
SISTEM KLASIFIKASI PENYAKIT JANTUNG BERBASIS PARTICLE SWARM OPTIMIZATION DAN NAÏVE BAYES DENGAN 5-FOLD CROSS VALIDATION

Oleh

Mufti Ari Bianto*¹, Kusrini²

¹Teknik Komputer, Universitas Muhammadiyah Lamongan

²Teknik Informatika, Universitas Amikom Yogyakarta

E-mail: ¹muftiari10@gmail.com, ²kusrini@amikom.ac.id

Article History:

Received: 21-12-2023

Revised: 29-12-2023

Accepted: 24-01-2024

Keywords:

Klasifikasi, Penyakit Jantung,
Naïve Bayes

Abstract: Serangan Jantung adalah salah satu penyakit yang paling mematikan tercatat di dunia, terdapat jumlah kasus baru Penyakit Jantung sebanyak 1,5% serta jumlah kematian sebanyak 14,38%. Pada tahun 2022 jumlah penderita Penyakit Jantung di Indonesia sejumlah 139.891 orang, pada umumnya jumlah penderita penyakit ini terus meningkat dikarenakan kurangnya pengetahuan atau informasi tentang penyakit jantung tersebut. Oleh karena itu dibutuhkan sebuah sistem yang dapat memberikan informasi serta klasifikasi penyakit secara dini yang dapat digunakan untuk klasifikasi apabila seseorang ingin mengetahui informasi ataupun gejala awal serangan jantung. Metode naïve bayes merupakan salah satu metode yang digunakan untuk melakukan klasifikasi berdasarkan probabilitas atau kemungkinan dari data sebelumnya, selain pendekatannya sederhana metode tersebut juga dapat melakukan klasifikasi secara baik. Mekanisme pengujiannya yaitu membagi 303 data kedalam 5 subset yang akan divalidasi dengan 5-fold cross validation. Hasil akhir dari penelitian ini adalah penerapan sistem klasifikasi dengan menggunakan metode naïve bayes yang akan menghasilkan nilai rata-rata akurasi sebesar 90,61%, presisi sebesar 87,44 %, dan recall sebesar 87,95%.

PENDAHULUAN

Penyakit Serangan jantung merupakan salah satu penyakit mematikan di dunia termasuk Indonesia [1]. Penyakit penyakit yang banyak penderitanya adalah Penyakit Serangan jantung. Di Indonesia, Penyakit Serangan jantung adalah kasus penyakit yang paling umum dijumpai pada orang dewasa [2]. Berdasarkan data kasus penyakit di Indonesia dari Global Burden of Disease (GBD) dan Organisasi Kesehatan Dunia (WHO) pada tahun 2019, terdapat jumlah kasus baru Penyakit Serangan jantung sebesar 1,5% dan persentase kematian sebesar 14,38%. Pada tahun 2022 jumlah penderita penyakit Penyakit Serangan jantung sebesar 139.891 orang, dengan penderita terbanyak berdasarkan provinsi, yaitu Kalimantan Utara dengan prevalensi kasus sebesar 2,2% kasus. Dari data diatas dapat

diketahui bahwa banyak orang yang belum menanggapi penyebab penyakit ini dengan serius, dan setelah melakukan pemeriksaan kesehatan, dokter mendeteksi adanya penyakit dengan stadium yang sudah tinggi [3]. Banyak alternatif cara untuk mencegah bahkan menyembuhkan penyakit, seperti operasi, penyinaran dan khemoterapi [4]. Namun, kurangnya akses informasi/media menjadi alasan penderita terlambat untuk memeriksakan diri ke dokter [5].

Terdapat hubungan antara kurangnya akses informasi/media dengan keterlambatan pemeriksaan awal Penyakit Serangan jantung [5]. Berdasarkan permasalahan diatas, dibutuhkan sebuah sistem yang dapat membantu dalam mendiagnosa penyakit Penyakit Serangan jantung, yaitu sistem klasifikasi dengan kemudahan menginput langsung gejala yang dirasakan lalu mendapatkan klasifikasi penyakit yang dialami [4]. Sistem klasifikasi membutuhkan metode yang tepat pula dalam mengelola pengetahuan yang diadopsi dari pakar sehingga diperoleh hasil yang akurat.

