

**IMPLEMENTASI METODE K-NEAREST NEIGHBORS (KNN)
UNTUK KLASIFIKASI PENYAKIT JANTUNG**

SKRIPSI



Oleh:

Ahmad Yogianto

2020503008

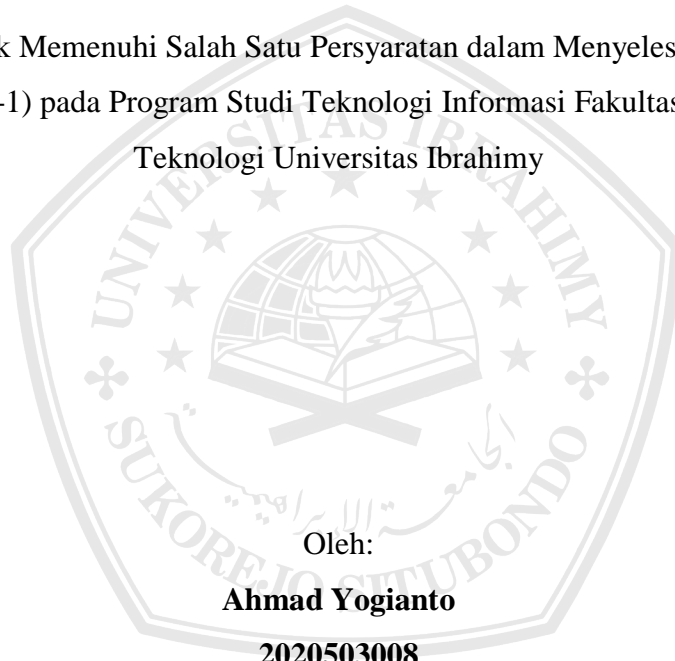
**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI UNIVERSITAS IBRAHIMY
SITUBONDO**

2024

**IMPLEMENTASI METODE K-NEAREST NEIGHBORS (KNN)
UNTUK KLASIFIKASI PENYAKIT JANTUNG**

SKRIPSI

Diajukan untuk Memenuhi Salah Satu Persyaratan dalam Menyelesaikan Program Sarjana (S-1) pada Program Studi Teknologi Informasi Fakultas Sains dan Teknologi Universitas Ibrahimy



Oleh:

Ahmad Yogianto

2020503008

**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI UNIVERSITAS IBRAHIMY
SITUBONDO**

2024

PERNYATAAN KEASLIAN TULISAN
PERNYATAAN KEASLIAN TULISAN


Saya yang bertanda tangan di bawah ini:

Nama : **Ahmad Yogiarto**
NPM : 2020503008
Prodi : S-1 Teknologi Informasi
Fakultas : Fakultas Sains dan Teknologi

Menyatakan dengan sebenarnya, bahwa tugas skripsi ini secara keseluruhan adalah hasil penelitian atau karya saya sendiri, kecuali pada bagian-bagian yang dirujuk sebagai sumber referensi dan disebutkan dalam daftar pustaka. Apabila di kemudian hari terbukti atau dapat dibuktikan bahwa skripsi ini hasil plagiarasi, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Situbondo, 15 Juli 2024

Saya yang menyatakan,



Ahmad Yogiarto
Ahmad Yogiarto

**PERSETUJUAN PEMBIMBING
PERSETUJUAN PEMBIMBING**

Nama : **Ahmad Yogianto**
NPM/NIRM : 2019503050
Judul : **Implementasi Metode K-Nearest Neighbors (KNN)
Untuk Klasifikasi Penyakit Jantung**



Telah disetujui oleh :

Pembimbing I,

Ahmad Homaidi, M.Kom
NIDN: 0705078901

Pembimbing II,

Zaehol Fatah, M.Kom
NIDN: 0715057801

PENGESAHAN

SKRIPSI

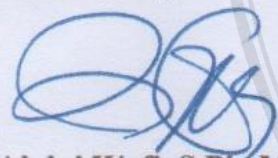
**IMPLEMENTASI METODE K-NEAREST NEIGHBORS (KNN)
UNTUK KLASIFIKASI PENYAKIT JANTUNG**

Ahmad Yogianto
2020503008

Telah dipertahankan di depan dewan penguji Sidang/Munaqasyah Skripsi pada hari Rabu, tanggal 07 Agustus 2024 sebagai salah satu syarat memperoleh gelar Sarjana (S.Kom) pada Fakultas Sains dan Teknologi Universitas Ibrahimi.

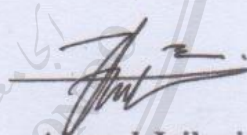
Tim Penguji,

Ketua Sidang,



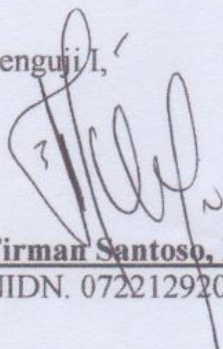
Abdul Wafi, S.Pi., M.P
NIDN. 0705049103

Sekretaris Sidang,



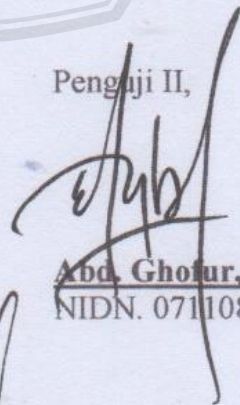
Ahmad Jailani, S.Kom

Penguji I,



Firman Santoso, M.Kom
NIDN. 0722129201

Penguji II,



Abd. Ghofur, M.Kom
NIDN. 0711088303

Mengetahui
Mekan,



Abd. Ghofur, M.Kom
NIDN. 0711088303

MOTTO

لَهُ مُعَقَّبَاتٌ مِنْ بَيْنِ يَدَيْهِ وَمِنْ خَلْفِهِ يَحْفَظُونَهُ مِنْ أَمْرِ اللَّهِ إِنَّ اللَّهَ لَا يُغَيِّرُ مَا بِقَوْمٍ
حَتَّى يُغَيِّرُوا مَا بِأَنْفُسِهِمْ وَإِذَا أَرَادَ اللَّهُ بِقَوْمٍ سُوءًا فَلَا مَرَدَّ لَهُ وَمَا لَهُمْ مِنْ دُونِهِ
مِنْ وَالٍ

“Bagi manusia ada malaikat-malaikat yang selalu mengikutinya bergiliran, di muka dan di belakangnya, mereka menjaganya atas perintah Allah. Sesungguhnya Allah tidak mengubah keadaan sesuatu kaum sehingga mereka mengubah keadaan yang ada pada diri mereka sendiri. Dan apabila Allah menghendaki keburukan terhadap sesuatu kaum, maka tak ada yang dapat menolaknya; dan sekali-kali tak ada pelindung bagi mereka selain Dia.”

(Q.S. Ar-Ra'd: 11)

HALAMAN PERSEMBAHAN

Dalam penulisan skripsi/ tugas akhir ini, penulis persembahkan untuk beberapa orang yang berkontribusi banyak dalam membantu penulis menyelesaikan tugas ini. Mereka adalah:

1. Segala puji bagi Allah SWT, Tuhan Semesta Alam, yang telah memberikan rahmat, hidayah, dan kekuatan kepada saya untuk menyelesaikan skripsi ini. Hanya dengan izin-Mu, ya Allah, segala rintangan dapat terlewati, dan langkah ini bisa sampai pada tahap akhir. Semoga apa yang telah saya capai ini menjadi berkah dan bermanfaat bagi saya, keluarga, dan orang-orang di sekitar saya. Terima kasih, ya Allah, atas segala nikmat yang tak pernah henti Engkau curahkan.
2. Shalawat dan salam semoga selalu tercurah kepada Nabi Muhammad ﷺ, suri teladan umat manusia, yang melalui ajarannya saya dapat menemukan arah dan makna dalam setiap langkah kehidupan. Semoga saya selalu berada dalam jalan yang beliau ajarkan dan mendapat syafaatnya di hari kemudian !.
3. Untuk diri ini sendiri yang telah melewati segala rintangan, kuat dan tetap tegar dalam menjalani setiap napas perjuangan, hingga sampai pada tahap akhir perkuliahan ini. Terima kasih telah kuat dan sabar sepanjang jalan. Kamu sungguh luar biasa.
4. Kepada kedua orang tua tercinta, **Aba Fendi Yanto** dan **Umi Rusmiyanti Ningsih**, terima kasih atas cinta, dukungan, dan kesabaran yang tak terbatas.

Perjalanan ini takkan mungkin terwujud tanpa doa dan pengorbanan kalian. Segala yang telah kulalui tak sebanding dengan kasih sayang yang kalian berikan. Kalian adalah segalanya bagiku.

5. Untuk seluruh guru dan dosen yang telah membagikan ilmu dan bimbingannya, Meskipun kadang ilmu tersebut terasa sulit, semoga bisa menjadi bekal yang membawaku menuju masa depan yang lebih baik..
6. Untuk teman-teman seperjuangan S3TV, tawa dan kebersamaan kita yang telah kita lalui. terima kasih atas dukungan dalam menjalani dan menyelesaikan skripsi ini. Juga kepada seluruh teman TI Putra Angkatan 2020, semoga kita bisa meraih gelar sarjana bersama dan menjadi individu yang bermanfaat bagi sesame dan semoga kita selalu diberkahi dan sukses dalam segala hal!

KATA PENGANTAR

Segala puji syukur peneliti sampaikan kepada Allah SWT, karena atas Rahmat dan Hidayah-Nya, perencanaan, pelaksanaan dan penyelesaian tugas akhir/skripsi dengan judul “Implementasi Metode K-Nearest Neighbors (KNN) untuk Klasifikasi Penyakit Jantung” sebagai salah satu syarat penyelesaian program diploma/sarjana dapat terselesaikan dengan baik dan lancar.

Kesuksesan ini dapat peneliti peroleh karena dukungan beberapa pihak. Peneliti menyampaikan terima kasih kepada :

1. **KHR. Ach Azaim Ibrahimy, S.Sy** selaku Pengasuh Pondok Pesantren Salafiyah Syafiiyah Sukorejo Situbondo.
2. **KH. Ach. Fadhail, S.H., M.H** selaku Rektor Universitas Ibrahimy.
3. **Bapak Abd. Ghofur, M.Kom** selaku Dekan Fakultas Sains dan Teknologi.
4. **Bapak Firman Santoso, M.Kom**, selaku Ka. Prodi Tekonologi Informasi.
5. **Bapak Ahmad Homaidi, M.Kom** dan **Bapak Zaehol Fatah, M.Kom** selaku pembimbing I dan II.

Semoga amal baik yang telah diberikan oleh Bapak/Ibu kepada peneliti mendapat balasan yang sebaik mungkin dari Allah SWT, Amin.

Situbondo, 15 Juli 2024


Peneliti

DAFTAR ISI

	Halaman
SKRIPSI	i
PERNYATAAN KEASLIAN TULISAN	ii
PERSETUJUAN PEMBIMBING	iii
PENGESAHAN	iv
MOTTO	v
HALAMAN PERSEMBAHAN	vi
KATA PENGANTAR	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
ABSTRACT	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	3
1.3 Rumusan Masalah	4
1.4 Batasan Masalah	4
1.5 Tujuan Penelitian	5
1.6 Manfaat Penelitian	5
1.7 Metodologi Penelitian	6
1.7.1 Jenis Penelitian	6
1.7.2 Metode Pengumpulan Data	6
1.8 Sistematika Pembahasan	7
BAB II TINJAUAN PUSTAKA	9
2.1 Penelitian Terdahulu	9
2.2 Landasan Teori	13
2.2.1 Data Mining	13
2.2.2 Metode K-Nearest Neighbour (K-NN)	13
2.2.3 Klasifikasi	14
2.2.4 Penyakit Jantung	14

2.3 Perangkat Lunak yang Digunakan	15
BAB III METODELOGI PENELITIAN	18
3.1 Metode Penelitian.....	18
3.2 Tahapan Penelitian	19
3.2.1 Pengumpulan Data.....	20
3.2.2 Penentuan Atribut.....	20
3.2.3 Proses <i>Data Mining</i>	31
3.3 Perancangan Sistem	34
3.3.1 Metode K-Nearest Neighbors (K-NN).....	34
BAB IV HASIL DAN PEMBAHASAN	39
4.1 Hasil Penelitian.....	39
4.2 Perhitungan Manual Metode K-Nearest Neighbors	39
4.2.1 Tentukan Nilai K.....	39
4.2.2. Hitung jarak	39
4.2.3. Tentukan Tetangga Terdekat.....	43
4.2.4. Tentukan Kelas Prediksi.....	43
4.2.5 Menghitung Nilai <i>Accuracy, Sensitivity, Specificity, PPV dan NPV</i>	44
4.3 Perhitungan RapidMiner.....	47
4.3.1 Pemrosesan Data	47
4.3.2 Perhitungan dan Validasi	50
4.3.3 Hasil Performance Vector	52
BAB V PENUTUP.....	56
5.1 Kesimpulan	56
5.2 Saran	56
DAFTAR PUSTAKA	15
LAMPIRAN	17
CURRICULUM VITAE	

DAFTAR TABEL

Tabel	Halaman
3.1: Atribut Umur.....	19
3.2 : Atribut Jenis Kelamin	21
3.3 : Atribut Jenis Nyeri Dada.....	21
3.4 : Atribut Tekanan Darah	22
3.5 : Atribut Tensi Darah	23
3.6 : Atribut Kadar Gula Darah.....	23
3.7 : Atribut Reesting Electrocardiographic results	24
3.8 : Atribut Keadaan setelah olahraga	25
3.9 : Atribut Slope	25
3.10 : Atribut Pembuluh Darah Utama	26
3.11 : Atribut Tes Stress Thallium.....	27
3.12 : Atribut Prediksi Penyakit Jantung.....	27
4.1 : Tabel Data Set.....	36
4.2 : Tabel Data Test.....	37

DAFTAR GAMBAR

Gambar	Halaman
Gambar 3. 1 Tahapan Penelitian	19
Gambar 3. 2 Perhitungan Manual K-Nearest Neighbors(KNN)	34
Gambar 3. 3 <i>Flowchart Perhitungan K-Nearest Neighbors (KNN)</i>	36
Gambar 3. 4 <i>User Interface RapidMiner</i> versi 9.10.....	37
Gambar 3. 5 Kerangka Pemikiran pada <i>RapidMiner</i> 9.10	38
Gambar 4. 1 Proses Pembuatan Perhitungan Baru.....	48
Gambar 4. 2 Penentuan jenis Data	48
<i>Gambar 4. 3 Data Value</i>	49
Gambar 4. 4 <i>Transformasi Data</i>	50
Gambar 4. 5 Model Metode K-Nearest Neighbors (KNN).....	50
Gambar 4. 6 Hasil data testing dari Algoritma <i>K-Nearest Neighbors</i>	51
Gambar 4. 7 Hasil Akurasi	52
Gambar 4. 8 Perhitungan Precision.....	53
Gambar 4. 9 Perhitungan Recall	54
Gambar 4. 10 AUC (<i>Area Under Curve</i>).....	55

ABSTRAK

Ahmad Yogianto. 2024. **Implementasi Metode K-Nearest Neighbors (KNN) Untuk Klasifikasi Penyakit Jantung**. Skripsi. Program Studi Teknologi Informasi. Universitas Ibrahimi. Pembimbing: (1) Ahmad Homaidi, M.Kom., (2) Zaehol Fatah, M.Kom.

