

**PENERAPAN FUZZY K-NEAREST NEIGHBOR (FK-NN) UNTUK  
MENGKLASIFIKASI PENYAKIT JANTUNG BERDASARKAN  
SPECTF HEART DATASET**

**SKRIPSI**

Untuk memenuhi sebagian persyaratan  
memperoleh gelar Sarjana Komputer

Disusun oleh :  
Hadistria Massalli  
0910963113



**PROGRAM STUDI TEKNIK INFORMATIKA  
JURUSAN TEKNIK INFORMATIKA  
FAKULTASI ILMU KOMPUTER  
UNIVERSITAS BRAWIJAYA  
MALANG  
2016**

## PENGESAHAN

### PENERAPAN FUZZY K-NEAREST NEIGHBOR (FK-NN) UNTUK MENGLASIFIKASI PENYAKIT JANTUNG BERDASARKAN SPECTF HEART DATASET

SKRIPSI

Diajukan untuk memenuhi sebagian persyaratan  
memperoleh gelar Sarjana Komputer

Disusun Oleh :  
HADISTRIA MASSALLI  
NIM: 0910963113

Skripsi ini telah diuji dan dinyatakan lulus pada

18 Agustus 2016

Telah diperiksa dan disetujui oleh:

Dosen Pembimbing I

Dosen Pembimbing II

M. Tanzil Furqon, S.Kom, M.Sc  
NIP. 19820930 200801 1 004

Budi Darma Setiawan, S.Kom, M.Cs  
NIP. 198410152014041002

Mengetahui  
Ketua Jurusan Teknik Informatika

Tri Astoto Kurniawan, S.T.,M.T.,Ph.D  
NIP. 19710518 200312 1 001

## PERNYATAAN ORISINALITAS

Saya menyatakan dengan sebenar-benarnya bahwa sepanjang pengetahuan saya, di dalam naskah skripsi ini tidak terdapat karya ilmiah yang pernah diajukan oleh orang lain untuk memperoleh gelar akademik di suatu perguruan tinggi, dan tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis disitasi dalam naskah ini dan disebutkan dalam daftar pustaka.

Apabila ternyata didalam naskah skripsi ini dapat dibuktikan terdapat unsur-unsur plagiasi, saya bersedia skripsi ini digugurkan dan gelar akademik yang telah saya peroleh (sarjana) dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku (UU No. 20 Tahun 2003, Pasal 25 ayat 2 dan Pasal 70).

Malang, 11 Agustus 2016

Hadistria Massalli

NIM: 0910963113

## KATA PENGANTAR

Segala puji bagi Allah atas segala rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul "PENERAPAN FUZZY K-NEAREST NEIGHBOR(FK-NN) UNTUK MENGLASIFIKASI PENYAKIT JANTUNG BERDASARKAN SPECTF HEART DATASET . kripsi ini disusun untuk memperoleh gelar Sarjana Pada program studi Teknik Informatika Fakultas ilmu computer di Universitas Brawijaya.

Atas bimbingan dari dosen pembimbing dan Bantuan dari rekan – rekan terbaik serta dukungan moral dari keluarga maka disusunlah skripsi ini. Maka dari itu dengan segala kerendahan hati penulis ingin menyampaikan ucapan terima kasih pada :

- 1.Orang tua yang selalu mendukung dan mendoakan agar lancarnya skripsi ini
- 2.Bapak M.Tanzil Furqon selaku pembimbing I, arahan dan masukan yang selama ini diberikan
- 3.Bapak Budi Darma S selaku pembimbing II, arahan dan masukan yang selama ini diberikan
- 4.Teman teman yang memberikan bantuan,dukungan dan doa agar lancarnya proses pengerjaan skripsi ini

Semoga semua bantuan yang diberikan akan mendapat pahala kebaikan dari Allah. Dan penulis berharap semoga dengan adanya skripsi ini dapat bermanfaat untuk semua pihak yang memerlukan.

Malang, 11 Agustus 2016

Hadistria Massalli

Varna\_hadistria@yahoo.com

## ABSTRAK

Penelitian ini membahas tentang algoritma FKNN (Fuzzy K-nearest Neighbor) untuk mendiagnosa penderita penyakit jantung. Penelitian ini dibangun menggunakan data pasien penderita penyakit jantung (SPECT). SPECT (Single proton emission computer tomography adalah teknik pencitraan dibidang radionuklida dari dunia kedokteran. Cara kerja SPECT yaitu melihat proses darah mengalir melalui pembuluh darah di jantung, sehingga dapat mendeteksi secara rinci pada bagian yang tersumbat serta kerja jantung secara keseluruhan. Algoritma FKNN adalah algoritma yang memberikan nilai keanggotaan kelas pada vector sampel dan bukan menempatkan vector pada kelas tertentu. Data latih yang digunakan adalah data pasien penyakit jantung di Medical collage of ohio, U.S.A pada tahun 2001.

Hasil dari pengujian ini untuk mengetahui pengaruh data latih dan nilai k pada sebaran kelas, serta mengetahui tingkat akurasi dari penelitian ini. Dari hasil system didapatkan akurasi sebesar 66,6%, akurasi sebesar 66,6% terhadap nilai K. Dapat disimpulkan bahwa penelitian menggunakan algoritma FKNN ini memiliki akurasi yang cukup baik untuk mendiagnosa penyakit jantung.

**Kata kunci** : Penyakit jantung, FKNN, akurasi, SPECT

## ABSTRACT

This study is discusses the algorithm FKNN (Fuzzy K-nearest Neighbor) for diagnose patients with cardiovascular disease. This study is built using the data of patients with heartdisease (SPECT). SPECTF (single proton emission computed tomography is an imaging technique in the field of radionuclides from the medical world. The workings of SPECT is to see the blood flowing through blood vessels in the heart, so it can detect in detail the occlusion as well as the work of the heart as a whole. The algorithm FKNN is an algorithm that value vector class membership in the sample and not put a vector in a particular class. the data used is the practice that patients with heart disease at the Medical collage of ohio, USA in 2001.

The results of this test to determine the effect of training data and k values on the distribution of class , as well as determine the accuracy of the study. the results of the system obtained an accuracy of 66.6%, 66.6% accuracy rate against K. Can conclude this studies using the FKNN algorithm has good accuracy for diagnosing heart disease.

**Keywords:** Heart disease, FKNN, accuracy, SPECTF

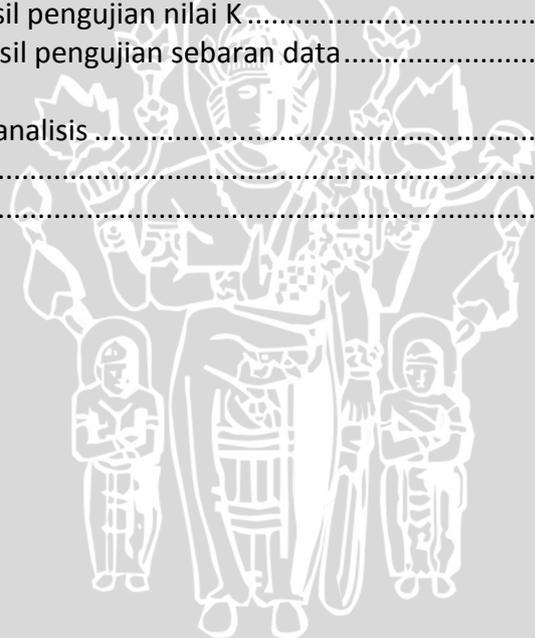


DAFTAR ISI

	Halaman
LEMBAR PERSETUJUAN .....	i
LEMBAR PENGESAHAN .....	ii
LEMBAR PERNYATAAN .....	iii
KATA PENGANTAR.....	iv
DAFTAR ISI .....	v
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah .....	3
1.3 Batasan Masalah .....	3
1.4 Tujuan.....	3
1.5 Manfaat .....	4
1.6 Sistematika Penulisan .....	4
BAB II KAJIAN PUSTAKA .....	5
2.1 Jantung .....	6
2.2 SPECTF .....	7
2.3 Logika Fuzzy.....	8
2.4 Data Mining.....	8
2.4.1 Pengertian Data Mining.....	8
2.4.2 Proses Data Mining.....	8
2.5 Klasifikasi Data Mining .....	9
2.6 Definisi K-Nearest Neighbor.....	9
2.6.1 Proses N-Nearest Neighbor .....	9
2.7 Fuzzy K-Nearest Neighbor .....	10
2.7.1 Definisi Fuzzy K-Nearest Neighbor .....	10
2.7.2 Proses Fuzzy K-Nearest Neighbor .....	11
2.8 Perhitungan Akurasi .....	11
BAB III METODOLOGI .....	12
3.1 Studi Literatur .....	13
3.2 Data Penelitian .....	13
3.2.1 Deskripsi Umum Sistem .....	13
3.2.2 Batasan Sistem.....	13
3.2.3 Perancangan Sistem .....	13
3.2.4 Proses K-Nearest Neighbor.....	14
3.2.5 Contoh perhitungan Manual .....	16
3.2.6 Data Latih dan Data Uji pada data Penderita Jantung .....	17
3.2.7 Menentukan K record terdekat .....	21
3.2.8 Menentukan maksimum membership dan kelas target .....	22
3.3 Perancangan Antarmuka.....	23
3.4 Perancangan Uji coba.....	25