Klasifikasi adalah proses pengidentifikasian obyek ke dalam sebuah kategori, kelas atau kelompok berdasarkan prosedur, definisi dan karakteristik yang telah ditentukan sebelumnya [6]. Klasifikasi bertujuan untuk menempatkan objek yang ditugaskan hanya ke salah satu kategori yang disebut kelas, terdapat beberapa metode klasifikasi diantaranya Algoritma CART (Classification and Regreesion Trees), Algoritma mean vector, Algoritma k-nearest neighbour, Algoritma ID3, Algoritma C4 dan naïve bayes. Tak hanya klasifikasi, proses pengelompokan obyek juga dapat dilakukan dengan menggunakan teknik clustering.

Clustering merupakan pengelompokan obyek berdasarkan kemiripan antar obyek. Perbedaan antara klasifikasi dan clustering terletak pada proses pengelompokan obyek. Jika pada klasifikasi proses pengelompokan obyek dilakukan dengan membagi obyek berdasarkan kelompok / kategori yang telah didefinisikan sebelumnya, maka proses pengelompokan obyek pada clustering dilakukan dengan melihat kemiripan antar obyek, sehingga kategori belum terdefinisi sebelumnya. Salah satu metode klasifikasi yang sering digunakan adalah Naïve Bayes Classifier yang pertama kali dikemukakan oleh Revered Thomas Bayes [7]. Penggunaan Naïve Bayes Classifier sudah dikenalkan sejak tahun 1702-1761. Menurut Lewis, Hand dan Yu, Naïve Bayes Classifier merupakan pendekatan yang sangat sederhana dan sangat efektif untuk pelatihan klasifikasi. Sedangkan Kononenko dan Langley menyimpulkan bahwa Naïve Bayes Classifier merupakan kemungkinan label kelas data atau bisa diasumsikan sebagai atribut ke006Cas yang diberi label [7][8].

LANDASAN TEORI

Data Mining

Data mining adalah suatu metode pengolahan data untuk menemukan pola yang tersembunyi dari data tersebut. Hasil dari pengolahan data dengan metode data mining ini dapat digunakan untuk mengambil keputusan di masa depan. Data mining ini juga dikenal dengan istilah pattern recognition [9]. Data mining merupakan metode pengolahan data berskala besar oleh karena itu data mining ini memiliki peranan penting dalam bidang industri, keuangan, cuaca, ilmu dan teknologi. Secara umum kajian data mining membahas metode-metode seperti, clustering, klasifikasi, regresi, seleksi variable, dan market basket analisis [13].

Naive Bayes

Naive Bayes adalah sebuah Naive dengan dua penafsiran berbeda, dalam penafsiran Bayes, Naive ini menyatakan seberapa jauh derajat kepercayaan subjektif harus berubah secara rasional ketika ada petunjuk baru. Dalam penafsiran frekuentis, Naive ini menjelaskan representasi invers probabilitas dua kejadian [9]. Naive Bayes ditemukan oleh Reverend Thomas Bayes pada abad ke-18. Naive ini merupakan dasar dari statistika Bayes dan dapat diterapkan dalam banyak bidang seperti sains, rekayasa, ilmu ekonomi, teori permainan kedokteran, hukum, dan lain sebagainya [14].

Evidence Tunggal (E) dan hipotesis tunggal (H)

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

Keterangan :

$P(H|E)$ = probabilitas hipotesis H terjadi jika evidence E terjadi

$P(E|H)$ = probabilitas munculnya evidence E jika hipotesis H terjadi

$P(H)$ = probabilitas hipotesis H tanpa memandang evidence apapun

$P(E)$ = probabilitas evidence E tanpa memandang apapun

1. Evidence tunggal (E) dan hipotesis ganda (H_1, H_2, \dots, H_n)

$$P(H_i|E) = \frac{P(E|H) \times P(H)}{\sum_{k=1}^n P(E|H) \times P(H_k)}$$

Keterangan

$P(H_i|E)$ = probabilitas hipotesis H_i benar terjadi jika diberikan evidence E

$P(E_i|H)$ = probabilitas munculnya evidence E jika diketahui hipotesis H_i Benar