Penyakit Jantung masuk ke dalam kelompok penyakit kardiovaskuler dan salah satu penyebab utama dari kehilangan nyawa secara global. Hal yang diakibatkan terjadinya penyumbatan terhadap aliran darah yang masuk ke dalam otot jantung, yang mengakibatkan terjadinya kerusakan berat pada jantung.

Meskipun penyakit ini tidak menular WHO menyebutkan bahwa penyakit jantung telah menyebabkan 17,9 juta orang meninggal pada setiap tahun. Data asli berasal dari database Cleveland dari UCI Machine Learning Repository yang terdapat 303 data sample dengan 14 atribut untuk memprediksi pasien yang positif dan negatif terkena penyakit jantung.

RapidMiner merupakan aplikasi yang digunakan dengan metode K-Nearest Neighbors (KNN), cara kerja metode KNN sendiri ialah dengan cara menentukan nilai dari data yang akan diuji berdasarkan pada K data tetangga yang paling dekat pada dataset yang akan di uji. Hasil pada metode KNN dengan parameter K=5 ini nilai Akurasi yang didapatkan sebesar 64,03%, nilai presesinya sebesar 64.58%, dan recall yang didapatkan sebesar 75.15%.

Kata Kunci : K-Nearest Neighbors (KNN), Klasifikasi, Penyakit Jantung, Rapidminer

ABSTRACT

Ahmad Yogiarto. 2024. **Implementation of Metode K-Nearest Neighbors (KNN) for Classification of Heart Disease..** Thesis.
Information Technology Study Program. Ibrahimi University. Supervisors:
(1) Ahmad Homaidi, M.Kom, (2) Zaehol Fatah, M.Kom.

Heart disease is included in the group of cardiovascular diseases and is one of the main causes of loss of life globally. This is caused by a blockage in blood flow entering the heart muscle, which results in serious damage to the heart.

Even though this disease is not contagious, WHO states that heart disease causes 17.9 million people to die every year. The original data comes from the Cleveland database from the UCI Machine Learning Repository which contains 303 data samples with 14 attributes to predict patients who are positive and negative for heart disease.

RapidMiner is an application used with the K-Nearest Neighbors (KNN) method. The way the KNN method itself works is by determining the value of the data to be tested based on the K data of the closest neighbors in the dataset to be tested. The results of the KNN method with parameter $K=5$ are that the accuracy value obtained is 64.03%, the precision value is 64.58%, and the recall value obtained is 75.15%.

Keywords: K-Nearest Neighbors (KNN), Classification, Heart Disease, Rapidminer

BAB I

PENDAHULUAN

1.1 Latar Belakang

Data Mining berasal dari dua kata, yaitu data dan mining, di dalam bukau tentang algoritma data mini dan pengajian, data merupakan sekumpulan dari fakta-fakta yang terekam atau suatu entitas yang memiliki informasi tertentu dan masih terabaikan, kegunaan dar mining sendiri iyalah bermakna proses penambangan. jika digabungkan data mining merupakan sebagai suatu proses penambangan atau penggalian suatu informasi tertentu dari data yang dicari, sehingga bisa menghasilkan suatu output berupa pengetahuan baru yang dapat diambil manfaatnya [1].

Data Mining merupakan rangkaian proses yang dilakukan untuk menggali suatu informasi atau pengetahuan yang tidak diketahui secara manual dari data yang besar, informasi yang didapatkan tersebut dapat diperoleh dengan cara melakukan ekstaksi dan mengenali pola atau atribut yang penting dari data yang terdapat pada basis data yang besar sehingga sering disebut dengan *Knowledge Discovery Databases* (KDD) [2].

Secara garis besar, terdapat lima metode utama dalam data mining, yaitu Estimasi, Prediksi, Klastering, Asosiasi, dan Klasifikasi. Metode Estimasi digunakan untuk dapat memprediksi nilai variabel target yang bersifat kontinu berdasarkan variabel prediktor, dengan metode yang digunakan seperti Support Vector Machine (SVM), Neural Network, dan Linear Regression. algoritma selanjutnya adalah Metode Prediksi atau forecasting yang mirip dengan Estimasi, namun yang membedakannya adalah lebih fokus pada proyeksi nilai numerik di masa depan. Metode Klastering bertujuan untuk dapat mengelompokkan data yang belum diketahui kelasnya ke dalam klaster tertentu berdasarkan karakteristik yang dimiliki, menggunakan algoritma klastering ini seperti metode K-means, Fuzzy C-Means, serta Self-Organizing Map (SOM). Sementara itu, Algoritma Asosiasi berfungsi untuk dapat menemukan

keterkaitan antara atribut-atribut yang sering muncul bersama dalam sebuah transaksi, algoritma ini yang sangat bermanfaat di dalam dunia bisnis, dan metode yang termasuk dalam algoritma adalah seperti metode Apriori Algorithm, GRI Algorithm, serta FP-Growth Algorithm. algoritma Terakhir adalah algoritma Klasifikasi, yang digunakan untuk dapat membedakan kelas atau mengidentifikasi objek berdasarkan data yang telah diketahui sebelumnya, metode yang biasa digunakan pada algoritma klasifikasi ini seperti metode Naive Bayes, Linear Discriminant Analysis, C4.5, serta K-Nearest Neighbors (KNN) [3]. Memahami kelima metode ini sangat penting dan bermanfaat terutama bagi peneliti untuk dapat memilih pendekatan yang tepat untuk menganalisis data secara efektif dan lebih efisien.

Seperti yang telah dijelaskan sebelumnya, Salah satu Algoritma yang dapat digunakan untuk menyelesaikan suatu permasalahan adalah dengan menggunakan Metode K-Nearest Neighbor. Metode K-Nearest Neighbor ini merupakan suatu algoritma yang digunakan untuk proses klasifikasi data mining berdasarkan pada kedekatan terhadap objek yang baru atau tertangga terdekatnya [4]. Banyaknya tetangga terdekat biasa disimbolkan dengan (K) memiliki pengaruh terhadap proses pengambilan keputusan pada algoritma K-NN. K-NN sendiri termasuk bagian dari pada algoritma terpelajar (*supervised learning*) yang artinya pada dataset diharuskan memiliki target ataupun kelas, pada K-NN penentuan nilai K berdasarkan dengan perhitungan nilai *euclidean distance* [5].

Jantung merupakan organ manusia yang berperan dalam sistem peredaran darah, penyakit jantung merupakan suatu kondisi dimana jantung tidak dapat melaksanakan tugasnya untuk peredaran darah tidak berjalan dengan baik [6]. Penyakit jantung merupakan suatu kondisi medis yang melibatkan berbagai gangguan pada jantung. Penyakit ini termasuk penyakit arteri koroner (penyakit arteri koroner), gagal jantung, penyakit katup jantung, aritmia jantung, dan penyakit jantung bawaan. Ini merupakan salah satu penyebab kematian paling umum di dunia [7]. Penyebab utama dari penyakit jantung ini sendiri bisa disebabkan oleh gaya hidup yang tidak sehat,

diantaranya adalah diet tinggi lemak dan garam, kurangnya aktivitas fisik, kebiasaan merokok, konsumsi alkohol yang terlalu berlebihan dan, stres. Faktor resiko lainnya yang menyebabkan penyakit jantung adalah obesitas, tekanan darah tinggi, kadar kolesterol tinggi, diabetes, dan riwayat keluarga yang pernah mengalami penyakit jantung [8].

Pada tahun 2012, Badan Kesehatan dunia atau The World Health Organization (WHO) menyebutkan bahwa 17,5 juta orang di dunia meninggal dunia dikarenakan penyakit kardiovaskular. Penyakit jantung ini merupakan salah satu jenis dari penyakit kardiovaskular yang ikut menyumbang 31% dari total 56,5 juta kasus kematian di dunia. Penyakit jantung terjadi ketika darah yang mengalir ke otot jantung berhenti, hal ini menyebabkan gangguan yang terjadi pada jantung. Hal ini menyebabkan adanya kebutuhan untuk mengklasifikasikan terhadap pasien yang memiliki penyakit jantung atau tidak, hal ini juga dapat membantu tenaga medis dalam mengambil keputusan untuk mengambil tindakan pencegahan terhadap penderita penyakit jantung [7]. Data dari UCI Machine Learning Repository menyebutkan bahwa terdapat 13 atribut yang dapat digunakan untuk dapat mengdiagnosis pasien apakah memiliki disinyalir positif memiliki penyakit jantung atau tidak [9].

Berdasarkan latar belakang yang telah dijelaskan diatas dapat menjadi bahan evaluasi dan saran kepada tenaga medis dalam mengambil keputusan untuk mengambil tindakan pencegahan terhadap penderita penyakit jantung sehingga perlu kiranya penulis mengangkat judul **IMPLEMENTASI METODE K-NEAREST NEIGHBORS (KNN) UNTUK KLASIFIKASI PENYAKIT JANTUNG** dalam penerapan penelitian ini.

1.2 Identifikasi Masalah

Berdasarkan latar belakang diatas dapat diidentifikasi beberapa masalah, berikut adalah masalah yang diidentifikasi pada penelitian ini yaitu:

- a. Mendeteksi seberapa besar kemungkinan pasien yang memiliki penyakit jantung dan tidak memiliki penyakit jantung

- b. Mendeteksi tingkat kesehatan pasien yang disinyalir memiliki penyakit jantung atau tidak sehingga dapat memudahkan penanganan yang dapat diambil oleh pengambil keputusan.

1.3 Rumusan Masalah

Berdasarkan uraian di atas baik dari latar belakang atau pada identifikasi masalah, maka rumusan masalahnya adalah sebagai berikut:

- a. Bagaimana implementasi metode K-Nearest Neighbor (KNN) untuk klasifikasi penyakit jantung?
- b. Bagaimana performa metode K-Nearest Neighbor (KNN) dalam mengklasifikasikan penyakit jantung?

1.4 Batasan Masalah

Dalam penelitian dengan teori Data Mining ini, banyak aspek yang dijadikan pertimbangan dalam penyusunan laporan tugas akhir ini maka perlu diberikan batasan masalah yaitu:

Adapun batasan masalah dalam penelitian ini adalah:

- a. Penelitian ini akan membahas data yang telah tersedia dalam data publik sehingga untuk perkembangan data terbaru dapat dilakukan penelitian yang baru mengikuti perkembangan data yang ada.
- b. Penelitian ini menggunakan penelitian yang bersifat cross-sectional sehingga hanya dapat meneliti pada satu waktu dan tidak dapat berhubungan dengan waktu yang berbeda atau tidak berkesinambungan sebagaimana penelitian longitudinal.
- c. Penelitian ini hanya menggunakan satu macam metode yakni Metode K-Nearest Neighbor dengan tanpa membandingkan dengan algoritma

data mining yang lainnya. Sehingga klasifikasi data hanya terbatas pada Metode K-Nearest Neighbors (KNN) saja.

1.5 Tujuan Penelitian

Berdasarkan perumusan masalah diatas maka tujuan dari penelitian ini adalah untuk mengetahui prediksi tingkat kesehatan pasien dalam Penyakit jantung sehingga dokter maupun pihak yang bersangkutan dapat mendiagnosis dan mengklasifikasikan penyakit jantung secara lebih akurat dan efisien sehingga dapat melakukan tindakan pencegahan yang tepat dan tepat.

1.6 Manfaat Penelitian

Sedangkan manfaat dari pengklasifikasian penyakit jantunga diantaranya adalah:

a. Bagi Peneliti

Penelitian ini memberikan pengalaman yang sangat berharga bagi peneliti untuk menerapkan secara langsung ilmu yang diperoleh selama menempuh pendidikan di perguruan tinggi.

b. Manfaat bagi Universitas

Dapat menjadikan penelitian ini salah satu sumber referensi maupun sumbangsih dalam pengakreditasian program studi sehingga dapat meningkatkan taraf akreditasi di Universitas Ibrahimi khususnya Fakultas Sains dan Teknologi.

c. Manfaat bagi peneliti lain

Hasil dari penelitian ini diharapkan dapat menjadi referensi bagi penelitian selanjutnya dalam pengembangan metode klasifikasi penyakit jantung yang lebih efektif.

d. Manfaat bagi pasien

Hasil dari penelitian ini, diharapkan dapat meningkatkan kesadaran terhadap masyarakat luas tentang pentingnya diagnosis dan klasifikasi penyakit jantung secara dini.

1.7 Metodologi Penelitian

1.7.1 Jenis Penelitian

a. *Library Online Research*

Peneliti menghimpun, mengolah, serta menganalisis data-data terkait Klasifikasi Penyakit jantung dengan Metode K-Nearest Neighbors (K-NN) yang bersumber dari artikel, jurnal, Perpustakaan online, serta buku.

b. *Penelitian Deskriptif Kuantitatif*

Metode penelitian deskriptif kuantitatif ini dipilih untuk menggambarkan, mendeskripsikan, dan menjelaskan sesuatu secara apa adanya serta menarik kesimpulan dari data berupa angka – angka yang telah disajikan.

1.7.2 Metode Pengumpulan Data

Metode pengumpulan data yang digunakan adalah Study Literature dengan jenis data sekunder yang didapatkan dari Website Kaggle dengan judul Klasifikasi Penyakit Jantung yang ditulis oleh Dirga Halim Susilo.

Tautan yang dapat diakses adalah sebagai berikut <https://www.kaggle.com/code/dirghalimsusilo/klasifikasi-penyakit-jantung>. Data terdiri dari 303 dataset dan Data asli berasal dari database Cleveland dari UCI Machine Learning Repository. Data ini dihimpun dengan 14 atribut atau fitur berkenaan dengan apakah Pasien memiliki penyakit jantung atau tidak.

1.8 Sistematika Pembahasan

Sistematika pembahasan yang akan dicantumkan dalam karya tulis ilmiah ini adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi tentang latar belakang penelitian, identifikasi, rumusan, dan Batasan masalah, tujuan dan manfaat penelitian, metode dan jenis penelitian, Teknik pengumpulan data, metode pengembangan data, metode pengembangan sistem, dan sistematika pembahasan.

BAB II TINJAUAN PUSTAKA

Pada Bab dua ini berisi, tentang beberapa literatur yang menjadi dasar pemikiran dalam penelitian yang dilakukan, diantaranya adalah penelitian terdahulu, landasan teori, dan perangkat lunak yang digunakan.

BAB III METODELOGI PENELITIAN

Bab ini pembahasannya tentang bagaimana peneliti mendapatkan data yang digunakannya dalam penelitian ini dan bagaimana cara untuk mendapatkannya berikut memprosesnya

sebelum melakukan Analisa lebih lanjut menggunakan algoritma yang terdapat dalam Data Mining.

BAB IV HASIL DAN PEMBAHASAN

Bab ini merupakan hasil dari Analisa yang telah diputuskan sebelumnya. Berisi tentang beberapa penjelasan maupun perhitungan dari data yang telah diolah sehingga menghasilkan sebuah prediksi berdasarkan kemungkinan yang telah didapatkan sebelumnya dan juga tingkat akurasi dari prediksi yang telah dihasilkan.

BAB V PENUTUP

Di bab terakhir ini berisi tentang kesimpulan maupun saran yang diberikan dari sistem yang telah diteliti, dirancang, dan diimplementasikan.

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Membaca dari penelitian terdahulu bertujuan untuk mendapatkan bahan dan referensi berbanding dan acuan. Selain itu untuk menghindari anggapan kesamaan dengan penelitian ini. Maka dalam kajian pustaka ini peneliti mencantumkan hasil penelitian terdahulu sebagai berikut

Klasifikasi Penyakit Gagal Jantung Menggunakan Algoritma K-Nearest Neighbor.

