fisher-Yates Untuk Mengacak Soal Penerimaan Forum Studi Mahasiswa Informatika Universitas Nasional,"

- JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 5, no. 2, p. 104, 2020, doi: 10.29100/jipi.v5i2.1808.
- [15] F. “kaggle,” 2021. [Online]. Available: <https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction>.
- [16] D. Lestari, “Metode Naïve Bayes Dalam Machine Learning Untuk Memprediksi Penyakit Jantung Dalam Tubuh,” *J. Teknol. Komput. dan Sist. Informasi* Februari, vol. 2022, no. 1, pp. 23–28, 2022, [Online]. Available: <http://jurnal.goretanpena.com/index.php/teknisi>.

BIOGRAPHY

Adrian Timotius lahir di tangerang 05 januari tahun 2000, Menyelesaikan kuliah strata 1 (S1) Program Teknik Informatika pada tahun 2023 di Universitas Buddhi Dharma.

Indah Fenriana *received her Bachelor Degree in Information Technology (S.Kom) from Buddhi Dharma University, Indonesia and Master Degree in Computer Sciences (M.Kom) concentration in Software Engineering from Eresha University, Indonesia. She is a lecturer at the Departement of Information Technology, Faculty of Science and Technology, Buddhi Dharma University.*