$P(E_i|H)$ = hipotesis H_i benar terjadi H_i (menurut hasil sebelumnya tanpa memandang evidence apapun

n = jumlah hipotesis yang mungkin

2. Evidence ganda dan hipotesis ganda

$$P(H_i|E_1 E_2 \dots E_m) = \frac{P(E_1|H_i) \times P(E_2|H_i) \times \dots \times P(E_m|H_i) \times P(H)}{\sum_{k=1}^n P(E_1 E_2 \dots E_m|H_k) \times P(H_k)}$$

Akan tetapi pengeplikasian tersebut tidak mungkin karena harus mengetahui semua probabilitas bersyarat dari semua kombinasi, maka persamaan tersebut diganti dengan persamaan :

$$P(H_i|E_1 E_2 \dots E_m) = \frac{P(E_1|H_i) \times P(E_2|H_i) \times \dots \times P(E_m|H_i) \times P(H)}{\sum_{k=1}^n P(E_1|H_k) \times P(E_2|H_k) \times \dots \times P(E_m|H_k) \times P(H)}$$

Partikel swam optimization

Particle Swarm Optimization (PSO) adalah metode optimasi global yang diperkenalkan oleh Kennedy dan Eberhart pada tahun 1995 berdasarkan penelitian terhadap perilaku kawanan burung dan ikan. Setiap partikel dalam Particle Swarm Optimization memiliki kecepatan partikel bergerak dalam ruang pencarian dengan kecepatan yang dinamis disesuaikan dengan perilaku historis mereka. Oleh karena itu, partikel memiliki kecenderungan untuk bergerak menuju daerah pencarian yang lebih baik selama proses pencarian [10].

Dalam algoritma PSO terdapat beberapa proses sebagai berikut:

1. Inisialisasi
 - a. Inisialisasi kecepatan awal
Pada iterasi ke-0, dapat dipastikan bahwa nilai kecepatan awal semua partikel adalah 0.
 - b. Inisialisasi posisi awal partikel
Pada iterasi ke-0, posisi awal partikel dibangkitkan dengan persamaan :
..... $x = x_{min} + rand[0,1] \times (x_{max} - x_{min})$
 - c. Inisialisasi pBest dan gBest
Pada iterasi ke-0, pBest akan disamakan dengan nilai posisi awal partikel. Sedangkan gBest dipilih dari satu pBest dengan fitness tertinggi.
2. Update kecepatan
Untuk melakukan update kecepatan, digunakan rumus berikut:
..... $vi,jt+1 = w.vi,jt + c1.r1(Pbesti,jt - xi,jt) + c2.r2(Gbestg,jt - xi,jt)$
3. Update posisi dan hitung fitness
Untuk melakukan update posisi, digunakan rumus berikut:
..... $xi,jt+1 = xi,jt + vi,jt+1$
 1. Precision : adalah tingkat ketepatan antara informasi yang diminta oleh pengguna dengan jawaban yang diberikan oleh sistem.
 2. Accuracy : didefinisikan sebagai tingkat kedekatan antara nilai prediksi dengan nilai aktual.

Cross-Validation

Pengujian dengan metode cross-validation adalah pengujian yang dilakukan dengan cara membagi data menjadi dua bagian, yaitu data training dan data test. Dalam K-Fold Cross-Validation, data awal secara acak dipartisi menjadi subset k yang terpisah dari yang lain atau "kelipatannya", D1, D2, ..., Dn, masing-masing memiliki ukuran yang sama. Data training dan data test dilakukan sebanyak n kali. Pada iterasi i, partisi Di diperuntukkan sebagai test set, dan partisi sisanya secara kolektif digunakan untuk training model [11]. Artinya dalam iterasi pertama subset-subset D2 sampai Dk secara kolektif digunakan sebagai training set untuk mendapatkan model pertama yang akan diuji pada subset di. Iterasi kedua dilatih pada himunan bagian D1, D3, ..., Dn dan diuji pada subset D2 serta seterusnya. Berbeda dengan metode Holdout And Random Subsampling, pada metode ini setiap sampel digunakan jumlah yang sama untuk setiap training data dan sekali untuk testing. Untuk rekomendasi, estimasi akurasi adalah jumlah keseluruhan rekomendasi yang benar dari iterasi-iterasi sebanyak n, dibagi dengan jumlah total tupel/record data awal [7].