Penelitian Penelitian ini dikerjakan oleh Yovi Pratama, Anton Prayitno, Defri Nazrian, Nur aini, Yoga Rizki R, dan Errisya Tasywir dari Prodi Teknik Informasika Fakultas Ilmu Komputer Unicersitas Dinamika Bangsa Kota Jambi, dimana tujuan penelitian ini iyalah dengan menggunakan algoritma K-Nearest Neighbor dalam klasifikasi penyakit gagal jantung, dengan fokus pada prediksi dan identifikasi kematian akibat gagal jantung. Penelitian bertujuan untuk menunjukkan bahwa algoritma KNN dapat digunakan untuk prediksi dengan akurasi tinggi, serta membantu dalam deteksi dini dan manajemen penyakit kardiovaskular. Atribut yang digunakan pada penelitian ini yaitu Age, Sex, ChestPainType, RestingBP, Colestrol, FastingBS, RestingECG, MaxHR, ExerciseAngina, Oldpeak, ST_Slope, HeartDiasease. aplikasi yang digunakan yaitu rapidminer versi 9.10 dengan operator retriive Heart, Split Data, , Performance, Algoritma K-NN, dan Apply Model . Dataset yang digunakan sebanyak 918 data dan 12 atribut. Kemudian hasil implementasi algortima KNN pada aplikasi ripedminer dilakukan dengan

menggunakan pergantian nilai k , dan hasil akurasi tertinggi didapat pada nilai $k=9$ dengan akurasi sebesar 70,65%, nilai presesinya adalah 75 %, sedangkan recall menghasilkan 70,73% [10].

Penerapan Metode Klasifikasi K-Nearest Neighbor pada Dataset Penderita Penyakit Diabetes

Jurnal ini dibuat oleh Andi Maulida Argina dari Universitas Muslim Indonesia, jurnal ini dan tujuan dari penelitian in adalah Tujuan dari penelitian ini adalah untuk menerapkan metode klasifikasi K-Nearest Neighbor pada dataset penderita penyakit diabetes dan mengukur akurasi, presisi, recall, dan F-Measure berdasarkan nilai K yang berbeda. Selain itu, penelitian ini juga bertujuan untuk mengevaluasi performa metode klasifikasi tersebut dalam memprediksi penyakit diabetes pada dataset yang digunakan. Data mining memiliki peran yang sangat penting dalam penelitian ini, hal ini karena data mining digunakan untuk mengekstraksi pola dan informasi yang bermanfaat dari dataset yang ada yakni penderita penyakit diabetes. Proses data mining membantu dalam analisis dan pengolahan data untuk mengaplikasikan metode klasifikasi K-Nearest Neighbor dan mengukur performanya dalam memprediksi penyakit diabetes.

Data mining merupakan suatu proses yang menggunakan statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat. Data Mining didefinisikan sebagai proses penemuan pola dalam data. Berdasarkan tugasnya, data mining dapat dikelompokkan menjadi deskripsi, estimasi, prediksi, klasifikasi, clustering, dan asosiasi. Proses dalam tahap data mining terdiri dari tiga langkah utama, yaitu data

preparation. Pada langkah ini, data dipilih, dibersihkan, dan dilakukan preprocessing mengikuti pedoman dan knowledge dari ahli domain yang menangkap dan mengintegrasikan data internal dan eksternal ke dalam tinjauan organisasi secara menyeluruh. Penggunaan algoritma data mining dilakukan pada langkah ini untuk menggali data yang terintegrasi untuk memudahkan identifikasi informasi bernilai.

Artikel ini membahas tentang penerapan metode klasifikasi K-Nearest Neighbor pada dataset penderita penyakit diabetes. Penelitian ini bertujuan untuk mengukur akurasi, presisi, recall, dan F-Measure berdasarkan nilai K yang berbeda. Hasil penelitian menunjukkan bahwa akurasi tertinggi yang dicapai adalah sebesar 39% saat nilai $K=3$. Selain itu, presisi tertinggi diperoleh pada nilai $K=3$ dan $K=5$ sebesar 65%, recall tertinggi adalah 36% pada $K=3$, dan F-Measure tertinggi adalah 46% pada $K=3$. Selain itu, artikel juga membahas berbagai metode klasifikasi lainnya seperti Decision Tree, Support Vector Machine, Naive Bayes, Random Forest, dan lainnya yang digunakan untuk berbagai aplikasi seperti klasifikasi penyakit, identifikasi kaca, readmisi pasien diabetes, sistem penentuan posisi dalam ruangan, deteksi tepi, dan lain-lain. Artikel juga mencakup topik seperti kurva presisi-recall-gain, dasar-dasar pembelajaran mesin, dan analisis kinerja sistem [11].

Penerapan Algoritma Klasifikasi Nearest Neighbor (K-NN) Untuk Mendeteksi Penyakit Jantung.

Penelitian ini dilakukan oleh Mei Lestari dari Program Studi Teknik Informatika, Fakultas Teknik, Matematika dan IPA Universitas Indraprasta Jakarta

Selatan. Penelitian ini melakukan dengan cara pengelompokan data yang telah dikumpulkan sebelumnya, hal ini dilakukan untuk mendiagnosa penyakit jantung menggunakan dengan menggunakan algoritma K-NN dengan $k = 9$ pada data pasien, serta dengan mengukur tingkat akurasi pendeteksian penyakit jantung. dalam penelitian ini, Terdapat 13 atribut yang digunakan untuk mendiagnosa penyakit jantung, atribut-atribut tersebut diantaranya adalah Age, Sex, Chest Pain Type, Tresting Blood Pressure, Serum cholestoral dalam mg/dl, Fasting blood sugar >120 mg/dl, Resting electrocardiographic result, The Slope of the peak exercise ST segment, Excercise Induced Angina, Old Peak, CA (Number of Major Vessels), Maximum Heart Rate, Achieved (Thalac), Thal dan dalam penelitan ini Terdapat 100 data pasien penyakit jantung yang diolah dengan menggunakan algoritma Klasifikasi Nearest Neighbor (KNN) dengan $k = 9$.

Tujuan dari penelitian ini adalah untuk menggunakan algoritma K-NN dengan $k = 9$ untuk mendiagnosa penyakit jantung pada pasien dan mengukur tingkat akurasi pendeteksian penyakit jantung. Penelitian ini juga bertujuan untuk mengevaluasi efektivitas algoritma K-NN dalam mendeteksi penyakit jantung serta memberikan rekomendasi untuk penelitian selanjutnya, termasuk perbandingan dengan algoritma data mining lainnya untuk menentukan keakuratan dan efisiensi yang lebih baik dalam mendeteksi penyakit jantung [4].

2.2 Landasan Teori

Landasan teori ini merupakan beberapa referensi atau pengertian dari beberapa kata kunci yang akan dibahas di dalam penelitian ini. Berikut ini merupakan beberapa dari pengertian tersebut.

2.2.1 Data Mining

Data mining merupakan suatu proses untuk mengetahui pengetahuan baru, Pola dan tren yang di pilah dari jumlah data yang begitu banyak yang disimpan dalam suatu repositori atau suatu tempat penyimpanan dengan menggunakan teknik pengenalan pola serta statistik dan teknik matematika. data mining akan memecahkan suatu masalah dengan cara mengidentifikasi, Mengestraksi dan menganalisis data yang telah ada dalam basis data [12].

Data Mining juga merupakan kegiatan untuk menemukan pola dari data dalam jumlah besar, data tersebut dapat disimpan dalam penyimpanan informasi, database dan data warehouse. Data Mining ini berkaitan dengan ilmu-ilmu yang lain, seperti Data Warehousing, Statistik, dan machine learning. selain itu data mining juga didukung dengan spatial data analysis, image database serta signal processing [13].

2.2.2 Metode *K-Nearest Neighbour (K-NN)*

Metode K-NN merupakan suatu metode dengan pembelajaran yang terawasi (supervised) dan banyak digunakan untuk memprediksi atau klasifikasi. Prinsip kerja dari algoritma K-Nearest Neighbour adalah dengan cara mencari jarak terdekat diantara data yang akan dievaluasi dengan K

tetangga (Neighbor) dalam data pelatihan algoritma ini bertujuan untuk mengklasifikasikan objek baru berdasarkan atribut data dan training sample. Kelebihan dari Metode K-Nearest Neighbour ini adalah memiliki akurasi tinggi, tidak ada asumsi tentang data serta intensif pada outlier, hal utama yang perlu dilakukan adalah dengan cara menentukan nilai K [14]. Kesamaan data dengan label ini digunakan jarak Euclidean distance dengan rumus :

$$d_i = \sqrt{\sum_{n=1}^p (x_{2i} - x_{1i})}$$

Dengan rincian sebagaimana berikut :

d_i	: Jarak
x_i	: Dataset pelatihan
y_i	: Dataset testing
i	: Dataset variabel
n	: Dataset dimensi

2.2.3 Klasifikasi

Klasifikasi merupakan teknik untuk mengklasifikasikan suatu data tertentu ke dalam kelas yang sebelumnya telah ditentukan kelasnya pada pola atau atribut yang terdapat pada data tersebut. Klasifikasi ini menjadi lebih populer dikarenakan dapat menangani rentang data yang lebih luas dibandingkan dengan regresi [15]. Tujuan utama dari klasifikasi ini sendiri yakni untuk memprediksi kelas atau label dari data yang belum dikenal berdasarkan data yang telah ada dari data yang sebelumnya telah diketahui..

2.2.4 Penyakit Jantung

Jika terdapat pasien yang penderita penyakit jantung memerlukan penanganan yang tepat dan baik, karena penyakit ini termasuk salah satu penyakit yang paling berbahaya dan meskipun tidak menular, penyakit jantung

termasuk dalam penyakit yang menyebabkan kematian nomer satu di dunia [16].

Penyakit jantung merupakan salah satu penyakit Kardivaskular dan merupakan suatu penyakit degenerative yang berkaitan dengan gaya hidup dan kebiasaan sehari-hari.,Data dari Badan Kesehatan Dunia atau World Health Organization (WHO) menyatakan bahwa lebih dari 7 juta penduduk dunia meninggal disebabkan oleh penyakit jantung [17]. meskipun penyakit ini tidak menular, penyakit jantung merupakan salah satu jenis penyakit yang sangat mematikan di dunia.

Penyakit ini disebut juga dengan penyakit jantung koroner. penyakit jantung ini terjadi apabila darah yang masuk ke otot jantung terhenti/tersumbat, akibat dari kondisi ini adalah mengakibatkan kerusakan berat pada jantung [6]. Penyebab utama dari Penyakit Jantung ini adalah Fisik yang tidak aktif, diet tidak sehat, Penggunaan tembakau berlebih, alkohol, resiko penyakit jantung bertambah dengan faktor meningkatnya usia, tekanan darah tinggi kelebihan berat badan serta mempunyai kolestrol yang tinggi [4].

2.3 Perangkat Lunak yang Digunakan

Berikut adalah beberapa perangkat lunak atau software yang akan dipakai dalam pembuatan aplikasi dan identifikasi data sehingga dapat terealisasi dengan baik dan mendapatkan hasil yang optimal.

a. RapidMiner Studio

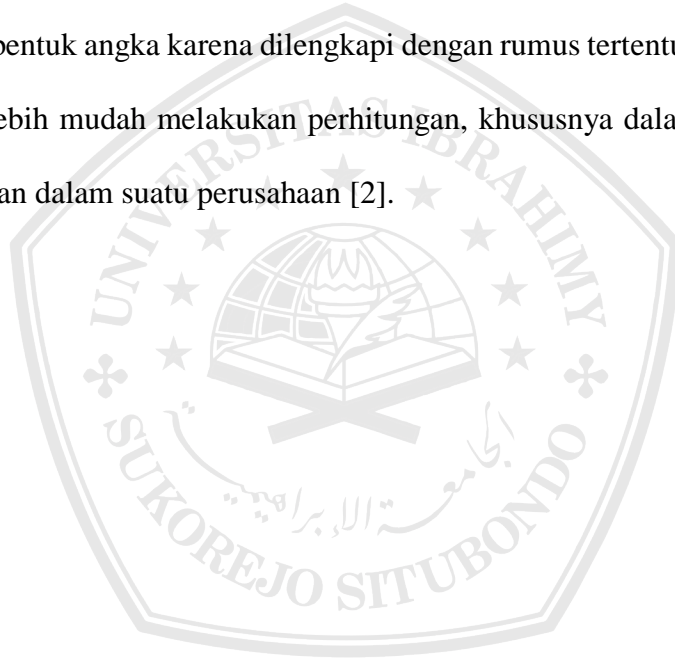
Pertama kali dikembangkan, RapidMiner bernama YALE (Yet Another Learning Environment) yang dikembangkan pada tahun 2001 oleh RalfKlinkenberg, Ingo Mierswa, dan Simon Fischer di Artificial Intelligence Unit dari University of Dortmund. RapidMiner sendiri merupakan perangkat lunak yang bersifat *open source* sehingga dapat dikembangkan oleh seluruh pengguna tanpa membayar royalti kepada penciptanya.

RapidMiner memiliki kurang lebih 500 operator *data preprocessing* dan visualisasi. RapidMiner ditulis dengan menggunakan Bahasa pemrograman Java sehingga bisa diaplikasikan di seluruh sistem operasi. RapidMiner menyediakan GUI (*Graphic User Interface*) untuk merancang sebuah pipeline analitis. GUI ini akan menghasilkan file XML (*Extensible Markup Language*) yang mendefinisikan proses analitis keinginan pengguna untuk diterapkan ke data. File tersebut kemudian dibaca oleh RapidMiner dan dianalisa secara otomatis. RapidMiner menempati peringkat pertama sebagai software data mining pada polling oleh KDnuggets, sebuah portal data mining pada tahun 2010-2011 [1] .

b. Microsoft Excel 2016

Microsoft Excel merupakan salah satu produk software yang diciptakan oleh serangkaian Microsoft Office bersama dengan Microsoft Word, Microsoft PowerPoint, dan lainnya. Perkembangan terakhir dari Microsoft Office ini adalah Microsoft Office 365.

Dalam pembuatan data mining ini aplikasi pertama yang digunakan dalam perekaman data adalah Microsoft Excel. Dan tipe yang peneliti gunakan adalah Microsoft Excel 2016. Microsoft Excel memiliki fungsi untuk pengolahan angka berbasis spreadsheet yang berisi baris dan kolom dalam membuat data yang diinginkan oleh user. Microsoft Excel ini termasuk dalam aplikasi yang banyak digunakan dalam membuat sebuah data di dunia karena kemudahannya dalam melakukan pengolahan data dalam bentuk angka karena dilengkapi dengan rumus tertentu sehingga user dapat lebih mudah melakukan perhitungan, khususnya dalam perhitungan keuangan dalam suatu perusahaan [2].



BAB III

METODELOGI PENELITIAN

3.1 Metode Penelitian

Metodologi penelitian merupakan cara sistematis untuk mengumpulkan data dengan tujuan tertentu. Metode penelitian ini merupakan suatu upaya atau prosedur untuk menemukan solusi terhadap suatu subjek atau masalah secara hati-hati, terencana, metodis, atau ilmiah, hal tersebut dilakukan dengan tujuan untuk menemukan fakta – fakta atau langkah-langkah yang ada. metode penelitian ini berguna sebagai fungsi signifikan untuk mendapatkan data data ataupun informasi yang dibutuhkan untuk menyelesaikan berbagai macam masalah dan memberikan suatu solusi atas masalah yang ada [18].

