

Perancangan Sistem Klasifikasi Penyakit Jantung Menggunakan Naïve Bayes

Designing a Heart Disease Classification System Using Naïve Bayes

Mufti Ari Bianto*¹, Kusri², Sudarmawan³

^{1,2,3}Magister Teknik Informatika, Universitas Amikom Yogyakarta

E-mail: *¹mufti.b@astudents.amikom.ac.id, ²kusri@amikom.ac.id,

³sudarmawan@amikom.ac.id

Abstrak

Serangan Jantung adalah salah satu penyakit yang paling mematikan tercatat di dunia, terdapat jumlah kasus baru Penyakit Jantung sebanyak 43,32% serta jumlah kematian sebanyak 12,91%. Pada tahun 2013 jumlah penderita Penyakit Jantung di Indonesia sejumlah 61.682 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%.

Kata Kunci — klasifikasi, penyakit jantung, naïve bayes Classifier

Abstract

Heart attack is one of the most deadly diseases recorded in the world, there are a number of new cases of heart disease as much as 43.32% and the number of deaths as much as 12.91%. In 2013 the number of sufferers of heart disease in Indonesia amounted to 61,682 people, in general the number of sufferers of this disease continues to increase due to lack of knowledge or information about heart disease, therefore we need a system that can provide information and classification of diseases early that can be used for classification if someone wants to find out information or early symptoms of a heart attack. Naïve Bayes method is one of the methods used to classify based on the probability or likelihood of previous data, in addition to a simple approach the method can also do a good classification. The testing mechanism is to divide 303 data into 5 subsets that will be validated by 5-fold cross validation. The final result of this study is the application of the classification system using the Naïve Bayes method which will produce an average accuracy value of 90.61%, a precision of 87.44%, and a recall of 87.95%.

Keywords — classification, heart disease, naïve bayes

1. PENDAHULUAN

1.1. Latar Belakang

Serangan Jantung adalah salah satu penyakit yang paling mematikan di dunia [1] dan salah satu penyakit yang banyak penderitanya adalah Penyakit Jantung. Di Indonesia, Penyakit Jantung adalah kasus penyakit yang paling sering dijumpai pada perempuan dewasa [2]. Berdasarkan data kasus penyakit di Indonesia dari GLOBOCAN (IARC) pada tahun 2012, terdapat jumlah kasus baru serangan Jantung sebanyak 43,30% dan jumlah kematian sebesar 12,90%. Pada tahun 2013 jumlah penderita penyakit Penyakit Jantung sebesar 61.682 orang, dengan penderita terbanyak berdasarkan provinsi, yaitu Jawa Tengah dengan jumlah kasus sebesar 11.511 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-penyakit tersebut, seperti operasi, penyinaran dan khemoterapi [2]. Namun, kurangnya akses informasi/media menjadi alasan penderita terlambat untuk memeriksakan diri ke dokter [1].

Terdapat hubungan antara kurangnya akses informasi/media dengan keterlambatan pemeriksaan awal Penyakit Jantung [1]. Kurangnya akses untuk mencari informasi tentang penyakit serangan jantung ini menyebabkan peningkatan angka kematian setiap tahunnya. Karena itu, dibutuhkan sebuah sistem klasifikasi yang dapat memberikan informasi tentang penyakit serangan jantung serta dapat melakukan pengecekan klasifikasi secara dini tentang penyakit serangan jantung yang dialami oleh seseorang [3]. Untuk melakukan sebuah klasifikasi sistem membutuhkan metode yang tepat dalam mengelola pengetahuan yang diadopsi dari pakar sehingga diperoleh hasil yang akurat. Salah satu metode yang dapat digunakan dalam penelitian ini adalah penerapan Naïve Bayes Classifier, dimana metode Naïve Bayes Classifier merupakan suatu pendekatan yang cukup sederhana dan baik dalam melakukan pelatihan data untuk klasifikasi [4]. Sedangkan Naïve Bayes Classifier juga dapat dikatakan menghitung kemungkinan kelas atau kategori klasifikasi data atau dapat dijelaskan seperti atribut kelas atau kategori yang diberi label [5].