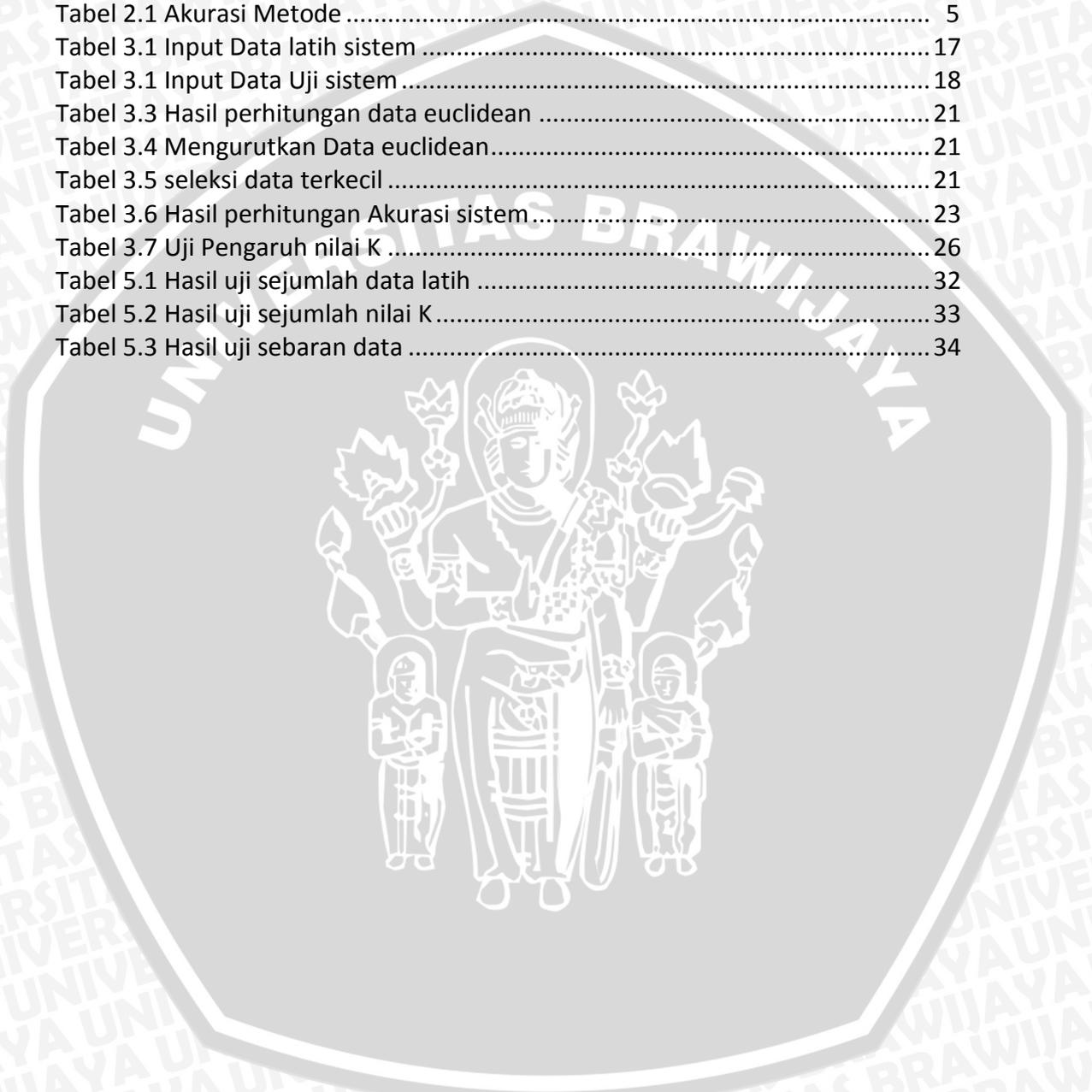


3.5 Uji pengaruh nilai K dan data latih terhadap tingkat akurasi .....	26
<b>BAB IV Implementasi</b> .....	27
4.1 Lingkungan implementasi .....	27
4.1.1 Lingkungan implementasi perangkat keras.....	27
4.1.2 Lingkungan implementasi perangkat lunak.....	27
4.2 Implementasi Program.....	27
4.2.1 Proses baca file .....	27
4.2.2 Proses FKNN.....	28
<b>BAB V Pengujian dan analisis</b> .....	32
5.1 Pengujian.....	32
5.1.1 Pengujian tingkat akurasi terhadap data latih .....	32
5.1.2 Pengujian tingkat akurasi terhadap nilai K .....	32
5.1.3 Pengujian tingkat akurasi terhadap sebaran data.....	33
5.2 Analisis.....	35
5.2.1 Analisis hasil pengujian data latih.....	35
5.2.2 Analisis hasil pengujian nilai K .....	35
5.2.3 Analisis hasil pengujian sebaran data.....	35
<b>BAB VI Pengujian dan analisis</b> .....	36
6.1 Kesimpulan.....	36
6.2 Saran.....	36



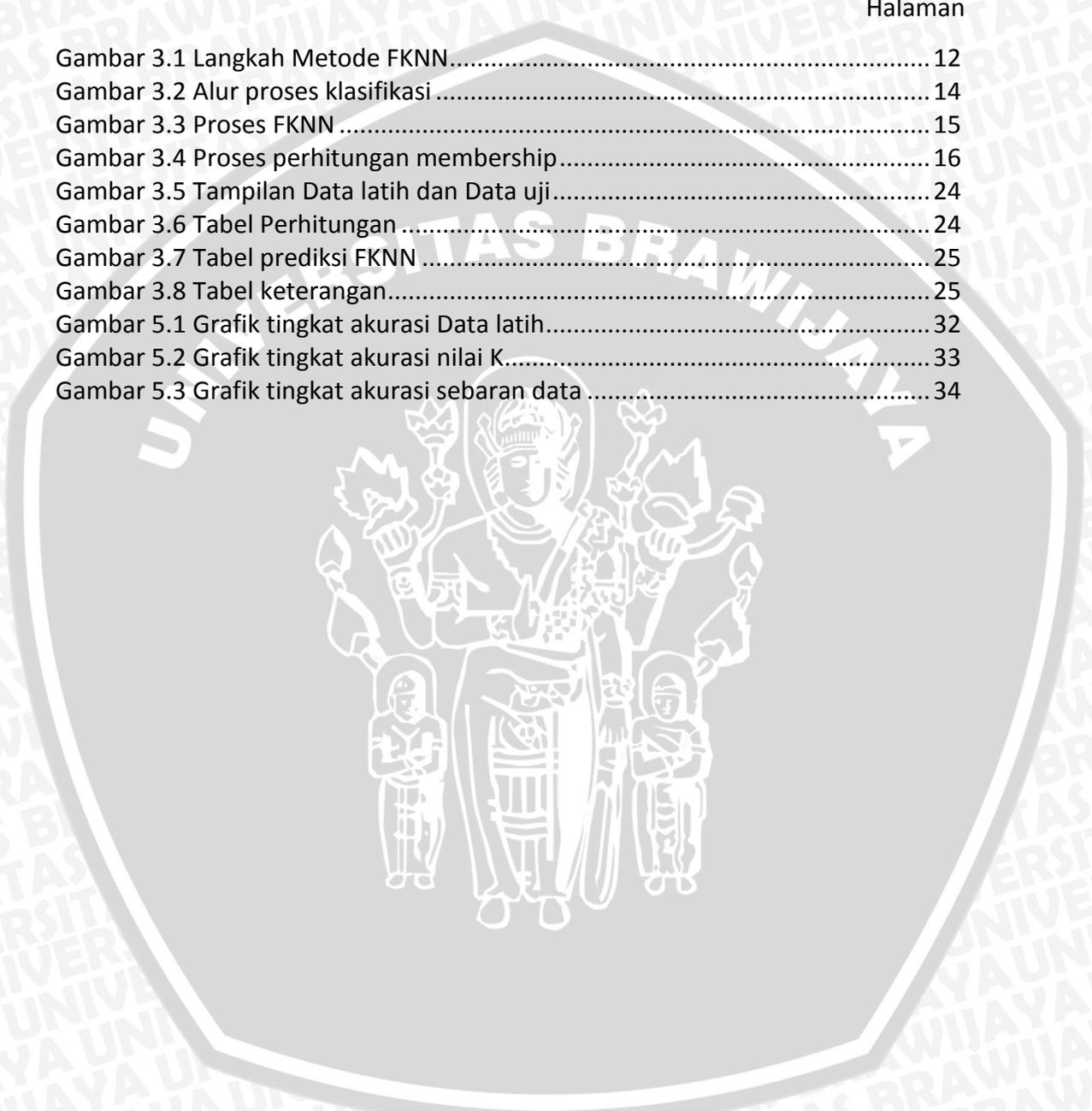
DAFTAR TABEL

	Halaman
Tabel 2.1 Akurasi Metode .....	5
Tabel 3.1 Input Data latih sistem .....	17
Tabel 3.1 Input Data Uji sistem .....	18
Tabel 3.3 Hasil perhitungan data euclidean .....	21
Tabel 3.4 Mengurutkan Data euclidean.....	21
Tabel 3.5 seleksi data terkecil .....	21
Tabel 3.6 Hasil perhitungan Akurasi sistem.....	23
Tabel 3.7 Uji Pengaruh nilai K .....	26
Tabel 5.1 Hasil uji sejumlah data latih .....	32
Tabel 5.2 Hasil uji sejumlah nilai K.....	33
Tabel 5.3 Hasil uji sebaran data .....	34



DAFTAR GAMBAR

	Halaman
Gambar 3.1 Langkah Metode FKNN.....	12
Gambar 3.2 Alur proses klasifikasi .....	14
Gambar 3.3 Proses FKNN .....	15
Gambar 3.4 Proses perhitungan membership.....	16
Gambar 3.5 Tampilan Data latih dan Data uji.....	24
Gambar 3.6 Tabel Perhitungan .....	24
Gambar 3.7 Tabel prediksi FKNN .....	25
Gambar 3.8 Tabel keterangan.....	25
Gambar 5.1 Grafik tingkat akurasi Data latih.....	32
Gambar 5.2 Grafik tingkat akurasi nilai K.....	33
Gambar 5.3 Grafik tingkat akurasi sebaran data .....	34



## BAB I PENDAHULUAN

### 1.1 Latar Belakang

Jantung merupakan salah satu organ terpenting yang ada di dalam tubuh. Jantung berfungsi sebagai pompa ganda atau disebut juga sistem kardiovaskular. Fungsi jantung adalah memompa darah keseluruhan tubuh kemudian dibersihkan oleh paru-paru dan menampungnya kembali. Disebelah kanan jantung memompa darah ke paru-paru dan disebelah kiri jantung memompa darah ke seluruh tubuh (Guyton, 2006). Salah satu permasalahan yang sering terjadi adalah penyumbatan pada pembuluh darah (pembuluh koroner) karena endapan lemak dan kolesterol atau yang biasa kita sebut Penyakit Jantung Koroner(PJK) (Wayne, 2013).