METODE PENELITIAN

Dalam penelitian ini peneliti melakukan studi literatur tentang metode klasifikasi partikel swam optimization dan naïve bayes. Kedua metode ini merupakan mesin pembelajaran yang sering digunakan untuk mengklasifikasi atau meramalkan sesuatu. Data yang akan diuji dalam penelitian ini yaitu: Akurasi, presisi, recall dan waktu pada hasil penggabungan algoritma partikel swam optimization dan naïve bayes dalam klasifikasi penyakit serangan jantung.

Metode Pengumpulan Data

Penelitian ini menggunakan tahap eksperimen sebagai metode pengumpulan data dengan data sekunder yang melibatkan data laboratorium dan penelitian terhadap penderita penyakit jantung, namun selain itu peneliti juga mengumpulkan data awal sebagai bahan referensi melalui studi pustaka yang terkait langsung dengan bidangnya. Tahap eksperimen dilakukan dengan melakukan penggabungan model algoritma Partikel swam optimization dan Naive bayes sebagai metode yang akan digabungkan, kemudian hasil penggabungan kedua model algoritma tersebut di evaluasi berdasarkan tingkat Akurasi, Presisi, dan Recall untuk dilihat hasil perbandingannya [15].

Proses pengumpulan data dimulai dari pengambilan data pada UCI Machine Learning untuk mendapatkan parameter-parameter yang akan diteliti, dan selanjutnya akan dilanjutkan proses pengambilan data pada objek yang akan diteliti, dalam kasus ini peneliti mengambil data klasifikasi penyakit serangan jantung dari UCI Machine Learning yang akan digunakan sebagai dataset. Dataset yang diperoleh tersebut akan dijadikan model data untuk melakukan penggabungan pada kedua model algoritma yang akan diteliti.

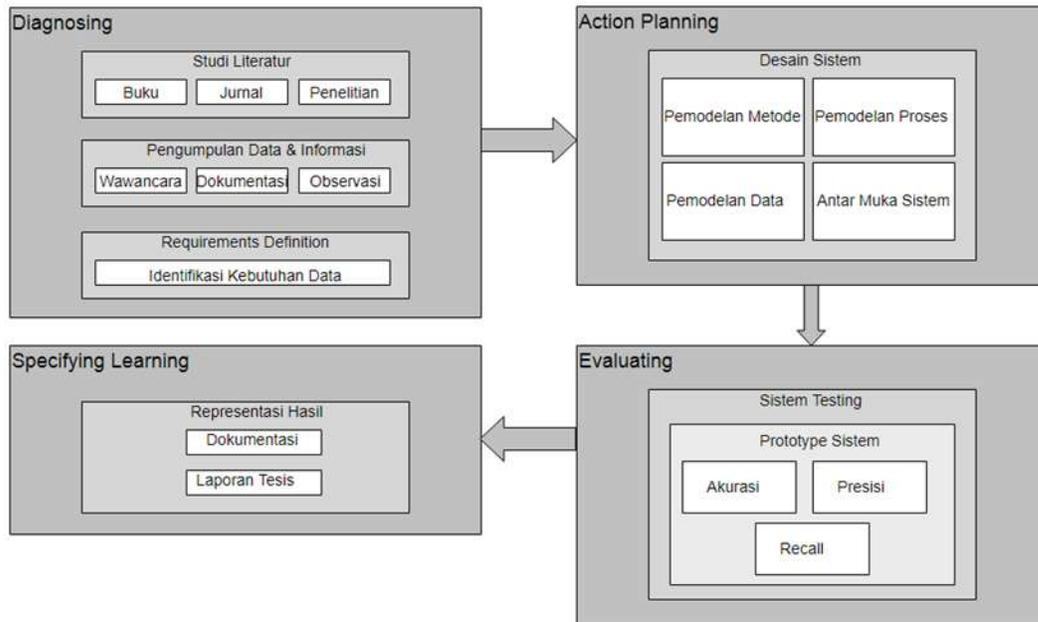
Metode Analisis Data

Metode analisis data pada penelitian ini adalah membandingkan hasil eksperimen mulai dari awal sampai akhir. Eksperimen dimulai dari melakukan cleaning atau pembersihan pada data yang telah dikumpulkan. Tahap selanjutnya akan dilakukan transformasi data, dimana pada tahap ini akan dilakukan pengclustering untuk dijadikan menjadi beberapa group atau kelompok data. Setelah data berhasil dikelompokkan maka proses selanjutnya akan dilakukan proses pemodelan data dimana data yang sudah dikelompokkan tersebut akan diubah menjadi nilai-nilai yang dipisahkan dengan tanda koma atau Comma Separated Value (CSV) sebagai format data yang akan dimasukkan ke dalam database.