Beberapa penelitian yang menerapkan Naïve Bayes Classifier sudah banyak diteliti. Seperti pada penelitian yang diimplementasikan oleh Muhamad, yang mengemukakan Naïve Bayes mempunyai banyak keunggulan, diantaranya memiliki akurasi yang tinggi dengan bentuk algoritma yang sederhana namun cepat menyelesaikan masalah, namun validasi pengujian yang dilakukan dalam penelitian tersebut hanya satu kali. Mekanisme pengujian dilakukan dengan menggunakan 123 dataset dengan parameter akurasi saja, peneliti akan melakukan pengujian dengan mekanisme membagi 303 data kedalam 5 subset yang akan divalidasi dengan 5-fold cross validation dengan parameter pengujian akurasi, presisi dan recall [6]. Penelitian selanjutnya adalah percobaan yang diimplementasikan oleh Li dkk yang menjelaskan bahwa algoritma naïve bayes dapat digunakan untuk mengelola dataset dengan jumlah besar, algoritma naïve bayes mampu menghasilkan akurasi lebih dari 78%, hasil tersebut didapatkan melalui seleksi fitur terlebih dahulu sebelum melakukan klasifikasi sehingga dapat meningkatkan performa dari algoritma naïve bayes [7]. Sedangkan pada penelitian yang dilakukan nuraini lebih baik diimplementasikan pada data dengan jumlah yang besar serta dapat melakukan penganganan data yang bersifat kurang lengkap dan lumayan baik pada suatu atribut yang memiliki noise dan tidak relevannya pada sebuah data [8].

Berdasarkan latar belakang yang telah dijelaskan sebelumnya maka pada penelitian ini akan dilakukan penelitian pembuatan sistem klasifikasi penyakit jantung menggunakan Naïve Bayes Classifier. Dengan tujuan memberikan kontribusi penelitian melalui konfigurasi penerapan model algoritma dengan mekanisme validasi pengujian yang berbeda dari penelitian sebelumnya untuk hasil tingkat akurasi, presisi, dan recall yang lebih akurat.

1.2. Landasan teori

1.2.1. Naïve Bayes

Naive Bayes merupakan sebuah metode dengan pendekatan 2 penafsiran yang sedikit berbeda. Menurut penjelasan Bayes, model ini menjelaskan seberapa jauh tingkat derajat kepercayaan pada suatu subjektif harus dapat merubah secara rasional ketika adanya suatu petunjuk atau tujuan baru. Sedangkan penjelasan frekuentis, model ini mengartikan bahwa representasi sebagai invers probabilitas melalui dua scenario kejadian [9].

Naive merupakan dasar dari statistika Bayes dan dapat diterapkan dalam banyak bidang seperti sains, rekayasa, ilmu ekonomi, teori permainan kedokteran, hukum, dan lain sebagainya.

1. Menentukan Evidence Tunggal (E) dan hipotesa tunggal (H) seperti pada persamaan 1

$$P(H|E) = \frac{P(E|H) \times P(H)}{P(E)} \quad (1)$$

Penjelasan:

$P(H|.E)$, Kemungkinan hipotesaH terjadi jika evidence E terpenuhi

$P(E|.H)$, Kemungkinan munculnya evidence E jika hipotesis H terjadi

$P(H.)$, KemungkinanhipotesaH tanpa melihat evidence apapun

$P(E.)$, Kemungkinan evidence E tanpa melihatapapun

2. Mencari Evidence tunggal (E) dan hipotesa ganda (H_1, H_2, \dots, H_n) seperti pada persamaan 2

$$P(H_i|E) = \frac{P(E|H_i) \times P(H_i)}{\sum_{k=1}^n P(E|H_k) \times P(H_k)} \quad (2)$$