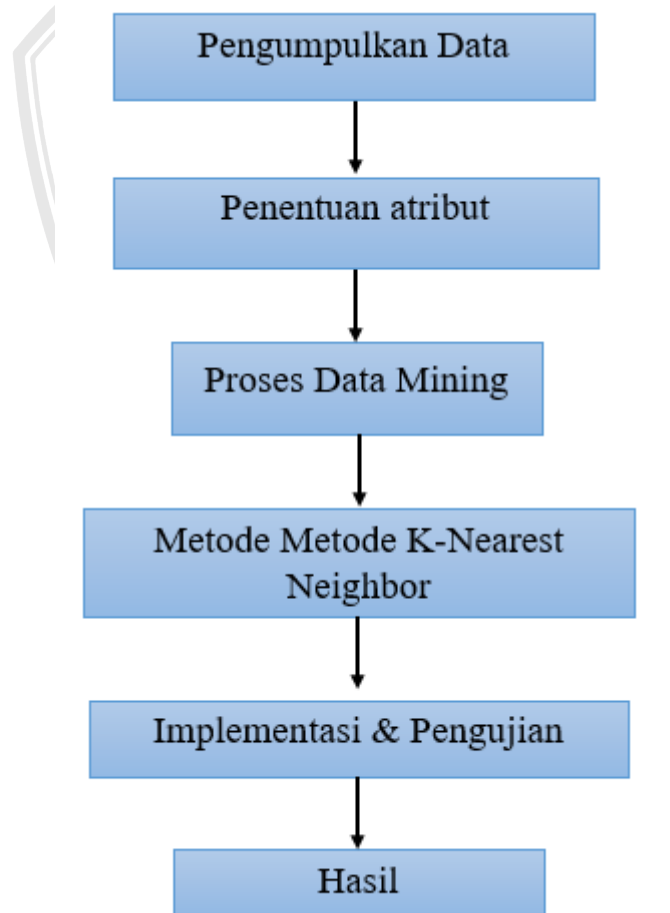
Metode penelitian menjadi lebih bagus jika metodologi penelitian tersebut disesuaikan dengan objek/subjek penelitian, keberhasilan atau kegagalan proyek penelitian bergantung pada pilihan metodologi penelitian, jadi penting untuk menggunakan tahap ini. Jika peneliti ingin karya ilmiah mereka memecahkan teka – teki dan mengungkap kebenaran, mereka harus akurat dalam menggunakan metodologi penelitian.

Data Prediksi Penyakit Jantung digunakan dalam penelitian ini dengan pendekatan kuantitatif dan juga studi perpustakaan secara online. Data yang digunakan merupakan data yang berasal dari website Kaggle. Pendekatan klasifikasi dengan Metode K-Nearest Neighbors akan digunakan untuk mengolah data tersebut.

3.2 Tahapan Penelitian

Pada bagian tahapan penelitian ini memaparkan tentang proses bisnis sebelum melakukan pemodelan data terkait. pada tahap pertama yakni Persiapan Data (prepossessing data) persiapan ini meliputi kegiatan-kegiatan seperti pemilihan data, transformasi data, pembersihan data dan validasi data. Langkah langkah tersebut dilakukan untuk memilih data yang relevan, dan membuang data yang tidak sesuai.

Terdapat langkah-langkah atau tahapan yang dilakukan dalam pelaksanaan penelitian ini, tahapan-tahapan tersebut sebagaimana pada gambar di bawah ini



Gambar 3. 1 Tahapan Penelitian

3.2.1 Pengumpulan Data

Data diperoleh dari repository online pada aplikasi Kaggle yakni KLASIFIKASI PENYAKIT JANTUNG dataset. link yang dapat diakses <https://www.kaggle.com/code/dirghalimsusilo/klasifikasi-penyakit-jantung>.

Dataset ini diperoleh dari database Cleveland dari UCI Machine Learning Repository. Data ini dihimpun dengan 14 atribut atau fitur berkenaan dengan apakah seseorang memiliki penyakit jantung atau tidak.

3.2.2 Penentuan Atribut

Berdasarkan Output data diatas, kami memiliki 303 entri atau baris dengan 14 atribut. terdapat 1 data yang kosong pada atribut umur, sehingga data tersebut akan diproses pada preprocessing data .

Atribut yang digunakan dalam memprediksi tingkat Kesehatan mental siswa di pelajaran matematika adalah sebagai berikut:

a. Age

Umur pasien yang ada pada survei ini, terdapat tingkatan umur pasien yang terdapat pada survei ini, yaitu :

Tabel 3.1
Atribut Umur

No	Umur	Jumlah
1	37 Tahun	3 Orang
2	38 Tahun	3 Orang
3	39 Tahun	4 Orang
4	40 Tahun	3 Orang
5	41 Tahun	10 Orang
6	42 Tahun	8 Orang

Tabel 3.1
(Lanjutan)

7	43 Tahun	8 Orang
8	44 Tahun	11 Orang
9	45 Tahun	8 Orang
10	46 Tahun	7 Orang
11	47 Tahun	5 Orang
12	48 Tahun	7 Orang
13	49 Tahun	5 Orang
14	50 Tahun	7 Orang
15	51 Tahun	12 Orang
16	52 Tahun	13 Orang
17	53 Tahun	8 Orang
18	54 Tahun	16 Orang
19	55 Tahun	8 Orang
20	56 Tahun	11 Orang
21	57 Tahun	17 Orang
22	58 Tahun	19 Orang
23	59 Tahun	14 Orang
24	60 Tahun	11 Orang
25	61 Tahun	8 Orang
26	62 Tahun	11 Orang
27	63 Tahun	9 Orang
28	64 Tahun	10 Orang
29	65 Tahun	8 Orang
30	66 Tahun	7 Orang
31	67 Tahun	9 Orang
32	68 Tahun	4 Orang
33	69 Tahun	3 Orang
34	70 Tahun	4 Orang

Tabel 3.1
(Lanjutan)

35	74 Tahun	1 Orang
36	76 Tahun	1 Orang
37	77 Tahun	1 Orang

b. Sex

Atribut ini menjelaskan tentang jenis kelamin pasien yang dibedakan 1 = Laki-Laki; 0 = Perempuan:

Tabel 3.2
Atribut Jenis Kelamin

Kode	Jenis Kelamin	Jumlah
1	<i>Laki-Laki</i>	207 Pasien
0	<i>Perempuan</i>	96 Pasien

c. CP

Chest Pain atau yang biasa dikenal dengan Atribut CP ini menjelaskan tentang tipe nyeri dada yang dialami oleh pasien, 0 = nyeri dada terkait mengurangi suplai darah ke jantung; 1 = nyeri dada tidak berhubungan dengan jantung; 2 = biasanya kejang esofagus (tidak berhubungan dengan jantung) 3 = nyeri dada tidak menunjukkan tanda-tanda penyakit :

Tabel 3.3
Atribut Jenis Nyeri Dada

Kode	Jenis Nyeri Dada	Jumlah penderita
0	Nyeri dada terkait mengurangi suplai darah ke jantung	51

**Tabel 3.3
(Lanjutan)**

1	nyeri dada tidak berhubungan dengan jantung	50
2	Kejang Esofagus (tidak berhubungan dengan jantung)	87
3	Nyeri dada tidak menunjukkan tanda-tanda penyakit	23

d. Trestbps

Resting Blood Pressure merupakan suatu istilah medis yang mengacu pada tekanan darah seseorang saat beristirahat. Biasanya, ini diukur dalam milimeter merkuri (mmHg) yang terdiri dari dua angka yakni tekanan tekanan sistolik (tekanan darah saat jantung berdetak) serta tekanan diastolik (tekanan darah saat jantung beristirahat di antara detak). TRESTBPS sering digunakan sebagai salah satu parameter dalam evaluasi risiko terkena penyakit kardiovaskular dan merupakan bagian penting dari pemeriksaan kesehatan rutin. Memantau tekanan darah istirahat dapat membantu dalam mendeteksi masalah jantung atau hipertensi sejak dini. Atribut ini menjelaskan tentang tekanan darah yang ada pada diri pasien, biasanya darah di atas 130-140 biasanya memprihatinkan.

Tabel 3.4
Atribut tekanan darah

No	Tekanan Darah	Jumlah Pasien
1	Darah Dibawah 130	169 Pasien
2	Darah Diatas 140	134 Pasien

e. Chol

Chol merupakan singkatan dari Cholesterol, yang mengacu pada jumlah kolesterol yang berbeda di dalam darah. Kolesterol ini merupakan suatu zat lemak yang penting bagi tubuh karena berperan dalam pembentukan sel dan produksi hormon tertentu. Namun, kadar kolesterol yang terlalu tinggi dalam darah dapat meningkatkan risiko penyakit jantung dan pembuluh darah, seperti serangan jantung dan stroke. Kolesterol dalam darah ini biasanya diukur dalam miligram per desiliter (mg/dL) dan dibagi menjadi 3 jenis, yakni Low-Density Lipoprotein (LDL), High-Density Lipoprotein (HDL) serta Triglycerides. Atribut ini merupakan atribut Kolesterol serum, atribut Kolesterol Serum ini dapat memberikan gambaran terhadap kadar Kolesterol seseorang biasanya diatas 200, memprihatinkan

Tabel 3.5
Atribut Tensi Darah

No	Tensi Darah	Jumlah Pasien
1	Serum cholestoral Dibawah 200	49 Pasien
2	Serum cholestoral Diatas 200	254 Pasien

f. Fbs

FBS adalah singkatan dari Fasting Blood Sugar, yang merupakan kadar gula darah yang diukur setelah seseorang berpuasa selama 8-12 jam. Tes ini digunakan untuk dapat mengevaluasi kadar glukosa yang berada dalam darah dan dapat membantu dalam diagnosis diabetes atau kondisi pra-diabetes. Pada Atribut ini menjelaskan tentang Fasting Blood Sugar atau kadar gula darah puasa merupakan tes darah sederhana dan umum yang dilakukan untuk mendiagnosis pradiabetes, diabetes, atau diabetes gestasional. Sebelum tes, tidak diperbolehkan makan selama 8 sampai 12 jam. biasanya tes lebih dari 126' mg/dL disinyalir menderita diabetes. Dijelaskan bahwa 1= True; 0 = False.

Tabel 3.6
Atribut Kadar Gula Darah

No	Kadar Gula darah	Jumlah Pasien
1	True	45 Pasien
2	False	258 Pasien

g. Restecg

Pada Atribut ini menjelaskan tentang pemeriksaan EKG yang dilakukan pada saat pasien dalam kondisi beristirahat (dalam posisi berbaring). Hasil dapat dikatakan normal berarti tidak terdapat kelainan aktivitas jantung pada saat pasien istirahat tersebut. pada dataset ini jika 0 maka Tidak ada yang perlu diperhatikan; jika 1 dapat berkisar dari gejala ringan hingga masalah parah dan sinyal detak jantung yang tidak normal dan jika 2 Ruang pompa utama jantung yang diperbesar.

Tabel 3.7
Atribut Reesting electrocardiographic results

No	Kode	Resting electrocardiographic results	Jumlah Pasien
1	0	Tidak ada yang perlu diperhatikan	147 Pasien
2	1	ST-T Wave abnormality	152 Pasien
3	2	Possible or Definite left ventricular hypertrophy	4 Pasien

h. Thalach

Atribut Pada Atribut ini menjelaskan tentang Thalach yang merupakan denyut jantung maksimal yang terjadi pada pasien

i. Exang

Atribut ini menjelaskan tentang Exang yang merupakan keadaan dimana pasien akan mengalami nyeri dada apabila beraktivitas atau berolah raga, 0 jika tidak nyeri, dan 1 jika menyebabkan nyeri.

Tabel 3.8
Keadaan setelah Olahraga

No	Kode	Keadaan setelah Olahraga	Jumlah Pasien
1	0	Tidak Nyeri	204 Pasien
2	1	Nyeri	99 Pasien

j. Oldpeak

Oldpeak merupakan suatu istilah yang digunakan dalam konteks analisis data medis, khususnya dalam pengukuran stres jantung. Oldpeak ini merujuk pada penurunan segmen ST dalam tes stres jantung, yang diukur dalam grafik elektrokardiogram (EKG). Penurunan segmen ST dapat menunjukkan adanya gangguan aliran darah ke jantung, seperti yang terjadi pada angina atau serangan jantung. Oldpeak digunakan sebagai salah satu indikator untuk menilai kesehatan jantung dan risiko penyakit jantung. Oldpeak ini merupakan depresi ST yang diakibatkan oleh latihan relative pad saat istirahat. hal tersebut bisa dengan cara melihat stres jantung saat berolahraga dan jantung yang tidak sehat akan lebih stres.

k. Slope

Slope dalam analisis dunia medis, khususnya dalam tes stres jantung, merujuk pada kemiringan segmen ST pada grafik elektrokardiogram (EKG) selama uji beban. Segmen ST ini menunjukkan perubahan aktivitas listrik jantung setelah terjadinya kontraksi, dan kemiringannya dapat memberikan informasi penting

tentang kesehatan jantung. Slope positif, di mana segmen ST naik atau tetap datar selama tes, biasanya dianggap normal dan menunjukkan bahwa jantung berfungsi dengan baik. Sebaliknya, slope negatif, yang menunjukkan penurunan segmen ST, dapat mengindikasikan adanya gangguan aliran darah ke jantung, seperti angina atau penyakit jantung koroner. Slope datarnya, di mana segmen ST tetap pada level yang sama, juga bisa menunjukkan masalah, meskipun tidak sebanyak penurunan. Dengan demikian, penilaian slope membantu dokter dalam mendiagnosis dan mengevaluasi risiko penyakit jantung. Slope dari puncak ST setelah berolahraga. Memiliki tiga nilai antara lain, diantaranya Upsloping (0) detak jantung yang lebih baik dengan olahraga (tidak biasa); Flatsloping (1) perubahan minimal (jantung sehat yang khas); Downsloping (2) yang menunjukkan tanda-tanda jantung yang tidak sehat.

Tabel 3.9
Slope dari puncak ST setelah berolahraga

No	Kode	Slope dari puncak ST setelah berolahraga	Jumlah Pasien
1	0	Upsloping	21 Pasien
2	1	Flatsloping	140 Pasien
2	2	Downsloping	142 Pasien

1. *Ca*

Atribut CT Calcium Score merupakan pemeriksaan kadar kalsium dalam pembuluh darah koroner melalui perangkat CT scan.

Pemeriksaan kadar kalsium di pembuluh darah dapat membantu dokter mendeteksi penumpukan zat berlemak atau plak yang dapat menyumbat aliran darah. jumlah pembuluh darah utama (0-3) diwarnai oleh flourosopy pembuluh berwarna berarti dokter dapat melihat darah yang melewatinya dan semakin banyak gerakan darah semakin baik (tidak ada gumpalan).