Setelah proses pemodelan data atau memasukan data ke dalam database selesai dilakukan, maka proses selanjutnya adalah melakukan penggabungan dua model algoritma yaitu Partikel swam optimization dan Naive bayes menggunakan data yang telah diperoleh. Kedua model algoritma ini akan digabungkan dengan menggunakan 5-FOLD Cross-Validation [16]. Metode ini membagi secara acak 5 subset dan kemudian setiap subset $n=1$ akan menjadi data training sedangkan $n=2,3,\dots,5$ akan menjadi data testing. Kemudian dilakukan sebanyak 5 iterasi dengan data testing $n+1$ dan syarat data testing tidak sama dengan data training. Penarikan kesimpulan dilakukan berdasarkan hasil evaluasi pada kedua model algoritma yang terpilih dengan menggunakan pengukuran nilai nilai performa Akurasi, presisi, recall dan waktu. Hasil pengukuran tersebut akan dijadikan acuan atau pedoman dalam menentukan hasil atau klasifikasi pada penelitian ini.

Alur Penelitian

Pada bagian ini berisi diagram alur langkah penelitian secara lengkap dan terinci termasuk di dalamnya mencakup algoritma, rute, pemodelan-pemodelan, desain yang terkait dengan perancangan sistem. Alur penelitian pada sistem prediksi klasifikasi penyakit serangan jantung ditunjukkan pada Gambar 1 di bawah ini [12].



Gambar 1. Metode Penelitian

HASIL DAN PEMBAHASAN

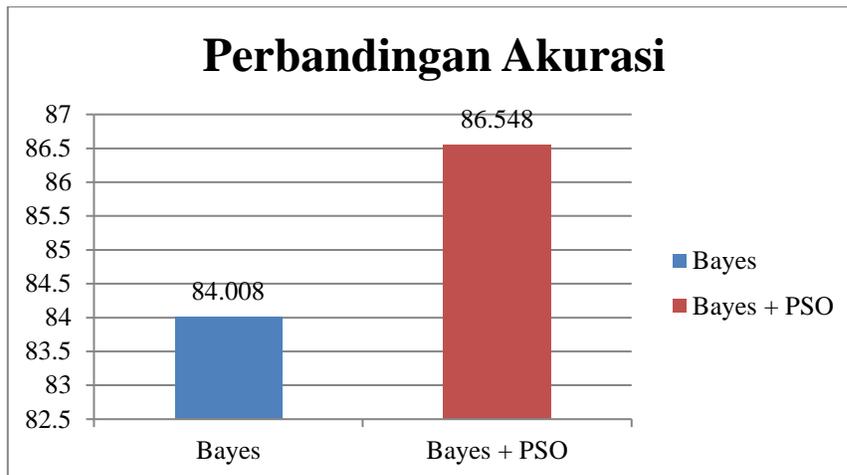
Hasil analisis perbandingan hasil pengujian melalui konfigurasi 303 data dan 14 parameter adalah sebagai berikut:

Pada bagian ini hasil percobaan dari implementasi Naive Bayes serta Naive Bayes dan Particle Swam Optimization akan dilakukan perbandingan. Hasil yang akan dibandingkan merupakan rata-rata terbaik dari keseluruhan percobaan pengujian dengan parameter Akurasi, Presisi, Recall dan waktu training yang telah dilakukan pada sub bab sebelumnya. Hasil nilai rata-rata performa pada percobaan terbaik dengan menggunakan dua algoritma dapat dilihat pada Tabel 2.