Penjelasan

$P(H_i|E)$, Kemungkinan hipotesa H_i benar terjadi jika diberikan evidence E

$P(E_i|.H)$, Kemungkinan munculnya evidence E jika diketahui hipotesa H_i Benar

$P(E_i|.H)$, Kemungkinan H_i benar terjadi H_i

N, jumlah kemungkinan yang akan terjadi

3. Mencari Evidence ganda dan tunggal ganda seperti pada persamaan 3

$$P(H_i|E_1 E_2 \dots E_m) = \frac{P(E_1|H_i) \times P(E_2|H_i) \times \dots \times P(E_m|H_i) \times P(H)}{\sum_{k=1}^n P(E_1 E_2 \dots E_m|H_k) \times P(H_k)} \quad (3)$$

Namun pada pengaplikasian yang dilakukan tidak akan mungkin terjadi dikarenakan harus memiliki pengetahuan semua kemungkinan bersyarat melalui semua kemungkinan yang dikombinasikan, maka rumus perhitungan tersebut diubah menjadi seperti persamaan 4.

$$P(H_i|E_1 E_2 \dots E_m) = \frac{P(E_1|H_i) \times P(E_2|H_i) \times \dots \times P(E_m|H_i) \times P(H)}{\sum_{k=1}^n P(E_1|H_k) \times P(E_2|H_k) \times \dots \times P(E_m|H_k) \times P(H)} \quad (4)$$

1.2.2. Klasifikasi

Klasifikasi merupakan metode untuk menentukan sebuah anggota kedalam suatu kelas tertentu yang telah ditentukan sebelumnya. Anggota tersebut dimasukan kedalam kelas tertentu berdasarkan persamaan karakter dari data tersebut. Teknik klasifikasi banyak digunakan dalam penerapan sistem klasifikasi untuk kasus tertentu. Dalam klasifikasi pembagian dataset dibagi menjadi 2 yaitu data latih dan data uji, dimana semua dataset akan dibagi kedalam dua cluster untuk dilakukan pelatihan dan pengujian terhadap data tersebut. Pembagian data tersebut juga akan menentukan hasil akurasi dari penerapan suatu metode dalam Teknik klasifikasi tersebut [10].

2. METODE PENELITIAN

Pada tahap ini peneliti akan melakukan studi literatur tentang metode klasifikasi naïve bayes. Penerapan metode ini merupakan mesin pembelajaran yang sering digunakan untuk mengklasifikasi atau meramalkan sesuatu. Parameter yang digunakan dalam pengujian penelitian ini yaitu: Akurasi, presisi, recall dan waktu, yang dihasilkan dari penerapan naïve bayes dalam klasifikasi penyakit serangan jantung.

2.1. Metode Penelitian

Metode penelitian yang akan digunakan dalam penelitian ini adalah metode penelitian kuantitatif, dimana peneliti akan melakukan penelitian secara tersistematis, terstruktur, dan dijelaskan secara rinci tentang perancangan sebuah sistem klasifikasi dengan menggunakan metode naïve bayes untuk klasifikasi penyakit jantung.

Penelitian dilakukan dengan cara pengumpulan dataset dari dataset pada UCI-Machine-Learning dengan format Comma Separated Value (CSV) yang berisikan 303 record dan terbagi menjadi 2 kelas serta 15 atribut, yaitu age, trestbps, chol, sex, cpfbsthalach, exang, oldpeak, slope, c, thal, restecg, num. Hasil akhir penelitian adalah Analisa hasil tingkat akurasi, presisi dan recall yang divalidasi dengan menggunakan 5-FOLD Cross-Validation yang akan dianalisa kedalam suatu bentuk angka, tabel dan grafik diagram untuk menjelaskan kesimpulan penelitian yang diperoleh.

2.2. Metode Pengumpulan Data

Pada tahap pengumpulan data menggunakan tahap eksperimen, namun selain itu peneliti juga mengumpulkan data awal sebagai bahan referensi melalui observasi, wawancara, dan studi pustaka kepada praktisi yang terkait langsung dengan bidangnya. Tahap eksperimen dilakukan dengan menerapkan Naive bayes sebagai metode yang akan diimplementasikan. Kemudian hasil penggabungan kedua model algoritma tersebut dievaluasi berdasarkan tingkat Akurasi, Presisi, dan Recall untuk dilihat hasil perbandingannya.