Tabel 3.10
Pembuluh Darah Utama

No	Pembuluh Darah Utama	Jumlah Pasien
1	0	175 Pasien
2	1	65 Pasien
3	2	38 Pasien
4	3	20 Pasien
5	4	5 Pasien

m. Thal

Thal merupakan singkatan dari thalassemia, yaitu suatu kondisi medis yang dapat mempengaruhi produksi hemoglobin yang berada di dalam darah. Hemoglobin ini merupakan protein dalam sel darah merah yang membawa oksigen ke seluruh tubuh. Thalassemia merupakan jenis dari anemia genetik yang ditandai oleh produksi hemoglobin yang tidak normal, serta dapat menyebabkan anemia berat dan masalah kesehatan

lainnya. Dalam konteks data medis dan analisis kesehatan, Thal biasanya merujuk pada indikator atau kategori yang menunjukkan adanya thalassemia pada pasien. Misalnya, dalam dataset medis, Thal bisa digunakan sebagai variabel yang menunjukkan hasil pemeriksaan untuk thalassemia, dengan nilai yang mungkin menunjukkan apakah pasien memiliki kondisi tersebut atau tidak, dan tingkat keparahannya. Tes Stres Thallium ini merupakan tes pencitraan yang bertujuan untuk menunjukkan seberapa baik darah yang mengalir ke dalam jantung saat sedang istirahat ataupun sedang berolahraga, 1,3 menunjukkan normal; 6 menunjukkan bahwa fixed defect artinya dulu cacat tapi sekarang baik-baik saja dan; 7 menunjukkan reversable defect yang artinya tidak ada gerakan darah yang tepat saat berolahraga.

Tabel 3.11
Tes Stres Thallium

No	Tes Stres Thallium	Jumlah Pasien
1	1 Sampai 3	303 Pasien
2	6	0 Pasien
3	7	0 Pasien

n. Target

Atribut ini menjelaskan apakah pasien membutuhkan perawatan khusus tentang memiliki penyakit jantung atau tidak, sehingga dengan hal tersebut peneliti dapat memprediksi seberapa besar Pasien yang

membutuhkan penanganan khusus dan tidak. Diklasifikasikan menjadi 2 class yaitu 1 = Yes; dan 0 = No

Tabel 3.12

Atribut Prediksi Penyakit Jantung

No	Simbol	Memiliki Penyakit jantung	Jumlah Pasien
1	1	Yes	165 Pasien
2	0	No	138 Pasien

3.2.3 Proses *Data Mining*

Nama lain dari data mining adalah sebuah disiplin ilmu yang memiliki tujuan untuk menemukan, mengekstrak, atau menambang pengetahuan dari data dan informasi yang sudah kita miliki sebelumnya. Proses ini disebut juga dengan proses penemuan pengetahuan. Yang dalam Bahasa Inggris disebut dengan *Knowledge Discovery in Database* (KDD). proses-proses dalam data mining ini melibatkan data dalam jumlah yang cukup besar. banyak metode yang dapat digunakan dalam data mining, diantaranya adalah Metode Pohon Keputusan (Decision Trees) K-Means, Metode Algoritma Apriori, Metode Naive Bayes, Metode Support Vector Machines (SVM) serta Metode K-Nearest Neighbors (k-NN). Proses data mining ini yang di gunakan untuk menggali informasi dari data dan menemukan model-model baru yang tersimpan dari data. model baru tersebut yang digunakan untuk memahami suatu data ataupun informasi yang terdapat pada sekumpulan data yang besar tersebut. Data mining pun dapat digunakan untuk

berbagai bidang ilmu yang menggunakan pengolahan data di dalamnya, seperti Statistik, database dan kesehatan.

Proses-proses ini melakukan ekstraksi informasi dari data yang sangat besar yang direpresentasikan dengan menggunakan Minimal Microsoft Excel sebagai toolnya. Microsoft Excel merupakan perangkat lunak yang termasuk dalam Microsoft Office yang digunakan untuk mengolah data yang bersifat Angka (angka) dan biasanya digunakan dalam proses data mining.

Adapun langkah-langkah yang digunakan dalam dalam memproses data mining sebagai berikut:

a. *Cleaning Data*

langkah pertama yang dilakukan dalam proses Knowledge Discovery in Database adalah dengan pembersihan data. pembersihan data ini merupakan suatu proses untuk menghilangkan noise dari data untuk data yang tidak konsisten ataupun hilang dan untuk menghilangkan nilai yang lebih. proses cleaning data ini dilakukan untuk mengumpulkan atribut yang di nilai penting dan menghapus atribut yang tidak relevan dengan hasil yang di inginkan

b. *Data Selection*

Langkah kedua ini dilakukan untuk memilih atribut-atribut yang di perlukan serta membuang data yang tidak diperlukan. dalam dataset Klasifikasi Penyakit jantung ini terdapat 14 atribut. atribut-atribut tersebut diantaranya adalah *Age, Sex, Cp, Trestbps, Chol, FBS (Fasting Blood Sugar)Restecg Resting 12-lead electrocardiography (ECG),*

Thalach, Exang, Oldpeak, Skope, Ca (CT Calcium Score), Thal (Tes Stres Thallium), dan Target.

c. Transformasi Data

Langkah selanjutnya adalah transformasi data, transformasi data ini adalah dengan mengubah beberapa kata yang terdapat penyingkatan dan penulisan yang kurang dimengerti. pada data klasifikasi Penyakit Jantung ini atribut-atribut yang digunakan merupakan atribut-atribut dan terkait dengan penyakit jantung seperti Cp, Trestbps, Chol fbs, Ca dan thal sehingga tidak perlu untuk di ubah.

d. Data Mining

Data yang besar disimpan dalam penyimpanan, dan pada data mining merupakan tindakan untuk mengidentifikasi hubungan, pola dan tren baru yang relevan dengan menggunakan algoritma pengenalan pola termasuk dalam prosedur statistik dan matematika. pada proses KDD ini, secara keseluruhan dan substansial dipengaruhi oleh bagaimana algoritma atau metode yang tepat dapat dipilih untuk prediksi yang diinginkan.

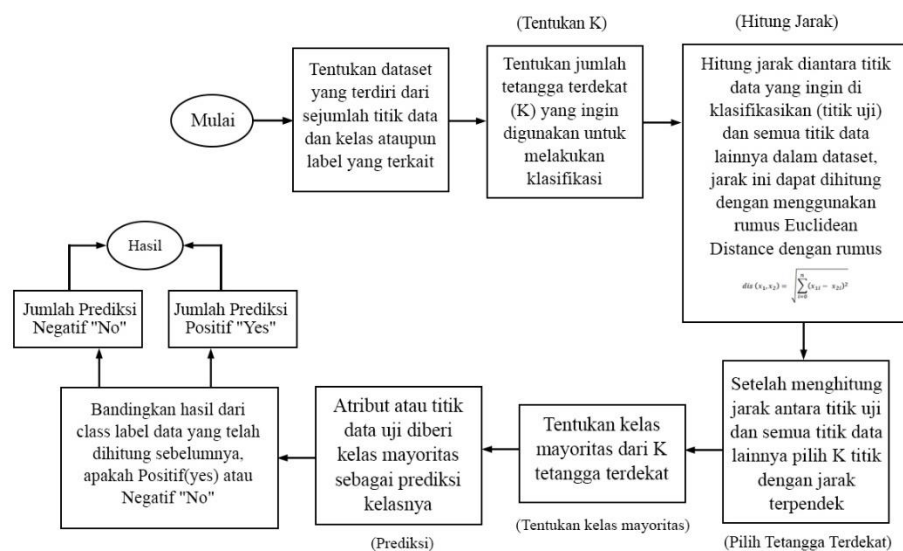
Pada penelitian ini, setelah segala proses persiapan data dilakukan dan data siap untuk diproses, maka kami memilih metode K-Nearest Neighbors (K-NN) untuk memprediksi pasien yang terkena penyakit jantung atau tidak sehingga dapat dilakukan penanganan lebih oleh pihak terkait. atribut yang menjadi label prediksi pada klasifikasi penyakit jantung ini adalah "target" yang menghasilkan nilai 0 bagi

pasien yang tidak terkena penyakit jantung dan menghasilkan nilai 1 untuk pasien yang terkena penyakit jantung dan membutuhkan penanganan lebih.

3.3 Perancangan Sistem

3.3.1 Metode K-Nearest Neighbors (K-NN)

Algoritma K-Nearest Neighbors (k-NN) merupakan suatu metode yang menggunakan algoritma supervised dimana hasil dari query instance yang baru diklasifikasikan berdasar pada mayoritas dari kategori pada



Gambar 3. 2 Perhitungan Manual K-Nearest Neighbors(KNN)

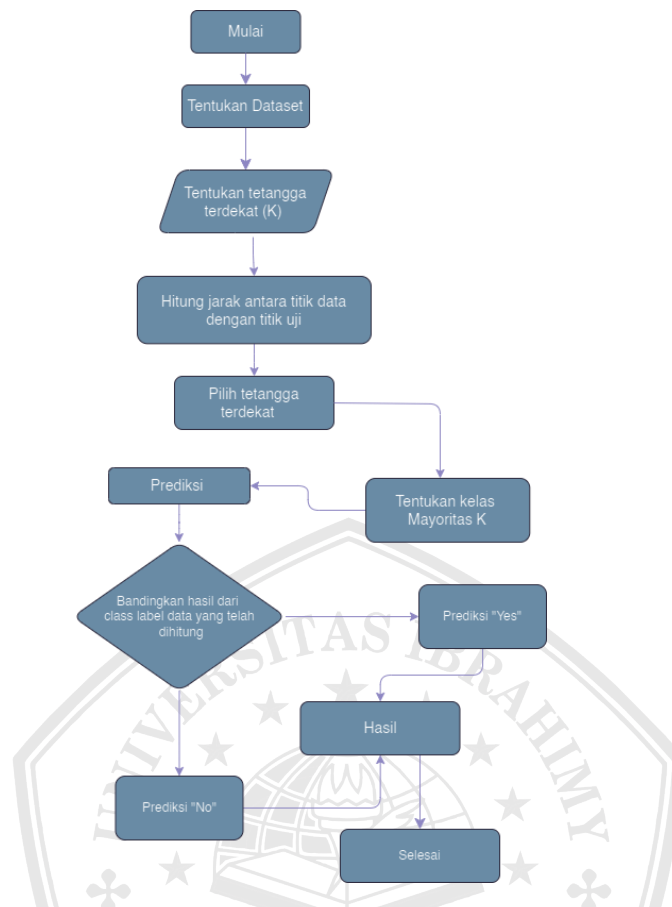
KNN. Prinsip kerja dari metode algoritma ini ialah dengan cara mencari jarak terdekat anatar data yang akan di evaluasi dengan k tetangga (Neighbors) dalam data pelatihan. tujuan dari algoritma ini ialah untuk mengklasifikasikan suatu objek baru berdasarkan atribut dan training sample. algoritma K-Nearest Neighbors memiliki beberapa kelebihan, diantaranya adalah ketangguhan terhadap training data yang memiliki

banyak noise dan efektif apabila training datanya 8 besar. untuk kelemahan dari KNN ini adalah perlu untuk menentukan nilai dari parameter K yang merupakan jumlah dari tetangga terdekat.[5] Berikut ini merupakan alur perhitungan manual dari algoritma K-Nearest Neighbors.

Tentukan dataset yang terdiri dari sejumlah titik data dan kelas ataupun label yang terkait, lalu Setelah ditentukan inputkan dataset dan Ketika dataset telah di inputkan, maka Tentukan jumlah tetangga terdekat (K) yang ingin digunakan untuk melakukan klasifikasi, setelah itu Hitung jarak diantara titik data yang ingin di klasifikasikan (titik uji) dan semua titik data lainnya dalam dataset, jarak ini dapat dihitung dengan menggunakan rumus Euclidean Distance dengan rumus :

$$d_i = \sqrt{\sum_{n=1}^p (x_{2i} - x_{1i})^2}$$

d merupakan jarak, X_{1i} adalah sample data dan X_{2i} merupakan data testing, i adalah symbol dari variable data dan n merupakan dimensi data. setelah itu hitunglah jarak antara titik uji dan semua titik data lainnya pilih K titik dengan jarak terpendek kemudian tentukan tentukan kelas mayoritas dari K tetangga terdekat, kemudian Atribut atau titik data uji diberi kelas mayoritas sebagai prediksi kelasnya dan bandingkan Bandingkan hasil dari class label data yang telah dihitung sebelumnya, apakah Positif (yes) atau Negatif "No" kemudian hasil yang terbesar dari perbandingan kedua label tersebut merupakan hasil dari prediksi metode K-Nearest Neighbors.

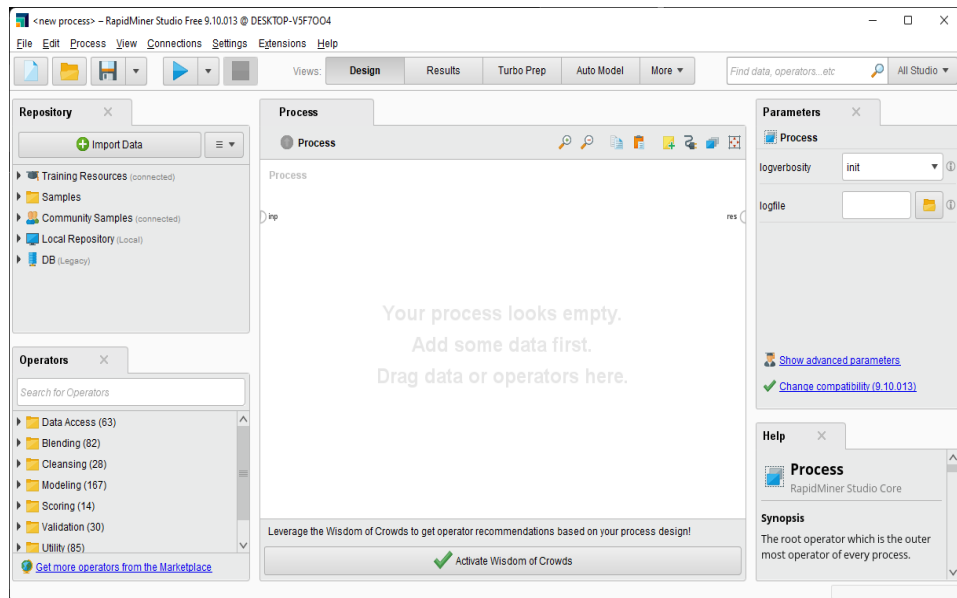


Gambar 3.3 *Flowchart Perhitungan K-Nearest Neighbors (KNN)*

3.4 Implementasi dan Pengujian Metode

Pada penelitian kali ini dipilih aplikasi RapidMiner versi 9.10 untuk implementasi dan pengujian akurasi pada proyek kali ini, RapidMiner merupakan suatu perangkat lunak yang digunakan sebagai alat pembelajaran dalam ilmu data mining. Platform RapidMiner ini dikembangkan oleh perusahaan yang didedikasikan untuk semua langkah yang melibatkan data yang besar seperti, penelitian, pendidikan, pelatihan, dan pembelajaran. RapidMiner memiliki sekitar 100 solusi pembelajaran untuk pengelompokan, klasifikasi dan analisis regresi. RapidMiner juga mendukung sekitar 22 format file, seperti .xls, .csv, dan format lainnya.

model model yang ditawarkan oleh RapidMiner yang berfokus pada Pengetahuan tentang Penemuan Data (*Knowledge of Data Discovery*) yang mencakup metode metode seperti Metode Pohon Keputusan (*Decision Trees*) K-Means, Metode , Metode Support Vector Machines (SVM) serta Metode K-Nearest Neighbors (k-NN). berikut ini merupakan user interface pada aplikasi RapidMiner versi 9.10

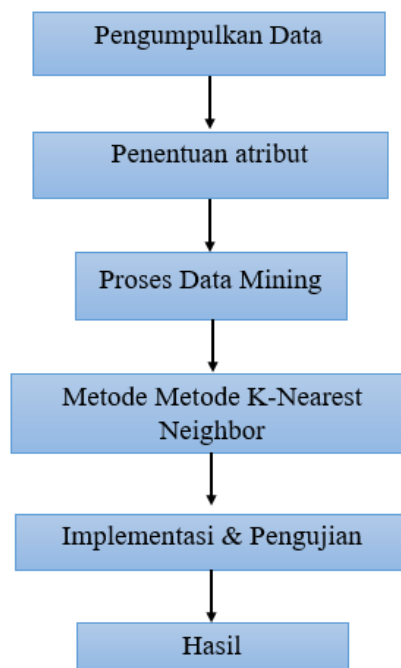


Gambar 3. 4 User Interface RapidMiner versi 9.10

Tahapan – tahapan yang akan dilakukan dalam implementasi *RapidMiner* 9.10 ini adalah sebagaimana berikut:

Tahap pertama adalah import data dalam format excel yang telah disiapkan sebelumnya. Hal ini dilakukan dengan fitur *Read Excel* dalam *RapidMiner 9.10*. Apabila terdapat data yang kosong, maka digunakanlah *tool Replace Missing Value* untuk menggantikan nilai yang hilang dengan nilai yang sesuai untuk mengisi nilai yang kosong tersebut. Kemudian pemilihan data *training* dan data *testing* dilakukan dengan menggunakan *tool split data* yang ada pada *RapidMiner*. Kemudian pelabelan dan penentuan tipe data dilakukan sebelum data siap untuk dihitung dengan algoritma *K-Nearest Neighbors* pada *RapidMiner 9.10*.