Tabel 2 . Hasil Perbandingan Tingkat Performa 14 Parameter

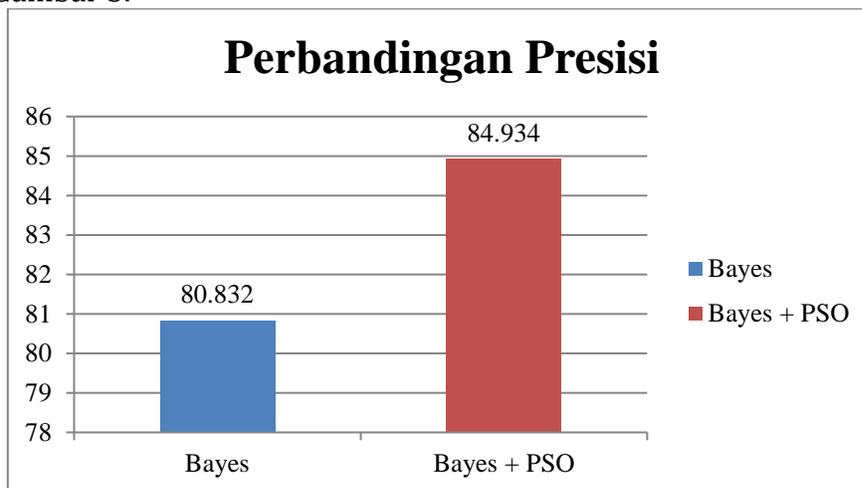
No	Algoritma	Akurasi (%)	Presisi(%)	Recall (%)	Waktu (Detik)
1	Naive Bayes	84.008	80.832	80.128	9,00
2	Naive Bayes dan PSO	86.548	84.934	84.688	40,60
Selisih		2.540	4.102	4,754	31.60

Berdasarkan hasil pada Tabel 2 dapat ditarik suatu kesimpulan bahwa nilai performa unuk akurasi model algoritma *Naive Bayes dan Particle Swam Optimization* memiliki tingkat akurasi yang lebih baik dengan nilai akurasi sebesar 85% dengan selisih 3,84% jika dibandingkan dengan model algoritma *Naive Bayes* dengan nilai yang lebih rendah sebesar 81,20% pada konfigurasi dataset sebanyak 303 data. Adapun grafik dari perbandingan nilai akurasi pada Gambar 2.



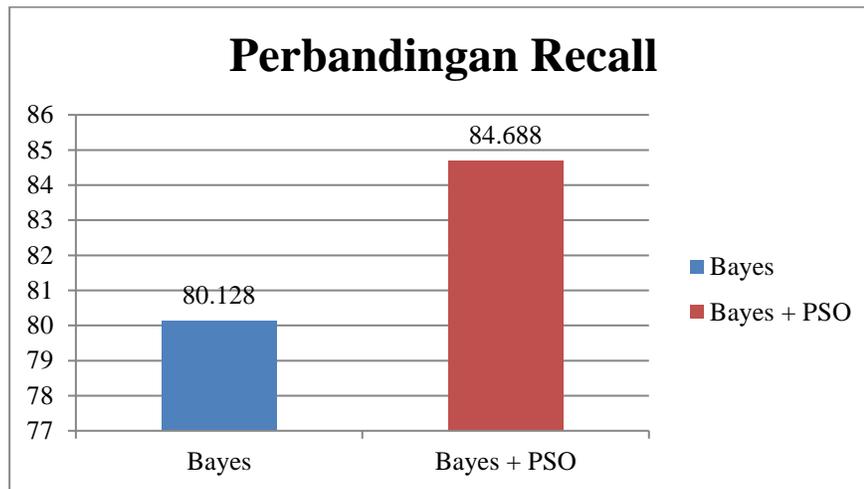
Gambar 2. Perbandingan Performa Akurasi 14 Parameter

Kemudian untuk nilai tingkat performa presisi *Naïve Bayes* dan *Particle Swam Optimization* memiliki lebih baik dengan nilai presisi sebesar 83,60% dengan selisih 4,20% jika dibandingkan dengan model algoritma *Naive Bayes* dengan nilai yang lebih rendah sebesar 79,40% pada konfigurasi dataset sebanyak 303 data. Adapun grafik dari perbandingan nilai presisi pada Gambar 3.