Proses pengumpulan data dimulai dari wawancara kepada pihak terkait untuk mendapatkan parameter-parameter yang akan diteliti, dan dilanjutkan dengan proses pengambilan data pada objek yang akan diteliti. Dalam kasus ini peneliti mengambil data klasifikasi penyakit jantung dari UCI-Machine-Learning yang akan diimplementasikan sebagai dataset. Dataset yang diperoleh tersebut akan dijadikan model data untuk melakukan penggabungan pada kedua model algoritma yang akan diteliti.

2.3. Metode Analisis Data

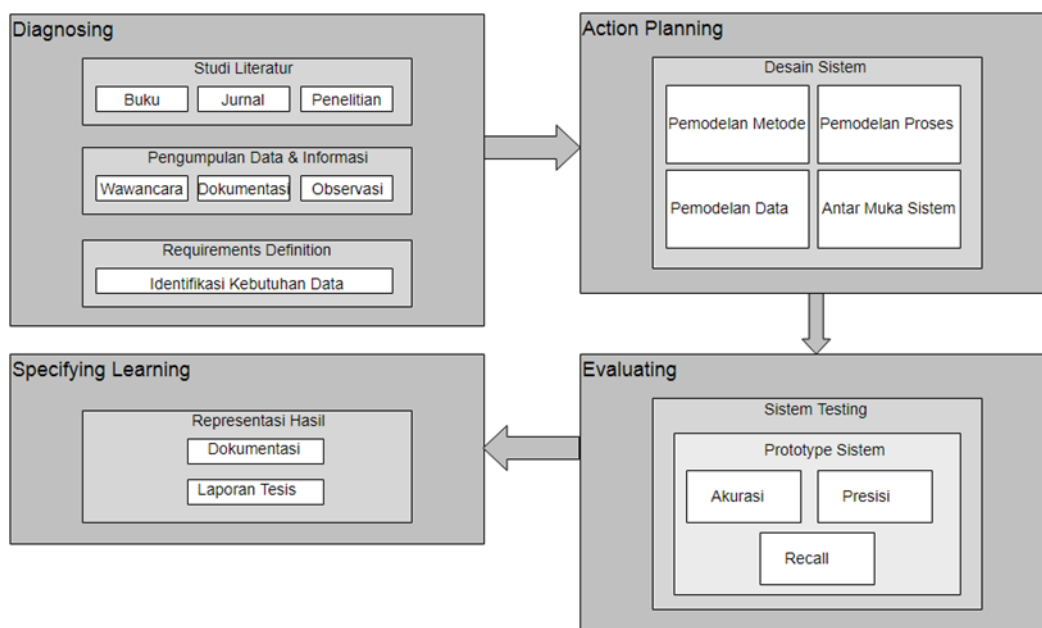
Tahap analisis data yang akan dilakukan adalah membandingkan hasil eksperimen mulai dari awal sampai akhir. Eksperimen dimulai dari melakukan cleaning atau pembersihan pada data yang telah dikumpulkan. Tahap selanjutnya akan dilakukan transformasi data, dimana pada tahap ini akan dilakukan pengclustering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

Setelah proses pemodelan data atau memasukan data di suatu database selesai dilakukan, maka proses yang dilakukan selanjutnya yaitu menerapkan model algoritma Naive bayes menggunakan data yang telah diperoleh. Kedua model algoritma ini akan digabungkan dengan menggunakan 5-FOLD Cross-Validation. 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.

2.4. Alur Penelitian

Alur pada tahap ini berisi diagram proses urutan alur penelitian secara rinci dan detail yang dapat mencakup algoritma, rute, pemodelan-pemodelan, desain yang terdapat pada perancangan sistem. Alur penelitian pada sistem prediksi klasifikasi penyakit jantung ditunjukkan di Gambar 1 dan Gambar 2.