Salah satu penyakit yang sering menyebabkan kematian dinegara negara besar dan berkembang yaitu penyakit jantung koroner (PJK), termasuk Indonesia. Secara global penyaki ini menjadi salah satu penyebab kematian di Negara Negara berkembang setelah infeksi pada tahun 2010 (Abdul, 2006) berikut ini adalah hasil penelitian para penderita PJK yang digolongkan berdasarkan umur, yaitu umur 50-59 tahun sebanyak 46,7% tidak menderita PJK, sedangkan yang berumur kurang dari 40 tahun sebanyak 78,6% tidak menderita PJK. Kemudian untuk tingkat pendidikan yaitu; Diploma: yang tidak terkena PJK sebesar 62,5%, SMP yang tidak terkena PJK sebesar 71,4%. yang terbiasa olahraga tenis meja mempunyai 99% tidak terkena PJK, tenis lapangan 80,0% tidak terkena PJK. (Hermansyah, 2012).

Pada tahun 2007 data yang didapat oleh riset kesehatan dasar(Riskesdas), pada tahun 2007 penyakit jantung adalah salah satu penyebab utama kematian. prevalensi Kematian akibat penyakit jantung secara nasional mencapai 7,2%, hipertensi dan stroke mencapai 31,9% sedangkan penyakit kardiovaskular mencapai 12% (Nila, 2014).

Meski menjadi pembunuh utama, tetapi masih sedikit sekali orang yang tahu tentang penyakit jantung koroner dan faktor resikonya. Jika faktor resiko suatu penyakit telah diketahui tentunya akan lebih mudah dilakukan tindakan pencegahan. Karena kurang jelasnya info membuat orang awam kesulitan mengetahui resiko penyakit jantung ini. Dengan kata lain, kesulitan untuk membedakan antara orang yang sehat dengan tidak sehat, sehingga sulit untuk membedakannya. Kemudian banyaknya faktor analisa untuk mendiagnosa penyakit jantung dari pasien membuat kita kesulitan. Jadi, untuk mengetahui seberapa parah penyakit jantung yang diderita masyarakat memerlukan bantuan alat, sehingga masyarakat bisa mengantisipasi penyakit tersebut dengan Dengan melakukan tindakan pencegahan seperti merubah gaya hidup atau melakukan pemeriksaan kesehatan. Penyakit jantung koroner banyak disebabkan oleh penyumbatan pembuluh darah koroner di jantung, untuk itu teknik diagnose yang dianggap paling akurat dalam mendeteksi penyakit ini adalah menggunakan teknik pencitraan dibidang radionuklida yaitu teknik pencitraan single proton emission computer tomography (SPECT). Salah satu teknik dari dunia kedokteran

nuklir yaitu untuk melihat proses darah mengalir melalui pembuluh darah di jantung, sehingga dapat mendeteksi secara rinci pada bagian yang tersumbat serta kerja jantung secara keseluruhan (Dian, 2013).

Meskipun telah ditemukannya teknik diagnosa penyakit jantung koroner dari segi medis, namun hasil ini pun belum dirasa akurat dikarenakan keterbatasan-keterbatasan dari proses pencitraan SPECT tersebut. Untuk itu, demi mendapatkan diagnosa yang baik diperlukan adanya perbandingan metode untuk melakukan identifikasi data berdasarkan gejala-gejala yang dialami dan berkaitan dengan data riwayat pasien yang terekam dalam data rekam medik. Data pasien yang sangat kompleks membutuhkan pengolahan dan analisa yang akurat untuk melakukan tindakan medis.

Data dari faktor penyakit Jantung merupakan data yang masih bersifat tidak jelas dan terdapat unsur yang masih belum pasti didalamnya. Untuk itu, terdapat beberapa algoritma yang dapat digunakan untuk membantu klasifikasi maupun identifikasi diagnosa penyakit pada ilmu kecerdasan buatan salah satunya metode fuzzy dan *K-Nearest Neighbor*. (Sriyanti, 2013) *Logika fuzzy* dapat dipakai untuk memodelkan proses berpikir manusia yang menyangkut unsur ketidakpastian, keraguan dan linguistik dimana metode klasik yang biasanya digunakan seperti metode matematis sulit untuk diterapkan karena kurang cukupnya pengetahuan (Agus, 2005).

Metode *k-Nearest Neighbor* (k-NN) merupakan metode klasifikasi data yang cara kerjanya lebih sederhana bila dibandingkan dengan metode klasifikasi pada data *mining* lainnya. Dimana metode ini melakukan klasifikasi kepada data baru yang masih belum diketahui masuk kedalam kelas mana, dengan menggunakan beberapa data dengan sejumlah *k* yang letaknya terdekat dengan data baru tersebut (Nils, 1996).

Metode k-NN dapat di gabungkan dengan beberapa metode lain, salah satu variasi dari penggunaan metode k-NN ini adalah *Fuzzy k-Nearest Neighbor* (Fk-NN) dimana metode ini menggabungkan teknik *fuzzy* dengan teknik data *mining* (Zhang, 2009). Fk-NN sebelumnya telah digunakan untuk memprediksi *IRIS*, *Vertebratal Column*, *Liver*, *Diabetes* oleh (James, 1985) dan *Web Classification Document* oleh (Zhang, 2009). Dimana, dalam implementasinya algoritma Fk-NN memberikan nilai keanggotaan kelas pada *vektor* sampel dan bukan menempatkan *vektor* pada kelas tertentu. Keuntungan yang didapat adalah nilai-nilai keanggotaan *vektor* memberikan tingkat jaminan pada hasil klasifikasi. Dasar dari algoritma Fk-NN adalah untuk menetapkan nilai keanggotaan sebagai fungsi jarak *vektor* dari k-NN dan keanggotaan tetangga mereka di kelas-kelas yang memungkinkan. Akurasi penelitian sebelumnya mencapai 88% (Reviangga, 2013).

Akurasi hasil dari suatu identifikasi data dalam dunia kedokteran sangatlah berpengaruh terhadap keputusan yang akan diambil oleh seorang dokter untuk melakukan tindakan terhadap pasien atau dalam melakukan analisa data untuk mendapatkan informasi yang akurat mengenai penyakit tersebut (Sri,2013), Untuk itu diharapkan dengan diterapkannya metode *fuzzy K-Nearest*

*Neighbor* dapat membantu mengklasifikasikan secara lebih objektif hasil dari diagnosa pada penyakit jantung koroner.

Berdasarkan uraian di atas, maka penelitian ini diberi judul “*PENERAPAN FUZZY K-NEAREST NEIGHBOR (FK-NN) UNTUK MENGLASIFIKASI PENYAKIT JANTUNG BERDASARKAN SPECT HEART DATASET*”. Data yang digunakan untuk klasifikasi pada Single Proton Emission Computed Tomography (SPECT) yang berupa 44 attribut.

## 1.2 Rumusan Masalah

1. Bagaimana menerapkan metode Fk-NN untuk klasifikasi pada penyakit jantung berdasarkan Single Photon Emission Computed Tomography (SPECTF) *Heart* dataset
2. Bagaimana tingkat akurasi hasil klasifikasi Single-photon emission computed tomography (SPECTF) menggunakan metode Fk-NN.

## 1.3 Batasan Masalah

Dari latar belakang di atas dapat diambil rumusan masalah sebagai berikut:

1. Data yang digunakan pada penelitian ini adalah *single photon emission computed tomography* yang diperoleh dari situs [ics.uci.edu](http://ics.uci.edu): dengan jumlah data sebanyak 267 *SPECT image set* (patients), dan 2 kelas yaitu 0 dan 1.
2. Parameter Penyakit Jantung yang digunakan berupa data attribut :  
F1R (ROI (region of interest) 1 in rest),F1S (ROI 1 in stress),F2R (ROI 2 in rest),F2S (ROI 2 in stress),F3R (ROI 3 in rest),F3S (ROI 3 in stress),F4R (ROI 4 in rest),F4S (ROI 4 in stress),F5R (ROI 5 in rest),F5S (ROI 5 in stress),F6R (ROI 6 in rest),F6S (ROI 6 in stress),F7R (ROI 7 in rest),F7S (ROI 7 in stress),F8R (ROI 8 in rest),F8S (ROI 8 in stress),F9R (ROI 9 in rest),  
,F9S (ROI 9 in stress),F10R (ROI 10 in rest),F10S (ROI 10 in stress),F11R (ROI 11 in rest),F11S (ROI 11 in stress),F12R (ROI 12 in rest),F12S (ROI 12 in stress),F13R (ROI 13 in rest),F13S (ROI 13 in stress),F14R (ROI 14 in rest),  
,F14S (ROI 14 in stress),F15R (ROI 15 in rest),F15S (ROI 15 in stress),F16R (ROI 16 in rest),F16S (ROI 16 in stress),17R (ROI 17 in rest),F17S (ROI 17 in stress),F18R (ROI 18 in rest),F18S (ROI 18 in stress),F19R (ROI 19 in rest),F19S (ROI 19 in stress),F20R (ROI 20 in rest),F20S (ROI 20 in stress),F21R (ROI 21 in rest),F21S (ROI 21 in stress),F22R (ROI 22 in rest),F22S (ROI 22 in stress)

## 1.4 Tujuan

Tujuan yang ingin dicapai dalam pembuatan skripsi ini adalah:

1. Menerapkan metode Fk-NN pada Single Proton Emission Computed Tomography (SPECT)
2. Menghitung akurasi klasifikasi yang dipengaruhi sejumlah nilai k pada Single Proton Emission Computed Tomography (SPECT) menggunakan metode Fk-NN

### 1.5 Manfaat

Manfaat yang dapat diambil dari skripsi ini yaitu dapat membantu proses diagnose penyakit jantung dari pasien berdasarkan data faktor resiko, dengan menggunakan Fk-NN. Sehingga nantinya dapat diketahui pasien masuk kategori normal atau abnormal.