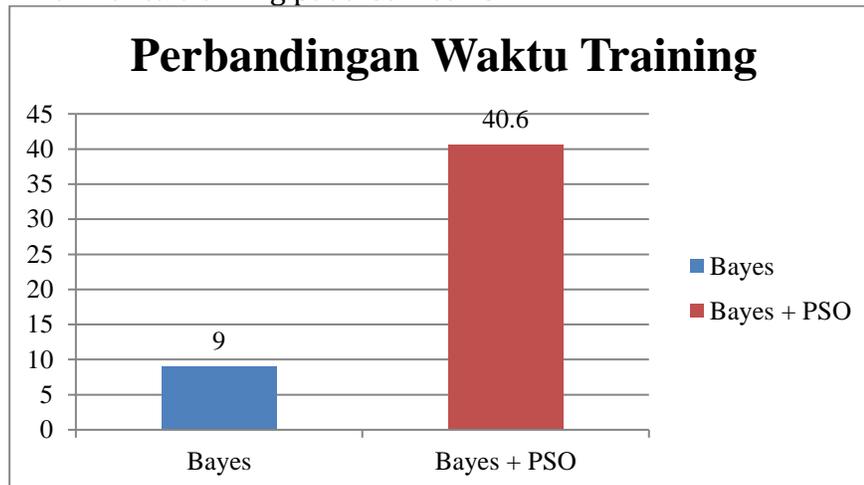
Gambar 3. Perbandingan Performa Presisi 14 Parameter

Kemudian untuk nilai tingkat performa recall *Naïve Bayes* dan *Particle Swam Optimization* memiliki lebih baik dengan nilai presisi sebesar 84% dengan selisih 4,40% jika dibandingkan dengan model algoritma *Naive Bayes* dengan nilai yang lebih rendah sebesar 79,60% pada konfigurasi dataset sebanyak 303 data. Adapun grafik dari perbandingan nilai recall pada Gambar 4.



Gambar 4. Perbandingan Performa Recall 14 Parameter

Kemudian untuk nilai tingkat performa waktu training *Naïve Bayes* memiliki lebih baik dengan nilai waktu sebesar 9,00 detik dengan selisih 31,60 detik jika dibandingkan dengan model algoritma *Naïve Bayes dan Particle Swam Optimization* dengan nilai yang lebih rendah sebesar 40,60 detik pada konfigurasi dataset sebanyak 303 data. Adapun grafik dari perbandingan nilai waktu training pada Gambar 5.



Gambar 5. Perbandingan Performa Waktu Training 14 Parameter

KESIMPULAN

Setelah melalui tahap pengujian tingkat performa dengan parameter akurasi, presisi, recall dan Waktu Training pada model algoritma *Naive Bayes* serta *Naive Bayes dan Partikel Swam Optimization* untuk prediksi mahasiswa pada diagnose penyakit serangan jantung, maka dapat diambil beberapa kesimpulan antara lain:

1. Pada konfigurasi 303 data dan 14 parameter model algoritma *Naive Bayes* memiliki nilai tingkat performa akurasi dengan nilai 84.008%, performa presisi dengan nilai 80.832%, performa recall dengan nilai 80.128% serta performa waktu training dengan nilai 9,00 detik. Sedangkan pada model algoritma *Naive Bayes dan Partikel Swam Optimazation* memiliki nilai tingkat performa akurasi dengan nilai 86.548%, performa presisi dengan

- nilai 84,934%, performa recall dengan nilai 84.688% serta performa waktu training dengan nilai 40.60 detik.
2. Pengaruh jumlah records terhadap model algoritma yang digunakan memiliki pengaruh baik terhadap nilai akurasi, presisi dan recall dimana nilai akurasi, presisi dan recall akan meningkat jika jumlah records semakin bertambah, akan tetapi berdampak buruk pada kecepatan waktu training dimana waktu training akan melambat jika jumlah records semakin bertambah.
 3. Pengaruh jumlah parameter terhadap model algoritma yang digunakan memiliki pengaruh baik terhadap nilai akurasi, presisi dan recall dimana nilai akurasi, presisi dan recall akan meningkat jika jumlah records semakin bertambah, akan tetapi tidak berdampak pada kecepatan waktu training.

PENGAKUAN/ACKNOWLEDGEMENTS

Dengan penuh rasa syukur, kami mengucapkan terima kasih kepada Universitas Muhammadiyah Lamongan atas dukungan fasilitas dan akademiknya. Penghargaan yang tulus juga kami sampaikan kepada Profesor Kusri atas bimbingan dan ilmunya yang berharga.