Gambar 1. Metode Penelitian

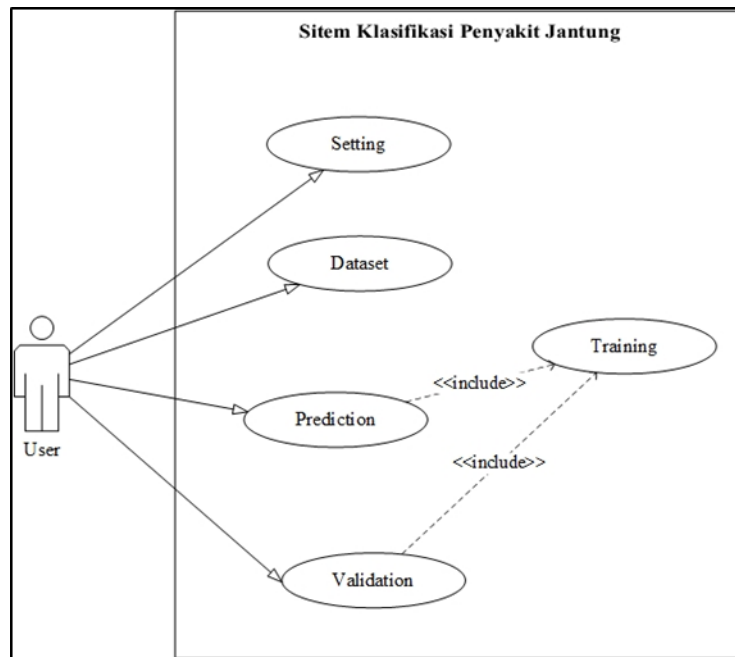
Penjelasan singkat alur atau metode penelitian seperti berikut:

1. *Diagnosing*
Pada tahapan ini peneliti melakukan sebuah studi literatur melalui proses membaca buku, jurnal, makalah serta laporan yang terkait dengan topik penelitian. Kemudian melakukan pengumpulan data dan informasi seperti melakukan wawancara, dokumentasi dan observasi. Selanjutnya dilakukan proses definisi kebutuhan dengan melakukan identifikasi data yang dibutuhkan, melihat prosedur yang sedang berjalan, menganalisis sistem yang sedang berjalan serta membuat hasil evaluasi sistem tersebut.
2. *Action Planning*
Pada tahap ini berisikan proses desain sistem dan *software* seperti pemodelan metode *Naive bayes*, pemodelan proses dengan *Unified Modeling Language (UML)* seperti *usecase diagram* dan *activity diagram* serta perancangan *User Interface*.
3. *Evaluating*
Pada tahap ini dilakukan pengimplementasian untuk didapatkan hasil penelitian dengan pengukuran performa Akurasi, presisi, recall dan waktu berupa nilai angka, tabel dan diagram.
4. *Specifying Learning*
Pada tahap ini dilakukan proses dokumentasi dan publikasi thesis berisi hasil penelitian yang sudah diterapkan.

3. HASIL DAN PEMBAHASAN

3.1. Perancangan Sistem

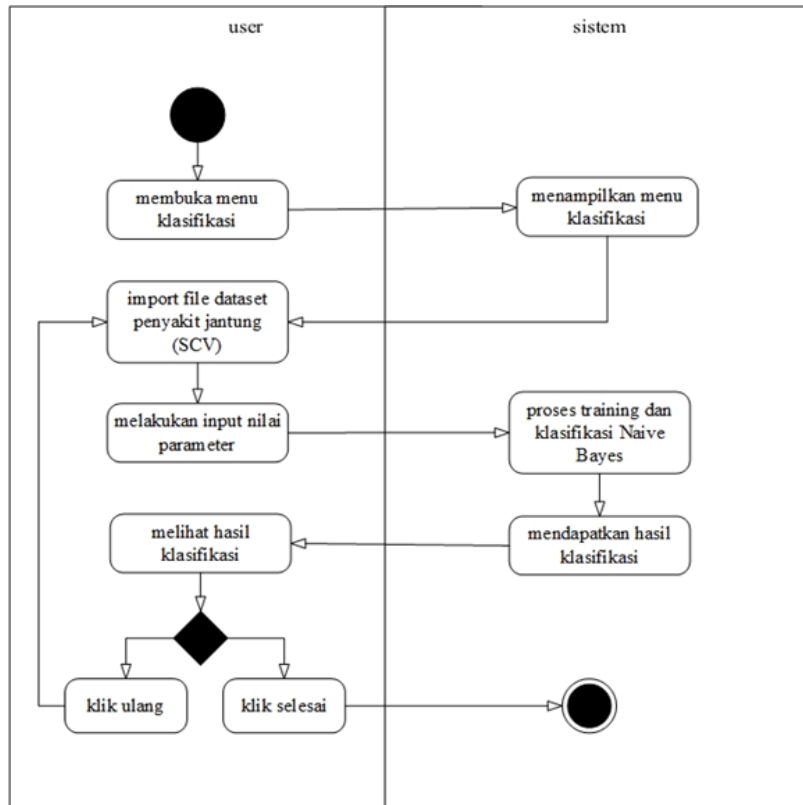
Pada proses ini akan dibuat suatu perancangan dari sistem klasifikasi penyakit serangan jantung melalui penerapan naïve bayes dengan menggunakan usecase pada Gambar 2 sebagai penjelasan kebutuhan fungsional, dan activity diagram pada Gambar 3 sebagai penggambaran alur atau aktifitas pada sistem.