### 1.6 Sistematika Penulisan

Sistematika penulisan dalam skripsi ini sebagai berikut:

#### **BAB 1 PENDAHULUAN**

Bab ini berisi latar belakang, rumusan masalah, batasan masalah, tujuan, manfaat, serta sistematika penulisan skripsi.

#### **BAB 2 TINJAUAN PUSTAKA**

Bab ini berisi tentang dasar teori dan referensi yang mendasari pembuatan klasifikasi data pada penyakit jantung yaitu kajian pustaka, jantung, klasifikasi, dan Fk-NN.

#### **BAB 3 METODOLOGI DAN PERANCANGAN**

Bab ini berisi tentang metode – metode yang digunakan untuk memprediksi tingkat resiko menggunakan klasifikasi Fk-NN, pada studi kasus diagnosis penyakit jantung.

#### **BAB 4 IMPLEMENTASI**

Membahas implementasi dari klasifikasi Fk-NN penyakit jantung yang digunakan dalam penelitian ini.

#### **BAB 5 HASIL DAN PEMBAHASAN**

Bab ini berisi tentang pembahasan dari implementasi klasifikasi Fk-NN pada penyakit jantung yang sesuai dengan system dan hasil pengujian yang akan dilakukan.

#### **BAB 6 PENUTUP**

Bab ini berisi kesimpulan yang diperoleh dari hasil pembuatan dan pengujian yang dikembangkan dalam skripsi ini serta saran untuk pengembangan lebih lanjut.

## BAB II

### KAJIAN PUSTAKA

Penelitian ini menerapkan metode *Fuzzy k-Nearest Neighbor* (Fk-NN) yang diterapkan pada dataset *Single Proton Emission Computed Tomography* (SPECT). Setiap pasien diklasifikasikan kedalam dua katagori, yaitu normal dan abnormal, kemudian atribut pada dataset akan dibagi menjadi dua yaitu rest dan stress, kemudian dihitung dengan FK-NN yang akan menghasilkan nilai keanggotaan, kemudian dipilih nilai keanggotaan terbesar untuk menentukan kelas target yang baru dan akan mendapatkan hasil.

#### 2.1 Kajian Pustaka

Pada penelitian pertama dengan judul “Implementasi Fuzzy Expert System Untuk Diagnosis Penyakit Jantung” oleh Hani Nurhayati. penelitian ini membahas tentang bagaimana menerapkan metode Fuzzy Expert dalam mendiagnosa penyakit jantung. Pada penelitian ini mempunyai 15 variabel input dan 1 variabel output. Pada pengujian ini Hasil akurasi dari uji system adalah membandingkan ouput system dengan hasil pemeriksaan dokter spesialis sebesar 70% ( Hani, 2012).

Pada penelitian kedua dengan judul “A Fuzzy Expert System for Heart Disease Diagnosis” oleh Ali adeli. penelitian ini membahas tentang bagaimana menerapkan metode *Fuzzy Expert* dalam mendiagnosa penyakit jantung. Pada penelitian ini mempunyai 13 variabel input dan 1 variabel output. Pada pengujian ini Hasil akurasi dari uji system adalah membandingkan ouput system dengan data yang ada di database dan mempunyai kesamaan sebesar 94% (Ali, 2010).

Pada penelitian Ketiga dengan judul “Penerapan *Fuzzy K-Nearest Neighbor*(FKNN) untuk mendiagnosa penderita *liver* berdasarkan *indian liver patient dataset(ILPD)*” oleh Hardika teguh. penelitian ini membahas tentang bagaimana menerapkan metode *Fuzzy K-Nearest Neighbor*(FKNN) dalam mendiagnosa penyakit Hati. Pada penelitian ini mempunyai 13 variabel input dan 1 variabel output. Pada pengujian ini Hasil akurasi dari uji system adalah membandingkan ouput system dengan data yang ada di database dan mempunyai kesamaan sebesar 94% (Hardika, 2010).

Tabel 2.1 Akurasi Metode

No	Judul	Metode	Objek(input)	Hasil
1	Implementasi Fuzzy Expert System Untuk Diagnosis Penyakit Jantung	<i>Fuzzy expert system</i>	Nyeri dada, tekanan darah, kolesterol (LDL),diabetes, data ECG, detak jantung maksimum	Perbandingan output system dengan hasil pemeriksaan dokter sebesar 70%

2	<i>A Fuzzy Expert System for Heart Disease Diagnosis</i>	<i>Fuzzy expert system</i>	Nyeri dada, tekanan darah, kolesterol, gula darah, detak jantung maksimum, data ECG, olahraga, jenis kelamin, usia	Perbandingan dengan data yang didapat sebesar 94%
3	Penerapan <i>Fuzzy K-Nearest Neighbor (FKNN)</i> untuk mendiagnosa penderita <i>liver</i> berdasarkan <i>indian liver patient dataset (ILPD)</i>	<i>Fuzzy K-Nearest Neighbor (FKNN)</i>	Data usia, jumlah <i>Bilirubin, Direct Bilirubin, protein total, albumin, rasio A/G, SGPT, SGOT dan Alk Phos.</i>	Hasil penelitian didapatkan akurasi sebesar 76% terhadap data latih, akurasi sebesar 74% terhadap nilai k

Sumber: ( Hani, 2012), (Ali, 2010), (Hardika, 2010).

## 2.2 Jantung

Jantung terletak pada rongga dada dan mempunyai berat 300 gram serta ukuran sebesar kepalan. Jantung terdiri dari otot jantung (*miokardium*) dan membrane jantung, membrane jantung bagian luar dilapisi dengan selaput jantung (*pericardium*), selaput jantung diisi cairan yang bertujuan menahan gesekan. Memompa darah keseluruh tubuh dan menampung kembali adalah fungsi utama jantung, Jantung mengalirkan darah hasil metabolisme ke seluruh tubuh yang kaya akan oksigen dari paru paru ke seluruh tubuh manusia. (Dudung, 2014)

Faktor penyebab penyakit jantung :

1. Faktor Usia dan Jenis kelamin

Pada wanita, pada usia dibawah 50 tahun wanita memiliki resiko yang lebih rendah dari laki laki akan tetapi setelah masa menopause resikonya meningkat dikarenakan penurunan hormone estrogen. Jadi bisa dikatakan bahwa salah satu factor usia dan kelamin mempengaruhi terjadinya penyakit jantung.

2. Faktor keturunan dari keluarga

Beberapa penelitian menunjukkan apabila keluarga dari wanita atau keluarga pria mempunyai riwayat penyakit jantung kemungkinan keturunan mereka memiliki peluang besar terkena penyakit jantung

3. Faktor perokok aktif atau pasif

Perokok pasif yang tinggal dengan perokok aktif memiliki resiko terkena penyakit jantung karena asap yang dikeluarkan oleh perokok aktif pasti akan terhirup oleh perokok aktif. Kira kira 20% kematian yang ada disebabkan

oleh penyakit jantung dan 50% berasal dari wanita yang usianya dibawah 55 tahun.

4. Gaya hidup kurang olahraga

Orang yang kurang olahraga memiliki resiko terkena serangan jantung yang lebih tinggi. Selain menurunkan tekanan darah, berolahraga sebanyak 4 kali seminggu dapat menurunkan resiko terkena serangan jantung. Olahraga juga dapat menurunkan tekanan darah

5. Stress/Emosi berlebihan

Stress dapat menyebabkan penyempitan arteri dan menurunkan aliran darah. Penyempitan arteri bahkan dapat terlihat pada penyakit ringan. Penelitian lain mengesankan bahwa stress berat dapat menyebabkan pecahnya dinding arteri yang dapat menyebabkan pemicunya penyakit jantung. Menghindari stress merupakan cara yang ampuh untuk mengurangi terjadinya penyakit jantung (Dudung, 2014)

### 2.3 SPECTF

*SPECTF* adalah pencitraan fungsional otak dengan tomografi emisi foton tunggal (*Single photon emission tomography/SPET*), atau dikenal dengan tomografi emisi foton tunggal terkomputerisasi (*Single photon emission computed tomography/SPECT*) yang memungkinkan melihat gambar tiga dimensi dari aliran darah serebral yang berasal dari data dua dimensi. Tomografi emisi positron ini menggunakan alat untuk mengukur metabolisme serebral regional dan karakteristik *neurotransmitter reseptor*. Secara sederhana, cara kerja pencitraan *SPECT* adalah dengan menyuntikkan pelacak radioaktif terlebih dahulu, kemudian dilakukan *scanning* dengan menggunakan kamera gamma sebanyak dua kali, yakni pada 10-15 menit setelah dilakukan penyuntikan, dan 2-3 jam sesudahnya. Dalam teknik tomografi dengan emisi terdapat keterbatasan fundamental yang harus diperhatikan, yaitu :

1. *collection efficiency* yaitu radiasi gamma yang dipancarkan ke segala arah lapisan, namun hanya yang masuk ke detektor yang dipakai untuk pencitraan. Oleh karenanya, efisiensinya sangat terbatas, kecuali bila pasien dapat dikelilingi oleh detektor.
2. *Atenuasi* yaitu penyederhanaan yang dilakukan untuk menjumlahkan pencacahan dari dua detector yang saling berhadapan dengan detector yang lain, namun terkadang tidak diperlukan dalam *SPECT*
3. Waktu koleksi yaitu bagian waktu yang ada saat radiasi gamma dipancarkan dan dibentuk dengan foton yang terbatas (Reviangga, 2013).