Kami juga berterima kasih kepada semua pihak yang telah berkontribusi, baik secara langsung maupun tidak langsung, dalam mendukung keberhasilan penelitian ini. Semoga segala bantuan yang diberikan mendapat balasan terbaik.

DAFTAR PUSTAKA

- [1] Pradana, Musthofa Galih, Pujo Hari Saputro, and Dhina Puspasari Wijaya. "Komparasi Metode Support Vector Machine Dan Naïve Bayes Dalam Klasifikasi Peluang Penyakit Serangan Jantung." Indonesian Journal of Business Intelligence (IJUBI) 5.2 (2022): 87-91.
- [2] Amriani, Eka Astri, Nurhikmawati Nurhikmawati, and Fadillah Maricar. "Karakteristik infark miokard akut pada usia muda." PREPOTIF: JURNAL KESEHATAN MASYARAKAT 8.2 (2024): 2996-3005.
- [3] Fauziah, Wardah, et al. "Pemeriksaan Kesehatan Jantung Dalam Rangka Pengembangan Aplikasi Screening Jantung Berbasis Android." Budimas: Jurnal Pengabdian Masyarakat 6.3 (2024).
- [4] Nurvan, Hadi, Arum Kesuma Wardani, and Nur Endah Palupi. "Karakteristik Pemeriksaan Pasien Di Instalasi Radiologi Rumah Sakit Ananda Babelan Bekasi Periode Agustus 2021–Juli 2022: Studi Retrospektif." Jurnal Pandu Husada 4.4 (2023): 1-14.
- [5] Prahmawati, Pira. "PENYULUHAN KESEHATAN TENTANG DETEKSI DINI KEGAWATAN DIABETES DAN JANTUNG." Bagimu Negeri: Jurnal Pengabdian Kepada Masyarakat 8.1 (2024): 23-29.
- [6] Abdullah, Said Noor, et al. "Klasifikasi Tingkat Keparahan Korban Kecelakaan Lalu Lintas Menggunakan Naïve Bayes." Progresif: Jurnal Ilmiah Komputer 19.2 (2023): 858-866.
- [7] Turban Efraim, 2005, Decision Support System and Intelligent Systems, edisi Bahasa Indonesia jilid 1, Yogyakarta :Penerbit Andi.
- [8] Han, Jiawei, Jian Pei, and Hanghang Tong. Data mining: concepts and techniques.

- Morgan kaufmann, 2022.
- [9] Kusumadewi, Sri. 2003. Artificial Intelligence Teknik dan Aplikasi. Yogyakarta :Grahallmu
- [10] Suyanto. (2010). Algoritma Optimasi (Deterministik atau Probabilitas). Indonesia: Graha Ilmu.
- [11] Tuntun, Ritham, Kusnawi Kusnawati, and Kusnawi Kusnawati. "Analisis Perbandingan Kinerja Algoritma Klasifikasi dengan Menggunakan Metode K-Fold Cross Validation." Jurnal Media Informatika Budidarma 6.4 (2022): 2111-2119.
- [12] Mubarog, Ilham, Arief Setyanto, and Heri Sismoro. "Sistem klasifikasi pada penyakit breast cancer dengan menggunakan metode Naïve Bayes." Creative Information Technology Journal 6.2 (2021): 109-118.
- [13] Alfatah, Dhika. "Application of the K-Means Clustering Algorithm in Mapping the Regional Voter Strategy for the Legislative Candidates for the DPR RI." Jurnal Komputer, Informasi dan Teknologi 1.2 (2021): 435-443.
- [14] Rosid, Danu Fajar Nur, and A. Sidiq Purnomo. "Sistem Pakar Diagnosa Penyakit Pada Tanaman Kedelai Hijau Menggunakan Metode Teorema Bayes." IKRA-ITH Informatika: Jurnal Komputer dan Informatika 8.2 (2024): 72-81.
- [15] Yusman, Yanti. "Naïve Bayes Classifier dengan Particle Optimize Weight Forward pada Dataset." Bulletin of Computer Science Research 3.6 (2023): 442-447.
- [16] Heristian, Sujiliani, Rian Septian Anwar, and Hanggoro Aji Al Kautsar. "Klasifikasi Perilaku Pemain Game Online Menggunakan Naïve Bayes Berbasis Particle Swarm Optimization." Computer Science (CO-SCIENCE) 4.2 (2024): 151-156.