Gambar 2. Usecase Diagram

Pada Gambar 2 menerangkan tentang fungsionalitas yang akan dibuat dan dirancang dengan pembuatan 5 menu yang diimplementasikan kedalam sistem, dimana terdapat menu pengaturan (setting) yang berfungsi untuk mengatur pembatasan jumlah dataset yang akan diolah dan dilakukan training. Pada system tersebut terdapat menu yang berfungsi untuk mengelola dataset seperti menambahkan dan menghapus dataset, menu prediction untuk melakukan prediksi atau klasifikasi penyakit jantung dan validation untuk melakukan validasi hasil berupa akurasi, presisi, dan recall. Akan tetapi menu prediction dan validation hanya dapat dilakukan ketika proses pada menu training dijalankan, menu training sendiri bertujuan untuk dapat menemukan bobot yang akan digunakan dalam proses klasifikasi dan validasi.

Pada Gambar 3 menerangkan alur klasifikasi penyakit jantung yang mana pengguna membuka dan mengklik menu klasifikasi dan sistem akan menampilkan menu klasifikasi, selanjutnya user harus mengimport dataset dengan format CSV. Proses selanjutnya user akan melakukan input nilai parameter age, trestbps, chol, sex, cpeftsthalach, exang, oldpeak, slope, c, thal, restecg, num. Kemudian sistem akan melakukan training dengan dataset untuk mendapatkan bobot untuk kemudian diklasifikasi menggunakan naïve bayes untuk menemukan hasil klasifikasi atau prediksi. Setelah mendapatkan hasil yang ditampilkan oleh sistem ke user, user dapat melakukan klasifikasi ulang atau menyelesaikan proses klasifikasi tersebut.



Gambar 3. ActivityDiagram Klasifikasi (Prediksi)

3.2. Analisis Model Data

Data yang akan diimplementasikan dalam penelitian merupakan data pada UCI Machine Learning dengan tema penyakit serangan jantung yang berjumlah 303 dataset dengan pembagian kelaskedalam 2 kelas yaitu terjangkau serta tidak terjangkau dan 15 atribut, dengan detail data seperti berikut:

1. (umur)
2. (jenis kelamin) jenis kelamin (1 = laki-laki; 0 = perempuan)
3. (cp) tipe nyeri dada
 - a. Nilai 1: khas angina
 - b. Nilai 2: angina atipikal
 - c. Nilai 3: nyeri non-angina
 - d. Nilai 4: tanpa gejala
4. (trestbps) tekanan darah istirahat (saat otot jantung istirahat) (dalam mm Hg saat masuk ke RSUD)
5. (chol) serum kolestoral (seluruh jumlah kolesterol dalam darah) dalam m/ dl
6. (fbs) (gula darah puasa / sebelum makan > 120m/dl) (1 = benar; 0 = salah)
7. (restecg) beristirahat hasil elektrokardiografi (alat pemeriksa otot jantung)
 - a. Nilai 0: normal
 - b. Nilai 1: memiliki kelainan gelombang ST-T (inversi gelombang T dan / atau ST
8. (thalach) denyut jantung maksimum tercapai
9. (exang) olahraga yang diinduksi angina (1 = ya; 0 = tidak)
10. (oldpeak) ST depresi yang disebabkan oleh olahraga relatif terhadap istirahat.
11. (slope/kemiringan) kemiringan segmen ST latihan puncak
 - a. Nilai 1: menaik
 - b. Nilai 2: datar
 - c. Nilai 3: downsloping