Cara kerjanya dengan mengambil data secara terus menerus dengan mengelilingi pasien dengan kamera dari berbagai sudut yaitu pada saat kamera berhenti pada suatu sudut tertentu maupun bergerak (*step and shoot acquisition*). Gerakan kamera dilakukan dari atas dan bawah berbarengan dengan gerakan 180°. (Edi, 2012)

## 2.4 Logika Fuzzy

*Fuzzy* bisa diartikan samar, jadi logika fuzzy dapat diartikan logika yang samar. nilai “*true*” dan “*false*” tergantung pada bobot keanggotaan yang dimilikinya, derajat keanggotaan mempunyai rentang antara 0 hingga 1, sangat berbeda jika dibandingkan dengan logika digital yang mempunyai dua keanggotaan 0 atau 1 saja pada satu waktu. Logika *fuzzy* sering digunakan untuk mengekspresikan suhu dalam ruangan apakah ruangan itu termasuk hangat, panas atau dingin. Dan sering digunakan untuk mengekspresikan nilai yang diterjemahkan dalam bahasa (*linguistic*). Pada tahun 1965, logika *fuzzy* dipublikasikan oleh seorang professor dari *University of California* bernama Lotfi A Zadeh di kota Berkeley. Alasan mengapa orang menggunakan logika *fuzzy* antara lain.

1. mudah dimengerti karena menggunakan konsep matematis yang sederhana.
2. Fleksibel
3. Mampu memodelkan fungsi non linier
4. Memiliki toleransi terhadap data yang tidak benar/tepat
5. Dapat dipadukan dengan teknik kendali secara konvensional (Kusumadewi, 2010).

## 2.4. Data Mining

### 2.4.1. Pengertian Data Mining

Data *mining* adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam *database*. Selain berguna untuk mencari informasi yang ada di dalam perusahaan, data *mining* juga bisa digunakan untuk mengambil keputusan berdasarkan sifat perilaku bisnis, analisa data *mining* dapat melebihi system pendukung keputusan tradisional karena data *mining* mengeksplorasi basis data untuk menemukan pola pola yang tersembunyi yang terlupakan oleh pelaku bisnis (Kusnawi, 2007).

### 2.4.2. Proses Data Mining

Prosedur yang umumnya digunakan untuk permasalahan data *mining* yaitu tahap-tahap sebagai berikut (Kantardzic, 2003):

1. Menentukan permasalahan dan merumuskan hipotesis  
Proses ditentukannya variabel-variabel dan hipotesis awal.
2. Mengumpulkan data  
Proses pengumpulan data yang nanti akan diolah.
3. *Preprocessing* data  
Proses pembersihan terhadap *outlier*, penanganan *missing value* maupun transformasi data.
4. Memperkirakan model  
Proses memilih teknik data *mining* yang sesuai untuk masalah yang ada.
5. Menafsirkan model dan menarik kesimpulan  
Proses penafsiran model untuk pengambilan keputusan.

Atribut cenderung memiliki nilai dengan rentang yang sangat bervariasi. Misalnya, dalam menentukan jarak antara dua *record*, atribut dengan rentang

nilai yang besar, memiliki lebih banyak pengaruh dalam menentukan jarak daripada atribut dengan rentang nilai yang kecil.

## 2.5 Klasifikasi Data Mining

Menurut (Kusnawi, 2007) menjelaskan bahwa klasifikasi adalah suatu fungsionalitas data *mining* yang akan menghasilkan model untuk memprediksi kelas atau kategori dari objek-objek di dalam baris data. Klasifikasi merupakan proses yang terdiri dari dua tahap, yaitu tahap pembelajaran dan tahap pengklasifikasian. Pada tahap pembelajaran, sebuah algoritma klasifikasi membangun sebuah model klasifikasi dengan cara menganalisis data *training*. Tahap pembelajaran dapat juga dipandang sebagai tahap pembentukan fungsi atau pemetaan  $y = f(x)$  dimana  $y$  adalah kelas hasil prediksi dan  $x$  adalah *tuple* yang ingin diprediksi kelasnya.

Selanjutnya, pada tahap pengklasifikasian, model yang telah dihasilkan digunakan untuk melakukan klasifikasi terhadap *unknown* data. Akan tetapi, sebuah model hanya boleh digunakan untuk klasifikasi jika akurasi model tersebut cukup tinggi. Akurasi dapat diketahui dengan cara menguji model tersebut dengan data *test*. Data *test* terdiri dari label kelas yang sudah diketahui, namun data *test* tidak boleh sama dengan data *training* karena menyebabkan pengujian tersebut menunjukkan akurasi yang tinggi, padahal belum tentu demikian.

## 2.6 Definisi k-Nearest Neighbor

Algoritma *k-Nearest Neighbor* (KNN) adalah Metode klasifikasi objek yang berdasarkan data yang jaraknya paling dekat dengan objek tersebut, dimana data diproyeksikan ke ruang berdimensi dan direpresentasikan fitur datanya, kemudian dibagi berdasarkan klasifikasi data. Bila titik pada ruangan ditandai kelas  $b$  dan jika kelas  $b$  merupakan klasifikasi yang banyak ditemui pada  $k$  tetangga yang paling dekat dengan titik tersebut. Dekat atau jauhnya biasanya dihitung dengan jarak Euclidean.

Berikut ini adalah keuntungan dari metode KNN :

- Sederhana penggunaannya.
- Dapat menangani data training yang mengandung noise.
- Efektif jika data training besar.

Kelebihan dari algoritma KNN itu sendiri adalah sederhana dan mudah diimplementasikan. Algoritma ini mencari  $k$  *training record* (tetangga) yang memiliki jarak terdekat dari *record* baru, untuk memprediksi kelas dari *record* baru tersebut (Yanita, 2013). Menurut (Mordian, 2009), KNN juga merupakan algoritma yang sering digunakan untuk klasifikasi pada teknik data *mining*.

### 2.6.2. Proses K-Nearest Neighbor (KNN)

cara kerja dari *K-Nearest Neighbor* (KNN) adalah mencari jarak yang paling dekat dengan tetangga dalam data latih. Kemudian dihitung jarak terdekat antara data uji dengan data latih terlebih dahulu. Untuk menghitung jarak digunakan fungsi jarak *euclidean* yang ditunjukkan oleh persamaan 2.1

$$d_i = \sqrt{\sum_{i=1}^p (x_{2i} - x_{1i})^2} \quad (2.1)$$

Keterangan:

$x_1$  = data latih

$x_2$  = data uji

$i$  = variabel data

$d$  = jarak

$p$  = dimensi data

Setelah diketahui jarak antar *record*, kemudian diambil sebanyak  $k$  tetangga terdekat untuk menentukan label kelas dari *record* baru menggunakan label kelas tetangga. (Mordian, 2009)

## 2.7. Fuzzy k-Nearest Neighbor

### 2.7.1. Definisi Fuzzy k-Nearest Neighbor

Fuzzy k-Nearest Neighbor (Fk-NN) merupakan metode klasifikasi yang menggabungkan teknik *fuzzy* dengan *k-Nearest Neighbor classifier*. Algoritma Fk-NN memberikan nilai keanggotaan kelas pada vektor sampel dan bukan menempatkan vektor pada kelas tertentu.

Sebuah data memiliki nilai keanggotaan yang berbeda pada setiap kelas dengan nilai interval  $[0, 1]$ . Teori himpunan *fuzzy* men-generalisasi teori k-NN klasik dengan mendefinisikan nilai keanggotaan data yang berbeda pada masing-masing kelas yang ada. Formula yang digunakan ditunjukkan oleh (James, 1985).

$$u_i(x) = \frac{\sum_{j=1}^k u_{ij} (||x - x_j||^{-2/(m-1)})}{\sum_{j=1}^k (||x - x_j||^{-2/(m-1)})} \quad (2.2)$$

$u_{ij}$  adalah nilai keanggotaan fuzzy pada contoh pengujian  $(x, x_j)$  yang merupakan satu dari contoh-contoh fuzzy k-nearest neighbor,  $k$  merupakan banyaknya nilai ketetanggaan terdekat yang diambil,  $j$  merupakan variable data untuk keanggotaan data latih,  $i$  merupakan variable data untuk keanggotaan data uji, sedangkan  $m$  merupakan berat kebalikan yang sebanding dengan jarak Antara  $x_j$  dan  $x$ . Variabel  $(m)$  merupakan penentuan seberapa banyak pemberian bobot pada jarak saat menghitung kontribusi jarak kedekatan pada masing-masing tetangga dengan nilai keanggotaan. Jika nilai  $m$  adalah dua, maka jarak kontribusi dari setiap titik tetangga (data latih) dibobotkan oleh nilai kebalikan dari jarak titik tetangga tersebut dengan titik yang sedang diklasifikasikan (data uji). Ketika nilai  $m$  naik, titik-titik tetangga tersebut dibobotkan lebih merata dan efek dari jarak relatif dari titik yang sedang diklasifikasikan akan berkurang.

Nilai  $m$  mendekati satu ketika semakin dekat tetangga maka akan dibobotkan lebih besar daripada tetangga yang lebih jauh (semakin besar nilai jarak maka semakin besar bobotnya), yang mana hal ini akan mempengaruhi pengurangan jumlah titik (tetangga) yang berkontribusi terhadap nilai keanggotaan dari titik yang sedang diklasifikasikan. Hasil yang ditampilkan pada jurnal ini, menggunakan nilai  $m=2$  tetapi perhatikan bahwa hampir tingkat kesalahan yang diperoleh pada data ini hampir sama dengan menggunakan nilai  $m$  yang beragam (James, 1985).