Kemudian jika terdapat data dengan penulisan yang beragam, akan disamakan dalam penulisannya, . Kemudian setelah data berhasil melalui proses



Gambar 3. 5 Kerangka Pemikiran pada *RapidMiner 9.10*

sebelum perhitungan (*preprocessing data*) kemudian dihitunglah menggunakan algoritma *K-Nearest Neighbors* dan juga proses Validasi dari algoritma tersebut.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Hasil Penelitian

Hasil penelitian yang telah dilakukan dalam mengklasifikasi penyakit jantung dengan menggunakan metode K-Nearest Neighbors (KNN). Dalam penelitian ini diperoleh sebanyak 14 atribut dengan 303 value data dari repositori online Kaggle.

4.2 Perhitungan Manual Metode K-Nearest Neighbors

4.2.1 Tentukan Nilai K

Langkah yang pertama kali dilakukan adalah dengan menentukan jumlah tetangga terdekat yang akan digunakan untuk data yang akan kita prediksi. Pada prediksi ini kita akan menggunakan nilai $K = 5$

4.2.2. Hitung jarak

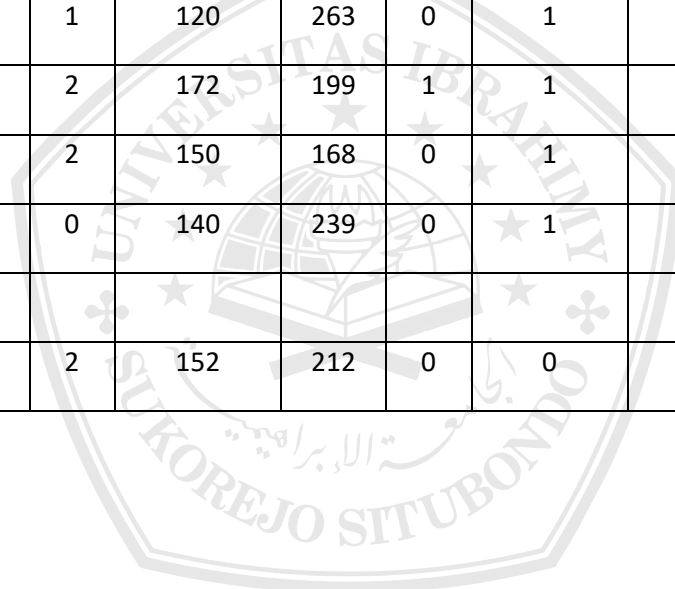
Langkah selanjutnya ialah mengukur jarak diantara data baru dengan data setiap data yang ada dalam dataset dengan menggunakan rumus Euclidean Distance

$$d_i = \sqrt{\sum_{n=1}^p (x_{2i} - x_{1i})^2}$$

Berikut ini merupakan data Set dari klasifikasi penyakit jantung yang diperoleh dari database Cleveland dari UCI Machine Learning Repository :

Tabel 4.1
Tabel Dataset

age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
63	1	3	145	233	1	0	150	0	23	0	0	1	1
37	1	2	130	250	0	1	187	0	35	0	0	2	1
41	0	1	130	204	0	0	172	0	14	2	0	2	1
56	1	1	120	236	0	1	178	0	8	2	0	2	1
57	0	0	120	354	0	1	163	1	6	2	0	2	1
57	1	0	140	192	0	1	148	0	4	1	0	1	1
56	0	1	140	294	0	0	153	0	13	1	0	2	1
44	1	1	120	263	0	1	173	0	0	2	0	3	1
52	1	2	172	199	1	1	162	0	5	2	0	3	1
57	1	2	150	168	0	1	174	0	16	2	0	2	1
54	1	0	140	239	0	1	160	0	12	2	0	2	1
.....													
67	1	2	152	212	0	0	150	0	8	1	0	3	0

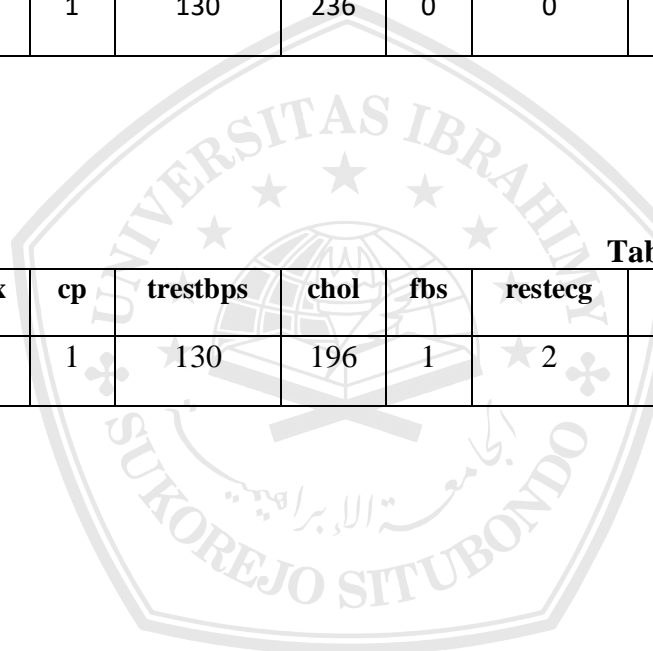


44	1	0	120	169	0	1	144	1	28	0	0	1	0
63	1	0	140	187	0	0	144	1	4	2	2	3	0
63	0	0	124	197	0	1	136	1	0	1	0	2	0
59	1	0	164	176	1	0	90	0	1	1	2	1	0
57	0	0	140	241	0	1	123	1	2	1	0	3	0
45	1	3	110	264	0	1	132	0	12	1	0	3	0
68	1	0	144	193	1	1	141	0	34	1	2	3	0
57	1	0	130	131	0	1	115	1	12	1	1	3	0
57	0	1	130	236	0	0	174	0	0	1	1	2	0

Data Test

Tabel 4.2
Tabel Data Test

age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
41	0	1	130	196	1	2	160	0	14	2	0	2	



Perhitungan :

Tabel 4.3
Tabel Perhitungan

Atribut				Hasil Perhitungan Euclidean Distance (41,0,1,130, ...,2)
X1 = Age	X2 = Sex	X13 = Thal	
63	1	1	47,66
37	1	2	64,11
41	0	2	14,59
56	1	2	47,83
57	0	2	159,36
57	1	1	24,9
56	0	2	99,92
44	1	3	70,48
52	1	3	44,53
57	1	2	40,55
54	1	2	46,11
67	1	3	39,51
44	1	1	36,03
63	1	3	32,16
63	0	2	36,03
59	1	1	83,42
57	0	3	62,45
45	1	3	76,4
68	1	3	41,28

57	1	3	80,73
57	0	2	47,49

4.2.3. Tentukan Tetangga Terdekat

Setelah selesai melakukan penghitung jarak ke semua data set, Langkah selanjutnya ialah dengan mengurutkan jarak dari yang terkecil ke terbesar dan pada penelitian kali ini, penulis K=5 tetangga terdekat. berdasarkan perhitungan diatas data yang memiliki positif memiliki penyakit jantung ditandai dengan :

Tabel 4.4
Tabel Tetangga Terdekat

Atribut				Hasil Perhitungan Euclidean Distance (41,0,1,130,2)	Jarak terkecil
X1 = Age	X2 = Sex	X13 = Thal		
57	0	2	159,36	1
56	0	2	99,92	2
59	1	1	83,42	3
45	1	3	76,4	5
57	1	3	80,73	4

4.2.4. Tentukan Kelas Prediksi

Langkah selanjutnya ialah dengan melihat pada label tetangga terdekat, dan setelah di cek menunjukkan bahwa pada data test :

Tabel 4.5
Tabel kelas Prediksi

Atribut				Hasil Perhitungan Euclidean Distance (41,0,1,130, ...,2)	Jarak terkecil	Target
X1 = Age	X2 = Sex	X13 = Thal			
57	0	2	159,36	1	1
56	0	2	99,92	2	1
59	1	1	83,42	3	0
45	1	3	76,4	5	0
57	1	3	80,73	4	0

Karena mayoritas dari 5 tetangga terdekat adalah 0, maka prediksi untuk data uji diatas adalah 0 atau *negative* atau tidak memiliki Penyakit Jantung

4.2.5 Menghitung Nilai *Accuracy*, *Sensitivity*, *Specificity*, *PPV* dan *NPV*

Berikut adalah rumus dari perhitungan akurasi, presisi, *recall* untuk mengetahui seberapa akurat perhitungan dari algoritma *K-Nearest Neighbors* yang telah kita lakukan sebelumnya, dengan menggunakan Matriks Konfusi (Confusion Matrix), matriks konfusi ini sendiri merupakan tabel yang digunakan untuk memahami bagaimana model klasifikasi (misalnya, model yang memprediksi apakah seseorang memiliki penyakit jantung atau tidak) telah bekerja. Ini menunjukkan berapa banyak prediksi model yang benar dan berapa banyak yang salah. Matriks ini memiliki empat komponen diantaranya adalah True Positive (TP) : Jumlah orang yang memang benar-benar sakit dan

model berhasil memprediksi pasien sebagai sakit. True Negative (TN): Jumlah orang yang memang benar-benar sehat dan model berhasil memprediksi pasien sebagai sehat. False Positive (FP): Jumlah orang yang sebenarnya sehat, tapi model salah memprediksi pasien sebagai sakit. False Negative (FN): Jumlah orang yang sebenarnya sakit, tapi model salah memprediksi pasien sebagai sehat.

1. Accuracy

Akurasi merupakan proporsi prediksi yang benar terhadap keseluruhan jumlah sampel yang ada. Ini mencakup baik prediksi positif maupun negatif yang benar. Berikut ini merupakan formula perhitungan dari Accuracy :

$$\text{Formula : } \frac{TP+TN}{TP+TN+FP+FN}$$

2. Sensitivity (Recall class "Yes")

Sensitivitas atau Recall ini mengukur kemampuan dari model untuk dapat mengidentifikasi semua sampel positif dengan benar. Ini adalah rasio antara True Positive dengan total jumlah sampel positif sebenarnya (TP + FN). Berikut ini merupakan formula perhitungan dari Sensitivity :

$$\text{Formula : } \frac{TP}{TP+FN}$$

3. Specificity (Recall class "No")

Specificity mengukur kemampuan dari model untuk dapat mengidentifikasi sampel negatif dengan benar. Ini adalah rasio antara True Negative dengan total jumlah sampel negatif

sebenarnya (TN + FP). Berikut ini merupakan formula perhitungan dari *Specificity* :

$$\text{Formula : } \frac{TN}{TN+FP}$$

4. PPV (*Positive Precision Vector for class "Yes"*)

PPV atau *Positive Precision Vector* juga dikenal sebagai Precision, mengukur proporsi sampel yang diprediksi positif yang benar-benar positif. Berikut ini merupakan formula perhitungan dari *Positive Precision Vector for class "Yes"*:

$$\text{Formula : } \frac{TP}{TP+FP}$$

5. NPV (*Negative Precision Vector for class "No"*)

NPV mengukur proporsi sampel yang diprediksi negatif yang benar-benar negatif. Berikut ini merupakan formula perhitungan dari *Negative Precision Vector for class "No"* :

$$\text{Formula : } \frac{TN}{TN+FN}$$

Berikut ini merupakan perhitungan dari Nilai Accuracy, Sensitivity, Specificity, PPV dan NPV :

a. Accuracy

$$\text{Formula : } \frac{TP+TN}{TP+TN+FP+FN} = \frac{124+68}{124+68+41+70} = \frac{192}{303} = 0,6403 = 64.03\%$$

b. Sensitivity (*Recall class "Yes"*)

$$\text{Formula : } \frac{TP}{TP+FN} = \frac{124}{124+70} = \frac{124}{194} = 0.5948 = 59.48\%$$

c. Specificity (*Recall class "No"*)

$$\text{Formula : } \frac{TN}{TN+FP} = \frac{68}{68+41} = \frac{59}{108} = 0.6391 = 63.91\%$$

d. PPV (Positive Precision Vector for class "Yes")

$$\text{Formula : } \frac{TP}{TP+FP} = \frac{124}{124+41} = \frac{124}{165} = 0.7515 = 75.15\%$$

e. NPV (Negative Precision Vector for class "No")

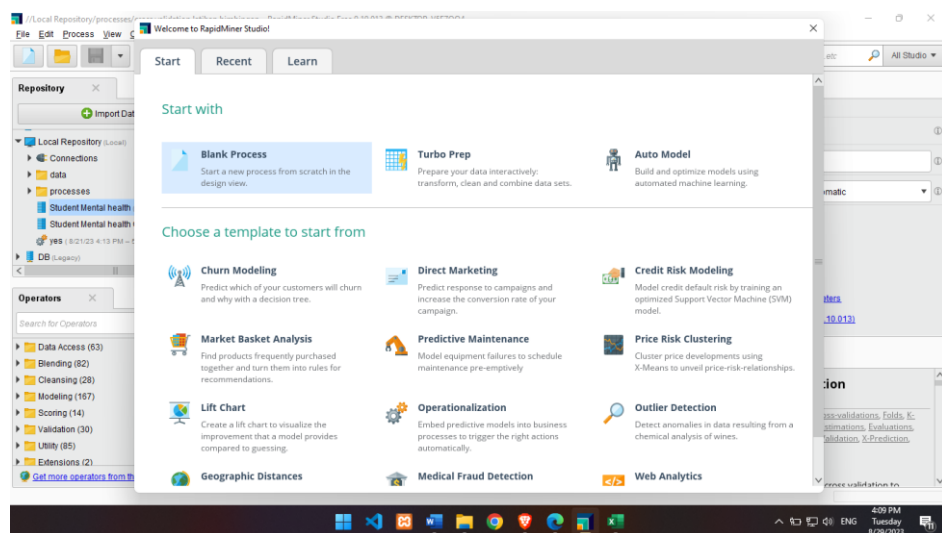
$$\text{Formula : } \frac{TN}{TN+FN} = \frac{68}{68+70} = \frac{68}{138} = 0.4927 = 49.27\%$$

4.3 Perhitungan RapidMiner

Berikut ini merupakan implementasi dari perhitungan klasifikasi penyakit jantung dengan menggunakan algoritma K-Nearest Neighbors. Berikut adalah Langkah – Langkah yang akan dilakukan dalam pengolahan data sampai hasil yang akan diperoleh dari prediksi tersebut.