12. (ca) jumlah pembuluh darah (0-3) dijelaskan dengan fluoroskopi
13. (thal) 3 = normal; 6 = cacat tetap; 7 = cacat yang dapat dibalik
14. (num) (atribut yang diprediksi) diagnosis penyakit jantung (status penyakit angiografi)
 - a. Nilai 0: <50% penyempitan diameter
 - b. Nilai 1:> penyempitan diameter 50%

3.3. Pengujian Sistem

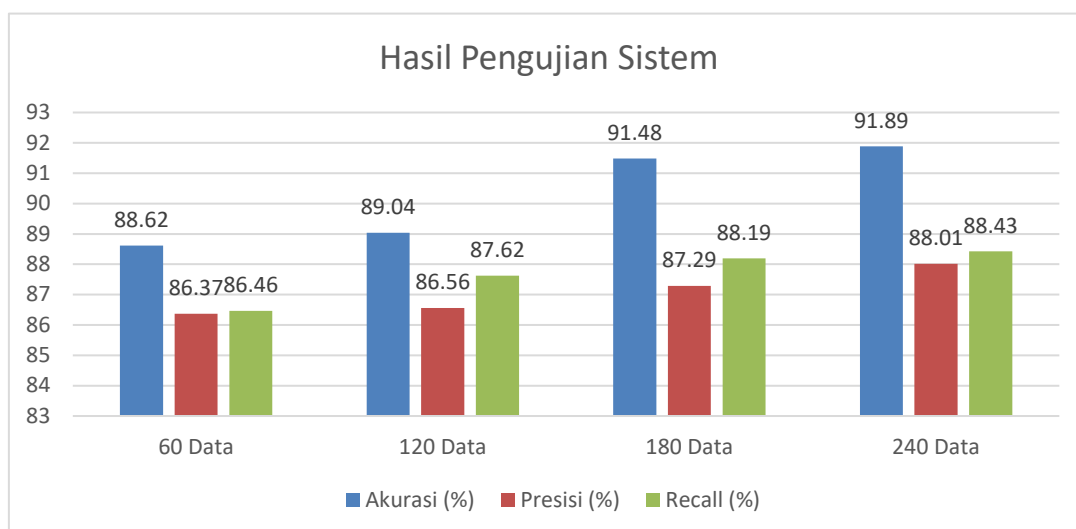
Model pengujian sistem dilakukan melalui beberapa percobaan dalam menentukan hasil nilai akurasi, presisi serta recall. Pengujian dilakukan dengan menggunakan 5-FOLD Cross-Validation, pembagian data dilakukan dengan membagi 303 data menjadi 5 bagian seperti pada Tabel 1.

Tabel 1. Hasil Pengujian Sistem

No	Data	Akurasi (%)	Presisi (%)	Recall (%)
1	60	88,62	86,37	86,46
2	120	89,04	86,56	87,62
3	180	91,48	87,29	88,19
4	240	91,89	88,01	88,43
5	303	92,02	88,98	89,03
Rata-Rata		90,61	87,44	87,95

Berdasarkan hasil pada Tabel 1 Diperoleh hasil subset 1 pada 60 data dengan nilai akurasi 88,62%, presisi 86,37% dan recall 86,46%, subset 2 pada 120 data dengan nilai akurasi 89,04%, presisi 86,56%, dan recall 87,62%, subset 3 pada 180 data dengan nilai akurasi 91,48%, presisi 87,29%, dan recall 88,19%. Dan pada subset ke 4, terdapat 240 data dengan nilai akurasi 91,89%, presisi 88,01%, dan recall 88,43%, sedangkan untuk hasil tertinggi pada subset 5 dengan 240 data dengan nilai akurasi 92,02%, presisi 88,98%, dan recall 89,03%.