Nilai  $u_{ij}$  pada  $u_i(x)$  terlebih dahulu diproses dengan menggunakan persamaan 2.3 (James,1985).

$$U_{ij} = \begin{cases} 0.51 + (n_j/K) * 0.49, & \text{jika } j=i \\ (N_j/K) * 0.49, & \text{jika } j \neq i \end{cases} \quad (2.3)$$

Keterangan:

$n_j$ = jumlah anggota kelas  $j$  pada suatu dataset  $K$

$K$ = total data latih yang digunakan

$j$ = kelas target (training/tidak training)

### 2.7.2. Proses Fuzzy K-Nearest Neighbor

Tahapan proses yang dilakukan pada algoritma *Fuzzy K-Nearest Neighbor* sebagai berikut.

1. Menghitung jarak antara dua *record* menggunakan *Euclidean distance* yang ditunjukkan oleh persamaan 2.3.
2. Menghitung nilai keanggotaan  $u_i(x)$  menggunakan persamaan 2.3 untuk setiap  $i$ , dimana  $1 \leq i \leq C$ ,  $C$  adalah jumlah kelas.
3. Mengambil nilai terbesar dari proses nomer 3 untuk semua  $1 \leq i \leq C$ ,  $C$  adalah jumlah kelas.
4. Memberikan label kelas baru pada proses nomer 4.

### 2.8 Perhitungan Akurasi

Akurasi merupakan seberapa dekat suatu angka hasil pengukuran terhadap angka sebenarnya (*true value atau reference value*). Di dalam penelitian ini akan menghitung jumlah diagnosis yang akan dibagi oleh data yang ada, kemudian akurasi yang diperoleh akan dihitung dengan persamaan 2.4 (Dany, 2006)

$$\text{Akurasi (\%)} = \frac{\sum \text{data uji benar}}{\sum \text{total data uji}} \times 100\% \quad (2.4)$$

Jumlah prediksi benar adalah jumlah *record* data uji yang diprediksi kelasnya menggunakan metode klasifikasi dan hasilnya sama dengan kelas sebenarnya. Sedangkan jumlah total prediksi adalah jumlah keseluruhan *record* yang diprediksi kelasnya (seluruh data uji). Berikut ini hasil akurasi dari berbagai macam metode yang pernah digunakan untuk mengklasifikasi penyakit jantung sebagai berikut:

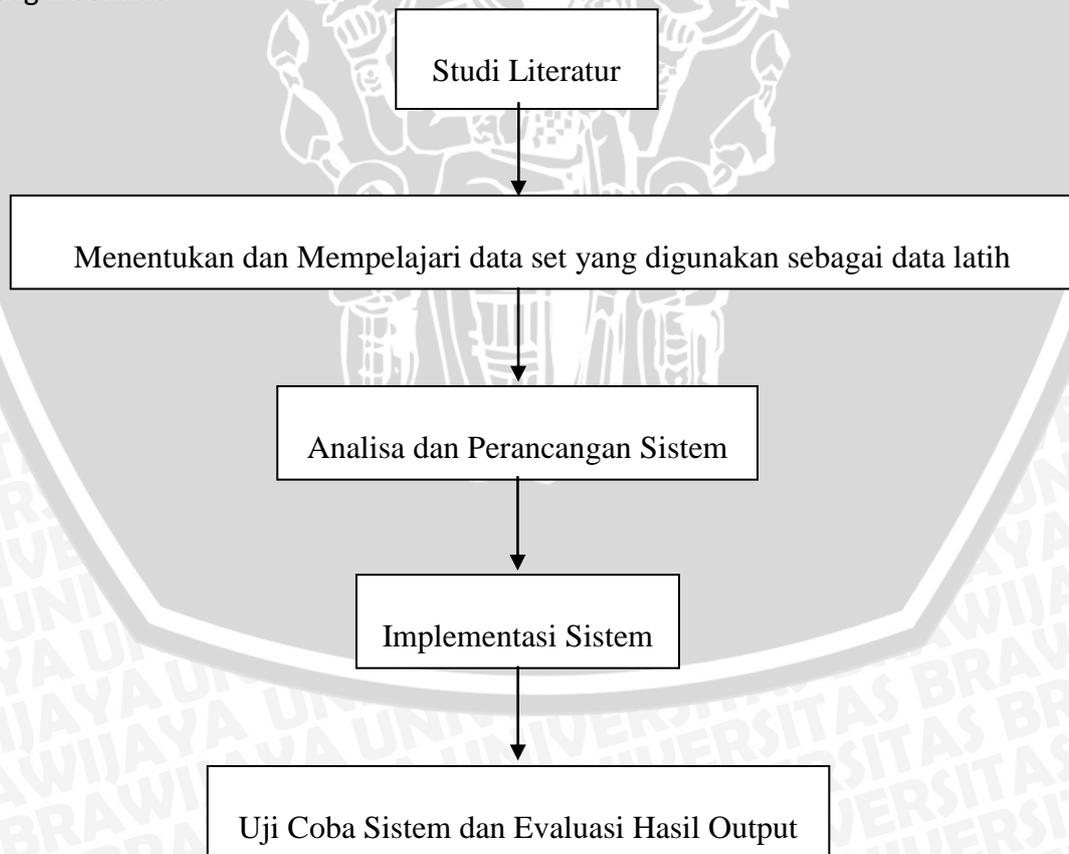
### BAB III METODOLOGI

Dalam bab metodologi dan perancangan ini akan dibahas metode perancangan dan langkah-langkah yang dilakukan dalam penelitian menggunakan metode *Fuzzy k-Nearest Neighbor* untuk diagnose penyakit Jantung.

Penelitian dilakukan dengan tahap-tahap sebagai berikut :

1. Mempelajari Studi literature yang terkait dengan masalah penyakit *Jantung* dan metode *Fuzzy k-Nearest Neighbor*.
2. Melakukan pengumpulan data – data dari *Single proton emission computer tomography (SPECTF)*
3. Menganalisis sistem dari data dan melakukan perancangan sistem yang meliputi pelatihan, dan pengujian.
4. Mengimplementasikan sistem untuk membuat perangkat lunak berdasarkan analisis dan perancangan yang telah dilakukan dengan menggunakan bahasa pemrograman Java.
5. Melakukan uji coba terhadap perangkat lunak.
6. Mengevaluasi uji coba terhadap perangkat lunak.

Langkah metode *Fuzzy k-Nearest Neighbor* dapat dilihat pada gambar 3.1 sebagai berikut.



Gambar 3.1 Langkah Metode *Fuzzy K-Nearest Neighbor*

### 3.1 Studi Literatur

Dalam penelitian ini dibutuhkan studi literatur yang dibutuhkan untuk merealisasikan tujuan dan penyelesaian masalah. Teori – teori mengenai penyakit Jantung, algoritma *fuzzy k-nearest neighbor* digunakan sebagai dasar penelitian yang berasal dari buku – buku, browsing dari internet dan jurnal, baik jurnal nasional maupun jurnal internasional. Kemudian data yang diperoleh diubah sehingga dapat digunakan untuk analisis. Setelah dianalisis maka dapat diimplementasikan ke dalam program.

### 3.2 Data Penelitian

Data yang digunakan dalam penerapan *fuzzy k-nearest neighbor* diambil dari situs <https://archive.ics.uci.edu/ml/datasets/SPECTF+Heart> dan menggunakan datapenyakit Jantung pada tahun 2012. Pada *dataset* Jantung, atribut yang digunakan adalah F1R,F1S,F2R,F2S,F3R,F3S,F4R,F4S,F5R,F5S,F6R,F6S,F7R,F7S,F8R,F8S,FF9R,F9S,F10R,F10S,F11R,F11S,F12R,F12S,F13R,F13S,F14R,F14S,F15R,F15S,F16R,F16S,17R,F17S,F18R,F18S,F19R,F19S,F20R,F20S,F21R,F21S,F22R,F22S. Sedangkan kelas output yaitu:

1. = Normal.
2. = Abnormal.

#### 3.2.1 Deskripsi Umum Sistem

Secara umum sistem yang dibuat merupakan perangkat lunak yang mengimplementasikan algoritma *Fuzzy k-Nearest Neighbor* (Fk-NN) untuk mengklasifikasikan pasien menjadi 2 kategori yaitu normal dan abnormal. Perangkat lunak ini akan menguji keakuratan hasil klasifikasi *dataset* penderita Jantung. Parameter uji yang berkaitan dengan nilai  $k$  (tetangga) dan data latih yang berpengaruh terhadap tingkat akurasi.

#### 3.2.2 Batasan Sistem

Batasan sistem prediksi penderita Jantung ini antara lain:

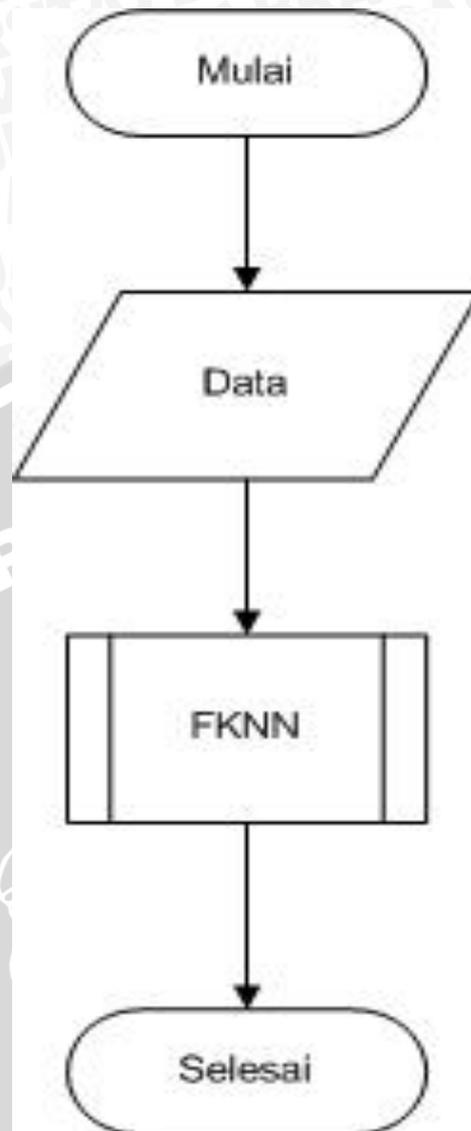
1. Input sistem berupa data penderita Jantung selama tahun 2012

#### 3.2.3 Perancangan Sistem

Tahap prediksi sistem FK-NN untuk prediksi penderita Jantung, langkah – Langkahnya adalah sebagai berikut:

1. Memasukan data latih dan data uji penderita Jantung dan  $k$ .
2. Menghitung Euclidean.
3. Menghitung FK-NN
4. Didapatkanlah hasil prediksi kelas penderita Jantung.