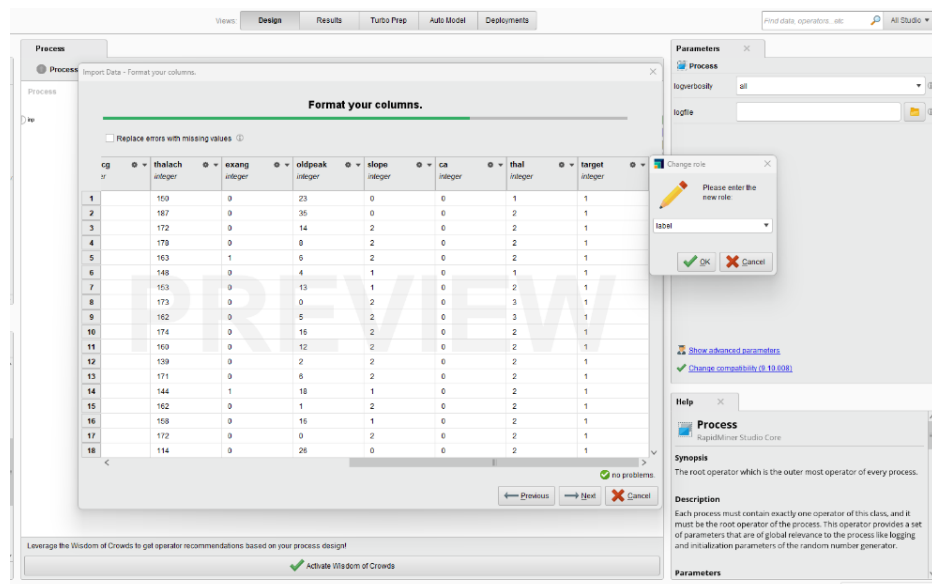
4.3.1 Pemrosesan Data

Langkah yang pertama kali dilakukan adalah dengan membuka aplikasi RapidMiner versi 9.10 dan berikut ini merupakan tampilan awal *RapidMiner* 9.10 klik *File*, kemudian *new process*, dan klik *blank process*



Gambar 4. 1 Proses Pembuatan Perhitungan Baru

Kemudian cari pada *tool operator* dan *drag tool Read Excel* untuk memasukkan data yang akan dihitung. Kemudian klik *Import Configuration Wizard* pada jendela kiri dan pilihlah data dalam bentuk excel yang akan dihitung. Kemudian dipilihlah kolom Excel dari atribut yang akan digunakan. Selanjutnya adalah memilih tipe data dari data



Gambar 4. 2 Penentuan jenis Data

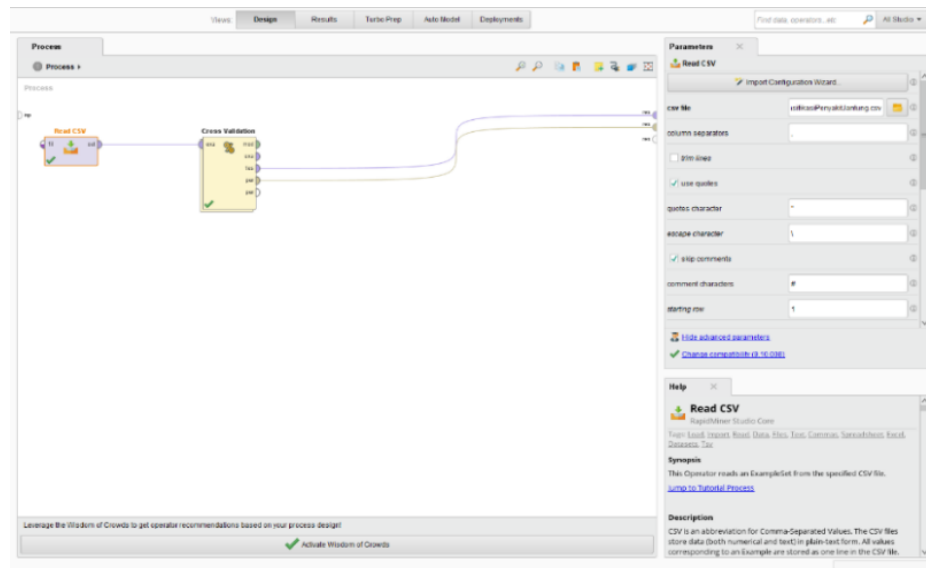
yang dimasukkan, apakah itu *integer*, *polynomial* atau *binominal*. Dan juga jangan lupa tentukan manakah atribut label dari data tersebut.

Kemudian Data yang mengalami missing value biasanya ditandai dengan (?) oleh system dan pada data diatas tidak terdapat data yang kosong atau mengalami missing value.

Row No.	target	h_ujro	sex	cp	trestps	chol	fbs	restngc	thalach	exang	oldpeak	slope
1	1	63	1	3	145	233	1	0	150	0	2.3	0
2	1	37	1	2	130	250	0	1	187	0	3.5	0
3	1	41	0	1	130	204	0	0	172	0	1.4	2
4	1	56	1	1	120	236	0	1	178	0	0.8	2
5	1	57	0	0	120	354	0	1	163	1	0.8	2
6	1	57	1	0	140	192	0	1	148	0	0.4	1
7	1	56	0	1	140	204	0	0	163	0	1.3	1
8	1	44	1	1	120	263	0	1	173	0	0	2
9	1	52	1	2	112	199	1	1	162	0	0.5	2
10	1	57	1	2	100	168	0	1	174	0	1.5	2
11	1	54	1	0	140	238	0	1	160	0	1.2	2
12	1	48	0	2	130	275	0	1	139	0	0.2	2
13	1	49	1	1	130	266	0	1	171	0	0.6	2
14	1	54	1	3	110	211	0	0	144	1	1.8	1
15	1	58	0	3	100	283	1	0	162	0	1	2
16	1	50	0	2	120	219	0	1	158	0	1.6	1
17	1	58	0	2	120	340	0	1	172	0	0	2
18	1	66	0	3	100	226	0	1	114	0	2.6	0
19	1	43	1	0	150	247	0	1	171	0	1.5	2
20	1	69	0	3	140	238	0	1	151	0	1.8	2
21	1	59	1	0	135	234	0	1	161	0	0.5	1
22	1	44	1	2	130	233	0	1	179	1	0.4	2
23	1	42	1	0	140	228	0	1	178	0	0	2
24	1	51	1	1	150	191	1	1	177	1	1	2
25	1	51	1	1	150	191	1	1	177	1	1	2

Gambar 4.3 Data Value

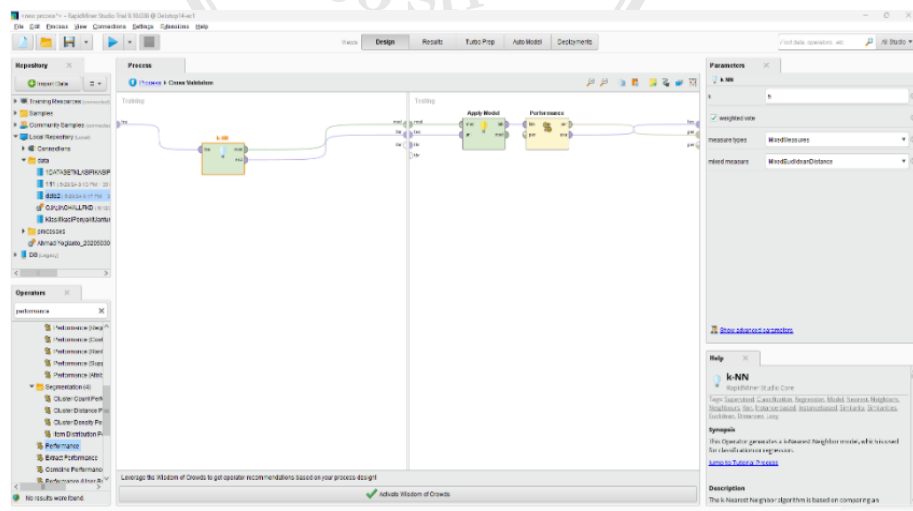
Tahap selanjutnya adalah proses transformasi data, Proses Transformasi data merupakan suatu cara yang dilakukan agar dapat memberikan inisialisasi kepada dataset yang mempunyai nilai dan disesuaikan dengan tipe data pada metode K-Nearest Neighbors. pada proses ini juga dapat ditentukan salah satu kriteria yang akan dijadikan label. pada dataset berikut ini yang dijadikan label adalah target.



Gambar 4. 4 Transformasi Data

4.3.2 Perhitungan dan Validasi

Dengan diinputkannya operator *Cross Validation* tersebut iyalah dengan masuk pada proses perhitungan menggunakan algoritma *K-Nearest Neighbors* (KNN). Kemudian menghitung tingkat akurasi dengan *Confusion Matrix* yang terdapat pada operator *Cross Validation* sebagaimana gambar dibawah ini.



Gambar 4. 5 Model Metode K-Nearest Neighbors (KNN)

Operator Read csv digunakan untuk membaca dataset yang disimpan dengan format csv dan operator cross validation merupakan suatu metode statistic yang dapat digunakan untuk mengevaluasi kinerja metode yang digunakan, dalam cross validation ini dipisahkan menjadi dua subset, yakni data testing dan data training[19].

Pada subset data training, digunakan operator KNN, hal ini dilakukan karena pada penelitian ini menggunakan metode KNN, dan tidak lupa untuk turut memasukkan tetangga terdekat (K) pada menu sebelah kanan atas, dan pada subset data testing terdapat apply model yang merupakan lanjutan dari metode KNN yang digunakan untuk mengaplikasikan model dengan dataset pelatihan pada data test[20], dalam subset data testing ini juga terdapat Performance yang berfungsi terhadap evaluasi kinerja, operator performance dapat memberikan daftar dari nilai kriteria kinerja.dibawah ini.

Row No.	target	prediction[...]	confidence(0)	confidence(1)	f1_zage	sex	cp	trestbps	chol	fbs	restecg	thalaci
1	1	1	0.214	0.786	64	1	3	110	211	0	0	144
2	1	1	0.384	0.616	58	0	2	120	340	0	1	172
3	1	1	0.395	0.605	54	1	2	125	273	0	0	152
4	1	1	0	1	45	1	0	104	208	0	0	148
5	1	1	0.400	0.600	53	0	2	128	216	0	0	115
6	1	1	0.444	0.556	62	1	2	130	231	0	1	146
7	1	1	0.201	0.799	48	1	0	122	222	0	0	186
8	1	1	0.204	0.796	57	0	0	128	303	0	0	159
9	1	1	0.399	0.601	58	1	2	140	211	1	0	155
10	1	0	0.803	0.197	54	1	2	120	258	0	0	147
11	1	1	0.405	0.595	55	1	1	130	262	0	1	155
12	1	1	0.195	0.805	41	1	2	130	214	0	0	168
13	1	0	0.583	0.417	54	0	2	160	201	0	1	163
14	1	0	0.599	0.401	67	0	0	106	223	0	1	142
15	1	1	0.204	0.796	70	1	1	156	245	0	0	143
16	1	1	0	1	47	1	2	130	253	0	1	179
17	1	1	0.399	0.601	56	1	1	130	221	0	0	163
18	0	0	1	0	67	1	0	160	286	0	0	108
19	0	0	0.600	0.400	60	1	0	130	206	0	0	132
20	0	0	0.600	0.400	50	1	0	150	243	0	0	128
21	0	0	0.592	0.408	51	0	0	130	305	0	1	142
22	0	0	0.814	0.186	62	1	1	120	281	0	0	103
23	0	0	0.597	0.403	70	1	0	130	322	0	0	109
24	0	1	0.414	0.586	60	1	0	140	293	0	0	170
...

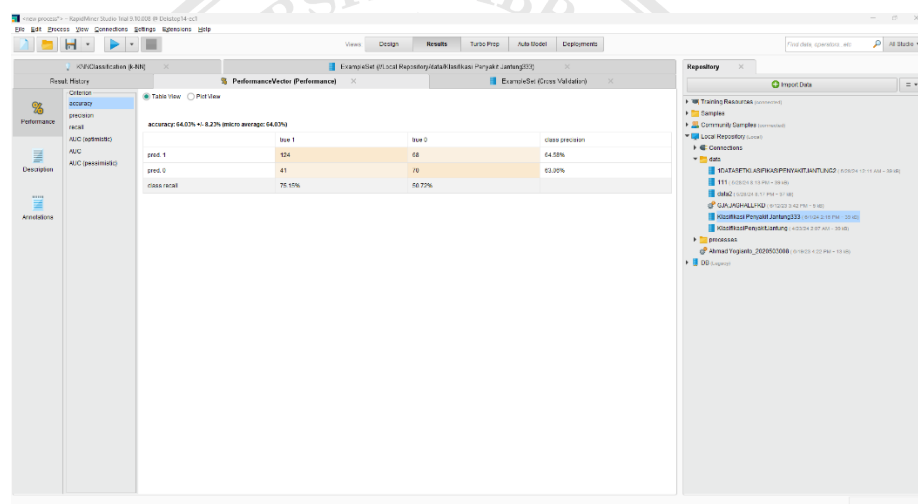
Gambar 4. 6 Hasil data testing dari Algoritma *K-Nearest Neighbors*

4.3.3 Hasil Performance Vector

Hasil dari proses akan menghasilkan performance vector yang memperoleh nilai *accuracy*, *recall*, *precision*, *simple distribution*, dan AUC (*Area Under Curve*) sebagai berikut :

a. Accuracy perhitungan RapidMiner

Dari hasil jumlah data yang telah diklasifikasikan dari data klasifikasi penyakit jantung maka dapat diketahui hasil akurasi dari perhitungan RapidMiner sebesar 64.34% dari hasil data testing.



Category	Value	Value 0	Class precision
accuracy	64.03%	8.23%	average: 64.03%
prec 1	64	68	64.58%
prec 0	41	70	63.00%
class recall	70.15%	60.73%	

Gambar 4. 7 Hasil Akurasi

Berdasarkan gambar diatas, dapat dijelaskan bahwa Penerapan metode KNN dengan parameter $K=5$ nilai akurasi sebesar 64,03% dengan Prediksi bahwa prediksi pasien memiliki penyakit jantung dan positif sebanyak 116 pasien dan prediksi pasien memiliki penyakit jantung dan negatif sebanyak 59 pasien. prediksi pasien yang tidak terkena penyakit jantung dan ternyata positif sebanyak 49 pasien dan prediksi pasien tidak terkena penyakit jantung ternyata negatif sebanyak 79 pasien.

b. Precision perhitungan RapidMiner

Hasil dari nilai precision ini adalah 64.58% untuk class “*negatif*” tidak terkena penyakit jantung dan 63.06 % untuk “*Positif*” positif terkena penyakit jantung.