Berdasarkan hasil yang telah diperoleh diatas diperoleh hasil rata-rata untuk pengujian 5 subset dengan nilai rata-rata akurasi sebesar 90,61%, presisi sebesar 87,44 %, recall sebesar 87,95%. Berdasarkan hasil tersebut terbentuk grafik pada Gambar 4.



Gambar 4. Hasil Pengujian Sistem

4. KESIMPULAN

Melalui hasil penelitian telah dijelaskan dan dilakukan pada pembuatan sistem klasifikasi penyakit jantung menggunakan naïve bayes. Pembuatan sistem ini menyimpulkan nilai hasil akurasi dengan rata-rata akurasi senilai 90,61%, rata-rata hasil nilai presisi senilai 87,44% dan rata-rata nilai recall senilai 87,95% dengan konfigurasi data yang terdapat pada UCI Machine Learning yang berisi 2 kelas klasifikasi dan 15 atribut dengan jumlah 303 data

5. SARAN

Berdasarkan hasil penelitian didapatkan saran untuk penambahan algoritma optimasi lain untuk menambah tingkat akurasi, presisi, dan recall. Penerapan algoritma baik berupa penggabungan ataupun optimasi algoritma diharapkan dapat meningkatkan nilai performa algoritma tersebut dalam melakukan klasifikasi.

DAFTAR PUSTAKA

- [1] Pemerintah Indonesia, 2009, Pedoman Pengendalian Penyakit Jantung dan Pembuluh Darah dalam Keputusan Menteri Kesehatan (KEMENKES), Lembaran RI Tahun 2019 No. 854, Sekretariat Negara, Jakarta.
- [2] Yahya, A. F., 2010, *Menaklukkan Pembunuh No. 1 Mencegah dan Mengatasi Penyakit Jantung Koroner Secara Tepat*, Qanita Publisher, Bandung.
- [3] Kasron, 2013, *Kelainan dan Penyakit Jantung: Pencegahan serta Pengobatannya*, Nuha Medika, Yogyakarta.
- [4] Praningki, T., Budi, I., 2017, Sistem Prediksi Penyakit Kanker Serviks Menggunakan CART, Naïve Bayes, dan K-NN., *Citec Journal*, No.2, Vol.4, Hal. 83-93
- [5] Prasetyo, Y., Haryanto, H., 2017, Visualisasi Berbasis Naïve Bayes Untuk Pemetaan Penyebaran Penyakit Infeksi Saluran Pernafasan Akut, *Jurnal Ilmiah SISFOTENIKA*, No.1, Vol 7, Hal. 74-84
- [6] Muhamad, H., Prasojo, A. A., Sugianto, N. A., Surtiningsih, L., Cholissodin, I., 2017, Optimasi Naïve Bayes Classifier Dengan Menggunakan Particle Swarm Optimization Pada Data Iris, *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, No. 3, Vol. 4, Hal. 180-184
- [7] Li, X., Ling, C. X., Wang, H, 2016, The Convergence behavior of Naive Bayes on Large Sparse Datasets, *Journal ACM Transactions on Knowledge Discovery from Data (TKDD)*, No. 1, Vol. 11, Hal. 853-858.
- [8] Nuraeni, N., 2017, Penentuan Kelayakan Kredit Dengan Algoritma Naïve Bayes Classifier: Studi Kasus Bank Mayapada Mitra Usaha Cabang PG, *Jurnal Teknik Komputer AMIK BSI*, No. 1, Vol. 3, Hal. 9-15
- [9] Kusumadewi, S., 2003, *Artificial Intelligence Teknik dan Aplikasi*, Graha Ilmu, Yogyakarta.
- [10] Turban, E., 2005., *Decision Support System and Intelligent Systems, edisi Bahasa Indonesia jilid 1*, Penerbit Andi, Yogyakarta.