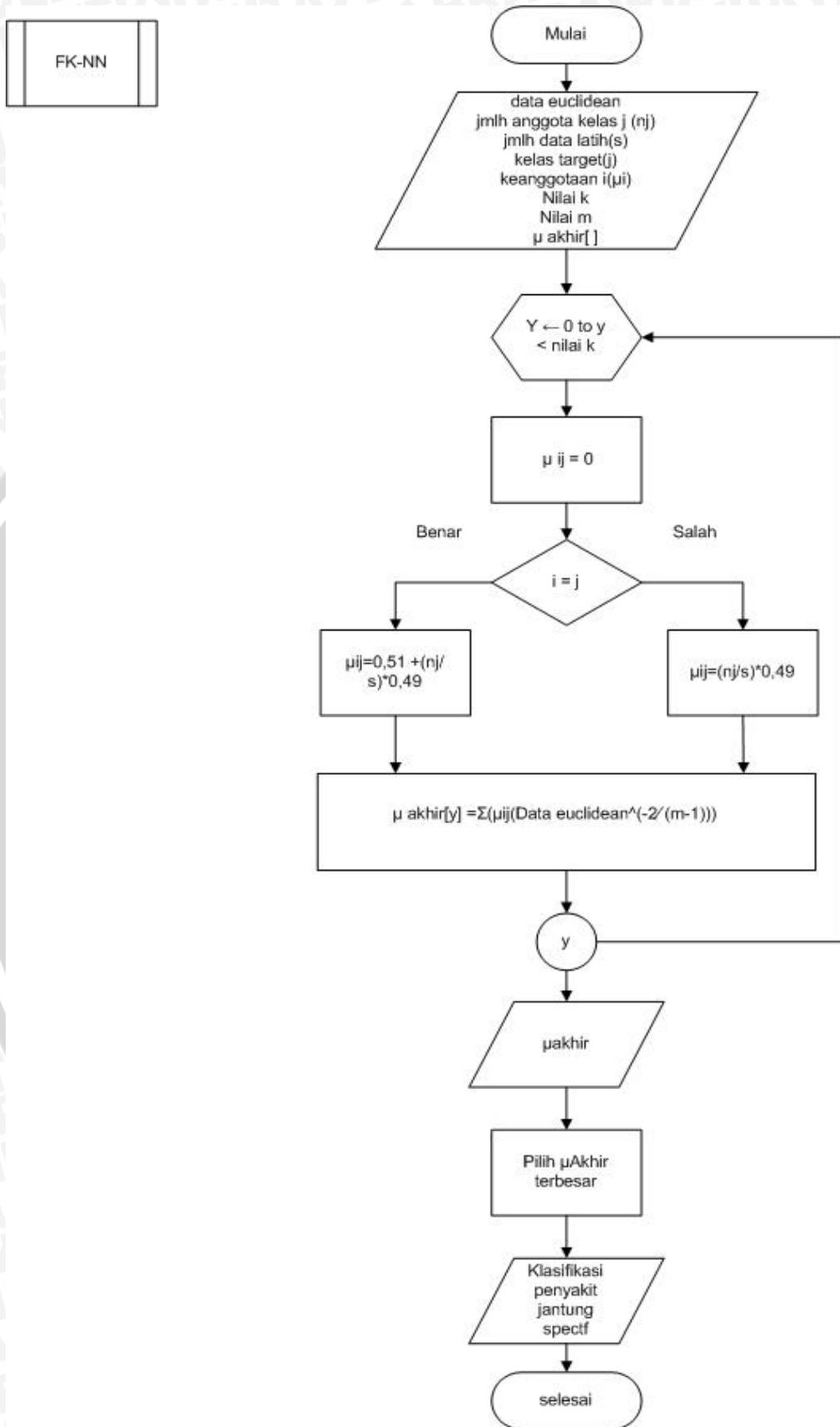
Tahapan proses klasifikasi dapat dilihat pada gambar 3.2:



Gambar 3.2 Alur proses klasifikasi

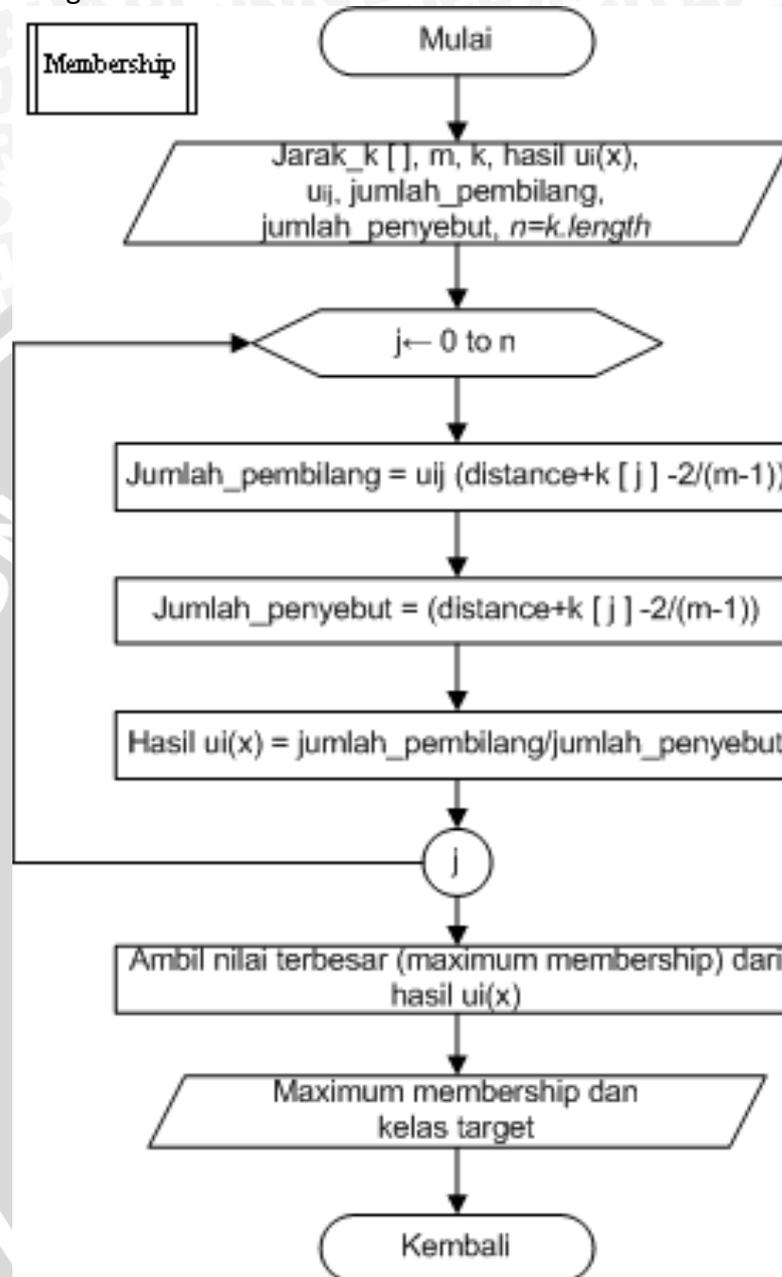
#### 3.2.4 Proses Fuzzy K-Nearest Neighbor

Tahapan ini merupakan proses perhitungan nilai jarak kedekatan tetangga data uji terhadap data latih menggunakan *Euclidean Distance*. Flowchart dari proses *Fuzzy k-nearest neighbor* ditunjukkan oleh gambar 3.3 :



Gambar 3.3 Proses Fuzzy K-nearest neighbor

Sedangkan flowchart untuk proses perhitungan membership ditunjukkan oleh gambar 3.4 sebagai berikut.



Gambar 3.4 Proses perhitungan membership

### 3.2.5 Contoh Perhitungan Manual

Pada subbab ini menampilkan perhitungan manual untuk proses diagnosa penyakit Jantung dan digunakan seluruh atribut dataset tersebut yang berjumlah 44 atribut. Data yang diambil sebanyak 17 record dengan rincian record ke-1 sampai 14 merupakan data latih dan record ke-15 sampai 17 merupakan data uji.