The screenshot shows the PerformanceView window in RapidMiner. The 'Precision' tab is selected, displaying a table with the following data:

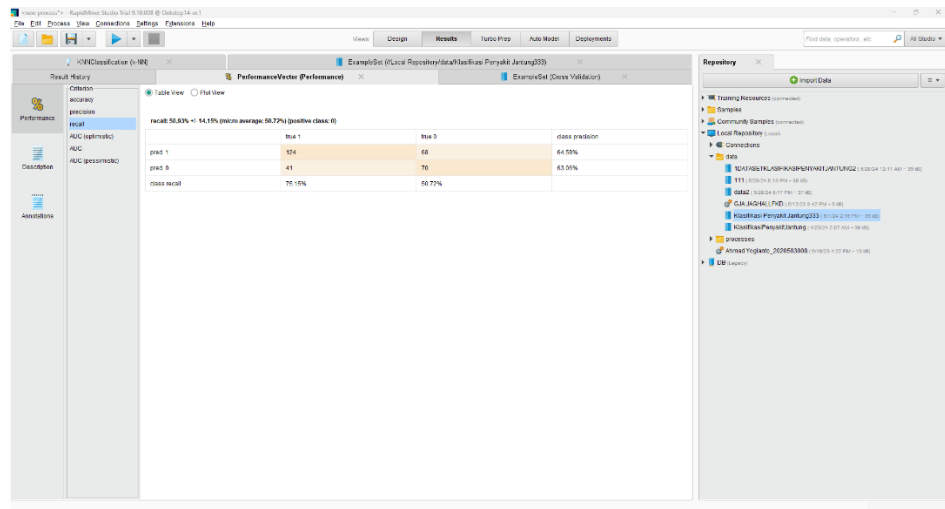
Class	Class Precision
negatif	64.58%
positif	63.06%

Below the table, the 'class recall' is shown as 75.15% for 'negatif' and 50.72% for 'positif'. The overall precision is 62.82% with a 0.39% micro average.

Gambar 4. 8 Perhitungan Precision

c. Recall perhitungan RapidMiner

Hasil dari perhitungan nilai recall pada RapidMiner sebesar 75.15% untuk class “*Positif*” memiliki penyakit jantung dan 50.72% untuk class “*Negatif*” memiliki penyakit jantung.



class	precision	recall	f1-score
class 1	64.55%	64.55%	64.55%
class 0	43.50%	43.50%	43.50%
class recall	56.82%	54.15%	55.72%

Gambar 4. 9 Perhitungan Recall

d. AUC (Area Under Curve)

Dapat diperhatikan dari gambar dibawah ini bahwa AUC (Area Under Curve) dari algoritma *K-Nearest Neighbors* (KNN) pada penelitian ini adalah sebesar 0.719%. AUC (Area Under Curve) terbagi menjadi 2 yaitu AUC *optimistic* dan AUC (Area Under Curve) *pessimistic*. Untuk AUC *optimistic* menghasilkan 0.723 % sedangkan untuk AUC *pessimistic* menghasilkan nilai 0.714%.



Gambar 4. 10 AUC (Area Under Curve)

Sehingga dapat disimpulkan dari proses validasi tersebut dengan menggunakan rumus perhitungan *Confusion Matrix* sehingga menghasilkan tingkat Akurasi sebesar 64,34% , Presisi 66.20% , Recall sebesar 70.30% , AUC (The Area Under Curve) sebesar 0.697 %.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil pembahasan yang telah diuraikan pada bab sebelumnya, maka diperoleh kesimpulan sebagai berikut:

1. Dengan dibuatnya model algoritma ini dapat mengoptimalkan pelajar yang membutuhkan pembinaan dan mengevaluasi pendidik tentang capaian pembelajaran berdasarkan data yang telah diprediksi sebelumnya.
2. Algoritma *K-Nearest Neighbors* (KNN) memanfaatkan data training untuk menghasilkan tetangga terdekat sehingga dapat menghasilkan keputusan apakah pasien tersebut memiliki penyakit jantung atau negative menderita penyakit jantung dengan menggunakan 14 atribut sebagai parameter dengan 1 label pengukur dengan menggunakan $K = 5$.
3. Metode *K-Nearest Neighbors* (KNN) dapat membantu tenaga medis dalam mengambil keputusan dan mengambil penanganan dini untuk pasien sehingga bisa mengatasi pasien yang memiliki penyakit jantung.

5.2 Saran

Berdasarkan hasil penelitian dan pembahasan serta kesimpulan yang telah dijelaskan, maka berikut ini merupakan sebuah saran yang diharapkan menjadi masukan dan pertimbangan untuk penelitian selanjutnya. Adapun beberapa saran sebagai berikut :

1. Metode *K-Nearest Neighbors* (KNN) ini hendaknya dapat diimplementasikan ke dalam sebuah aplikasi program yang dapat bermanfaat untuk membantu tenaga medis untuk dapat menganalisa pasien yang menderita penyakit jantung
2. Untuk meningkatkan hasil kualitas penelitian ini agar lebih baik lagi, diharapkan untuk penelitian yang akan dilakukan selanjutnya untuk dapat mengembangkan dan menambah atau merubah atribut sehingga mendapatkan hasil klasifikasi yang lebih baik lagi.
3. Diharapkan pula untuk penelitian selanjutnya dapat mencoba metode lain seperti metode Decision tree dan dalam pengklasifikasian data sehingga kita sebagai peneliti dapat membandingkan metode apa yang memiliki tingkat akurasi yang lebih tinggi antara metode Algoritma *K-Nearest Neighbors* (KNN) dengan metode yang lain dalam pengklasifikasikan suatu data.

DAFTAR PUSTAKA

- [1] Anggun Pastika Sandi and Vina Widya Ningsih, "Implementasi Data Mining Sebagai Penentu Persediaan Produk Dengan Algoritma Fp-Growth Pada Data Penjualan Sinarmart," *J. Publ. Ilmu Komput. dan Multimed.*, vol. 1, no. 2, pp. 111–122, 2022, doi: 10.55606/jupikom.v1i2.343.
- [2] R. T. Vulandari, *Data Mining Teori dan Aplikasi Rapidminer*. Surakarta: Penerbit Gava Media, 2017.
- [3] M. Ali, R. Sekolah, T. Teknik, S. Jalan, and N. Jaya, "Penggunaan Data Mining Dalam Memecahkan Permasalahan Dalam Pengelolaan Data," 2022, [Online]. Available: www.IJCSI.org.
- [4] D. Sitanggang, N. Nicholas, V. Wilson, A. R. A. Sinaga, and A. D. Simanjuntak, "Implementasi Data Mining Untuk Memprediksi Penyakit Jantung Menggunakan Metode K-Nearest Neighbor Dan Logistic Regression," *J. Tek. Inf. dan Komput.*, vol. 5, no. 2, p. 493, 2022, doi: 10.37600/tekinkom.v5i2.698.
- [5] T. K. Ningsih and H. Zakaria, "Implementasi Algoritma K-Nearest Neighbor Pada Sistem Deteksi Penyakit Jantung (Studi Kasus : Klinik Makmur Jaya)," *J. Ilmu Komput. dan Pendidik.*, vol. 2, no. 1, pp. 6–21, 2023, [Online]. Available: <https://journal.mediapublikasi.id/index.php/logic>
- [6] M. Lestari, "Penerapan Algoritma Klasifikasi Nearest Neighbor (K-NN) untuk Mendeteksi Penyakit Jantung," *Fakt. Exacta*, vol. 7, no. September 2010, pp. 366–371, 2014.
- [7] A. PS, A. Nugroho Sihananto, and D. Arman Prasetya, "Implementasi Metode K-NN dalam Klasterisasi Kasus Kesehatan Jantung," *ALINIER J. Artif. Intell. Appl.*, vol. 3, no. 2, pp. 18–21, 2022, doi: 10.36040/alinier.v3i2.5761.
- [8] L. Marleni and A. Alhabib, "Faktor Risiko Penyakit Jantung Koroner di RSI SITI Khadijah Palembang," *J. Kesehat.*, vol. 8, no. 3, p. 478, 2017, doi: 10.26630/jk.v8i3.663.
- [9] I. N. U. R. Septiani, J. Matematika, F. Matematika, D. A. N. Ilmu, P. Alam, and U. Sriwijaya, "IMPLEMENTASI ALGORITMA ITERATIVE DICHOTOMIZER THREE (ID3) PADA FUZZY DECISION TREE UNTUK KLASIFIKASI PENYAKIT JANTUNG," 2024.
- [10] Y. Pratama, A. Prayitno, D. Azrian, N. Aini, Y. Rizki, and E. Rasywir, "Klasifikasi Penyakit Gagal Jantung Menggunakan Algoritma K-Nearest Neighbor," *Bull. Comput. Sci. Res.*, vol. 3, no. 1, pp. 52–56, 2022, doi: 10.47065/bulletincsr.v3i1.203.
- [11] A. M. Argina, "Penerapan Metode Klasifikasi K-Nearest Neighbor pada

- Dataset Penderita Penyakit Diabetes,” *Indones. J. Data Sci.*, vol. 1, no. 2, pp. 29–33, 2020, doi: 10.33096/ijodas.v1i2.11.
- [12] A. M. Sholihah, N. Suarna, G. Dwilestari, and N. R., “Implementasi Metode K-means Clustering Untuk Menganalisa Penerima Bantuan di Desa Palasah,” *J. Inform. dan Teknol. Inf.*, vol. 1, no. 2, pp. 111–117, 2023, doi: 10.56854/jt.v1i2.121.
- [13] Anwar Pauji, S. Aisyah, A. Surip, R. Saputra, and I. Ali, “Implementasi Algoritma K-Nearest Neighbor Dalam Menentukan Penerima Bantuan Langsung Tunai,” *KOPERTIP J. Ilm. Manaj. Inform. dan Komput.*, vol. 4, no. 1, pp. 21–27, 2022, doi: 10.32485/kopertip.v4i1.114.
- [14] A. Yudhana, S. Sunardi, and A. J. S. Hartanta, “Algoritma K-Nn Dengan Euclidean Distance Untuk Prediksi Hasil Penggajian Kayu Sengon,” *Transmisi*, vol. 22, no. 4, pp. 123–129, 2020, doi: 10.14710/transmisi.22.4.123-129.
- [15] A. D. W. M. Sidik, I. Himawan Kusumah, A. Suryana, Edwinanto, M. Artiyasa, and A. Pradiftha Junfithrana, “Gambaran Umum Metode Klasifikasi Data Mining,” *Fidel. J. Tek. Elektro*, vol. 2, no. 2, pp. 34–38, 2020, doi: 10.52005/fidelity.v2i2.111.
- [16] W. S. Naomi, I. Picauly, and S. M. Toy, “Faktor Risiko Kejadian Penyakit Jantung Koroner,” *Media Kesehat. Masy.*, vol. 3, no. 1, pp. 99–107, 2021, doi: 10.35508/mkm.v3i1.3622.
- [17] S. Sahar, “Analisis Perbandingan Metode K-Nearest Neighbor dan Naïve Bayes Clasiffier Pada Dataset Penyakit Jantung,” *Indones. J. Data Sci.*, vol. 1, no. 3, pp. 79–86, 2020, doi: 10.33096/ijodas.v1i3.20.
- [18] D. P. Utomo and B. Purba, “Penerapan Datamining pada Data Gempa Bumi Terhadap Potensi Tsunami di Indonesia,” *Pros. Semin. Nas. Ris. Inf. Sci.*, vol. 1, no. September, p. 846, 2019, doi: 10.30645/senaris.v1i0.91.
- [19] P. Romadloni, B. Adhi Kusuma, and W. Maulana Baihaqi, “Komparasi Metode Pembelajaran Mesin Untuk Implementasi Pengambilan Keputusan Dalam Menentukan Promosi Jabatan Karyawan,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 6, no. 2, pp. 622–628, 2022, doi: 10.36040/jati.v6i2.5238.
- [20] A. Nurjanah and A. Rifai, “Penerapan Algoritma K-Nearest Neighbor Untuk Klasifikasi Kelayakan Status Penduduk Miskin Di Desa Susukan Tonggoh,” *J. Wahana Inform.*, vol. 2, no. 1, pp. 164–176, 2023.

LAMPIRAN**CURICULUM VITAE****A. Biodata Pribadi**

1. Nama : Ahmad Yogiarto
2. TTL : Bondowoso, 13 Oktober 2001
3. Jenis Kelamin : Laki-laki
4. Agama : Islam
5. Kebangsaan : Indonesia
6. Status : Belum Menikah
7. Tinggi, Berat Badan : cm, kg
8. No. Hp : -
9. Alamat : Desa Kalisat, Kecamatan Ijen,
Kabupaten Bondowoso, Jawa Timur
10. E-mail : ahmadyogiarto13@gmail.com

B. Riwayat Pendidikan

1. TK/RA : TK Kartini
2. SD : SDN 01 Sempol
3. SMP : SMPN 01 Ijen
4. SLTA : SMK Ibrahimy 1 Sukorejo

Lampiran

Lampiran A : Dataset

1	age	sex	cp	trestbps	chol	fb	restecg	thalach	exang	oldpeak	slope	ca	thal	target
2	63	1	3	145	233	1	0	150	0	23	0	0	1	1
3	37	1	2	130	250	0	1	187	0	35	0	0	2	1
4	41	0	1	130	204	0	0	172	0	14	2	0	2	1
5	56	1	1	120	236	0	1	178	0	8	2	0	2	1
6	57	0	0	120	354	0	1	163	1	6	2	0	2	1
7	57	1	0	140	192	0	1	148	0	4	1	0	1	1
8	56	0	1	140	294	0	0	153	0	13	1	0	2	1
9	44	1	1	120	263	0	1	173	0	0	2	0	3	1
10	52	1	2	172	199	1	1	162	0	5	2	0	3	1
11	57	1	2	150	168	0	1	174	0	16	2	0	2	1
12	54	1	0	140	239	0	1	160	0	12	2	0	2	1
13	48	0	2	130	275	0	1	139	0	2	2	0	2	1
290	57	1	0	110	335	0	1	143	1	3	1	1	3	0
291	55	0	0	128	205	0	2	130	1	2	1	1	3	0
292	61	1	0	148	203	0	1	161	0	0	2	1	3	0
293	58	1	0	114	318	0	2	140	0	44	0	3	1	0
294	58	0	0	170	225	1	0	146	1	28	1	2	1	0
295	67	1	2	152	212	0	0	150	0	8	1	0	3	0
296	44	1	0	120	169	0	1	144	1	28	0	0	1	0
297	63	1	0	140	187	0	0	144	1	4	2	2	3	0
298	63	0	0	124	197	0	1	136	1	0	1	0	2	0
299	59	1	0	164	176	1	0	90	0	1	1	2	1	0
300	57	0	0	140	241	0	1	123	1	2	1	0	3	0
301	45	1	3	110	264	0	1	132	0	12	1	0	3	0
302	68	1	0	144	193	1	1	141	0	34	1	2	3	0
303	57	1	0	130	131	0	1	115	1	12	1	1	3	0
304	57	0	1	130	236	0	0	174	0	0	1	1	2	0

Lampiran B : Kartu Bimbingan

Pembimbing I : Ahmad Haniadi, M.Ts.

NO	TANGGAL	CATATAN	PARAF
		Konsultasi Data	[Signature]
		Bimbingan Bab III	[Signature]
		Bimbingan Raport	[Signature]
		Perhitungan Algoritma KNN	[Signature]
		Bimbingan Bab IV	[Signature]
		Acc	[Signature]

Pembimbing II : Zaheer Fatah, M.Ts.

NO	TANGGAL	CATATAN	PARAF
		Bimbingan Bab III	[Signature]
		Acc Bab I-III	[Signature]
		Bimbingan Raport	[Signature]
		Perhitungan Algoritma KNN	[Signature]
		Revisi Laporan KTI	[Signature]
		Acc Bab IV	[Signature]
		Acc	[Signature]