### 3.2.6 Data Latih dan Data Uji pada data penderita Jantung

Pada contoh perhitungan kali ini digunakan 14 data latih dan 3 data uji. Data latih dan data uji yang dipakai seperti yang ditunjukkan pada tabel (3.1):

Tabel (3.1) Data latih sistem:

KELAS	F1R	F1S	F2R	F2S	F3R	F3S	F4R	F4S	F5R	F5S	F6R	F6S	F7R	F7S
1	59	52	70	67	73	66	72	61	58	52	72	71	70	77
1	72	62	69	67	78	82	74	65	69	63	70	70	72	74
1	71	62	70	64	67	64	79	65	70	69	72	71	68	65
1	69	71	70	78	61	63	67	65	59	59	66	69	71	75
1	70	66	61	66	61	58	69	69	72	68	62	71	71	71
1	57	69	68	75	69	74	73	71	57	61	72	74	73	69
1	69	66	62	75	67	71	72	76	69	70	66	69	71	80
0	62	67	64	70	59	58	67	74	60	66	68	68	73	71
0	62	67	68	70	65	70	73	77	69	70	69	73	71	74
0	59	68	69	67	69	59	78	73	66	65	77	73	74	66
0	75	75	70	77	67	75	75	75	67	66	74	73	68	72
0	77	79	79	77	74	76	76	81	65	68	66	66	74	73
0	68	64	74	80	76	72	78	75	67	64	75	80	78	77
0	76	73	74	76	60	69	76	76	68	69	78	79	57	62

F8R	F8S	F9R	F9S	F10R	F10S	F11R	F11S	F12R	F12S	F13R	F13S	F14R	F14S	F15R
66	65	67	55	61	57	68	66	72	74	63	64	56	54	67
70	71	72	75	66	65	73	78	74	79	74	69	69	70	71
61	61	73	71	75	74	80	74	54	47	53	37	77	68	72
65	58	60	55	62	59	67	66	74	74	64	60	57	54	70
63	59	74	75	70	69	83	77	73	70	41	37	39	40	58
61	58	60	55	71	62	79	70	77	71	65	63	69	55	61
66	64	71	77	65	61	72	67	71	69	65	57	69	65	68
60	63	64	74	64	65	74	77	69	73	59	58	58	67	65
71	71	76	75	66	67	73	73	70	74	63	67	58	68	66
66	55	71	66	69	68	75	73	80	79	69	65	69	66	68
64	70	76	70	67	63	74	75	72	68	69	68	75	69	71
72	68	67	73	63	62	72	67	76	69	68	64	64	61	69
66	64	67	67	70	60	78	82	70	68	63	60	64	60	54
69	69	67	66	73	69	80	81	58	68	75	69	73	70	58

F15S	F16R	F16S	F17R	F17S	F18R	F18S	F19R	F19S	F20R	F20S	F21R	F21S	F22R	F22S
54	76	74	65	67	66	56	62	56	72	62	74	74	64	67
69	72	70	62	65	65	71	63	60	69	73	67	71	56	58
59	72	68	60	60	73	70	66	65	64	55	61	41	51	46
73	69	76	62	64	61	61	66	65	72	73	68	68	59	63
46	75	73	65	66	67	69	70	66	70	64	60	55	49	41
68	75	74	63	64	63	58	69	67	79	77	72	70	61	65
65	76	73	63	64	69	70	72	72	69	68	70	73	63	59
69	78	76	61	62	64	67	72	74	71	71	71	69	66	61
69	78	79	69	70	71	73	72	71	73	77	72	76	64	66
65	75	71	59	61	65	64	73	71	81	75	74	65	69	66
74	75	76	63	70	71	69	66	63	70	73	66	68	58	59
68	73	75	70	66	64	70	70	70	73	76	79	73	65	63
56	70	73	59	65	55	58	50	51	73	70	69	65	42	41
65	79	76	74	71	66	64	65	62	78	68	75	68	62	60

**Keterangan :**

0 = Pasien Jantung Normal

1 = Pasien Jantung Abnormal

Dari 14 data tersebut, memiliki 2 buah kelas sesuai dengan nilai target output, dimana ke-2 kelas tersebut adalah kelas pasien Jantung normal dan kelas pasien Jantung abnormal.

Berikut ini merupakan table data uji :

Tabel (3.2) Data uji system

Kelas	F1R	F1S	F2R	F2S	F3R	F3S	F4R	F4S	F5R	F5S	F6R	F6S
1	67	68	73	78	65	63	67	60	63	62	71	68

F7R	F7S	F8R	F8S	F9R	F9S	F10R	F10S	F11R	F11S	F12R	F12S
76	73	56	61	62	56	74	73	78	76	79	79

F13R	F13S	F14R	F14S	F15R	F15S	F16R	F16S	F17R	F17S	F18R	F18S
70	70	68	67	65	67	76	75	63	61	61	56

F19R	F19S	F20R	F20S	F21R	F21S	F22R	F22S
76	75	74	77	76	74	59	68



Menghitung jarak antara record baru pada data uji dengan tiap record pada data latih. Dari data uji dengan data pelatihan, dicari terlebih dahulu jarak terdekat "Euclidean distance", dimana rumus untuk Euclidean distance dapat dilihat pada persamaan (2.3). Berikut ini adalah contoh, perhitungan nilai jarak terdekat dari *record* pertama data latih dengan record uji.

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

Keterangan:

$x_1$  = data latih

$x_2$  = data uji

$i$  = attribute ke -  $i$

$d$  = jarak

$p$  = dimensi data

UNIVERSITAS BRAWIJAYA



Jarak

$$\begin{aligned}
 & (((67 - 62)^2) + ((68 - 67)^2) \\
 & + ((73 - 64)^2) + ((78 - 70)^2) \\
 & + ((65 - 59)^2) + ((63 - 58)^2) \\
 & + ((67 - 67)^2) + ((60 - 74)^2) + \\
 & ((63 - 60)^2) + ((62 - 66)^2) \\
 & + ((71 - 68)^2) + (68 - 68)^2) \\
 & + ((76 - 73)^2) + ((73 - 71)^2) \\
 & + ((59 - 60)^2) + ((61 - 63)^2) + \\
 & ((62 - 64)^2) + ((56 - 74)^2) \\
 & + ((74 - 64)^2) + ((73 - 65)^2) \\
 & + ((78 - 74)^2) + ((76 - 77)^2) \\
 & + ((79 - 69)^2) + ((79 - 73)^2) + \\
 & ((70 - 59)^2) + ((70 - 58)^2) \\
 & + ((68 - 58)^2) + ((67 - 67)^2) \\
 & + ((65 - 65)^2) + ((67 - 69)^2) \\
 & + ((76 - 78)^2) + ((75 - 76)^2) + \\
 & ((63 - 61)^2) + ((61 - 62)^2) \\
 & + ((61 - 64)^2) + ((56 - 67)^2) \\
 & + ((76 - 72)^2) + ((75 - 74)^2) \\
 & + ((74 - 71)^2) + ((77 - 71)^2) \\
 & ((76 - 71)^2) + ((74 - 69)^2) \\
 & + ((59 - 66)^2) + ((68 - 61)^2)) \\
 = & \sqrt{\left( \begin{aligned} & 1.036 + 1.0009 + 1 + 0.154 + 1.06 + 1.018 + 0 + 1.335 + 1.024 \\ & + 1.03 + 1.022 + 0 + 1.013 + 1.0059 + 1.0032 + 1.010 + 1.008 \\ & + 1.461 + 0.234 + 0.115 + 0.031 + 0.002 + 0.095 + 0.026 + 0.064 \\ & + 1.110 + 0.049 + 0 + 0.0037 + 1.028 + 1.025 + 1.005 + 1.007 \\ & + 1.0048 + 1.015 + 1.256 + 0.015 + 2.0011 + 0.018 + 0.060 \\ & + 1.044 + 1.016 + 1.046 + 1.045 \end{aligned} \right)} \\
 = & 54.5802
 \end{aligned}$$

Selanjutnya dihitung lagi jarak antara record uji dengan record data latih yang lain. Hasil perhitungan nilai euclidean ditampilkan pada tabel 3.3

Tabel (3.3) Hasil perhitungan data

Jarak euclidean					
KELAS	DATA UJI 1	KELAS	DATA UJI 2	KELAS	DATA UJI 3
0	42.9418	1	58.634461	1	54.0278
0	49.5681	1	47.89572	1	41.9404
0	37.5633	1	56.0803	1	70.8096
0	50.1398	1	43.104524	1	48.2804
0	51.3225	1	61.619802	1	77.6788
0	66.4455	1	49.112117	1	45.5522
0	55.3353	1	40.224371	1	36.1248
1	54.5802	0	37.56328	0	39.0896
1	52.2015	0	48.6621	0	33.2114
1	85.4985	0	40.442497	0	42.2256
1	41.4849	0	40.435133	0	34.6699
1	85.6738	0	50.705029	0	40.4475
1	34.9142	0	55.928526	0	61.1146
1	50.2295	0	50.049975	0	37.3898

Setelah hasil perhitungan euclidean didapatkan maka di ambil nilai yang terkecil. Hasil perhitungan jarak setelah diurutkan dari yang terkecil ditunjukkan pada tabel 3.4

Tabel (3.4) Data setelah diurutkan dari yang terkecil

data ke-	Nilai euclidean	Kelas	Urutan Terkecil
13	34.9142	1	1
3	37.5633	0	2
8	41.4849	1	3
1	42.9418	0	4
2	49.5681	0	5
4	50.3225	0	6
14	50.2295	1	7
5	51.3225	0	8
9	52.2015	1	9
8	54.5820	1	10
7	55.3353	0	11
6	66.4455	0	12
11	85.4985	1	13
13	85.6738	1	14

### 3.2.7 Menentukan k record terdekat

Berdasarkan hasil perhitungan euclidean pada tabel 3.3, kemudian dilakukan pengurutan terhadap euclidean yang ditunjukkan pada tabel 3.4. kemudian diambil k record yang terkecil. Apabila ditentukan  $k = 3$ , maka record yang terpilih adalah record ke-10, 13, dan 8.

Tabel 3.5 seleksi 3 data dengan nilai terkecil

data ke-	Nilai	Kelas	Urutan Terkecil
13	34.9142	1	1
3	37.5633	0	2

11	41.4849	1	3
----	---------	---	---

3.2.8 Menentukan maximum membership dan kelas target

Proses menentukan maximum membership dan kelas target dengan mencari nilai membership untuk tiap kelas j dengan menggunakan persamaan 2.3. K=14 (jumlah data latih),  $u_0$ =pasien Jantung normal,  $u_1$ =pasien Jantung,  $n_0=7$ ,  $n_1=7$

$$U_{ij} = \begin{cases} 0.51 + (n_j/K) * 0.49, & \text{jika } j=i \\ (N_j/K) * 0.49, & \text{jika } j \neq i \end{cases} \quad (2.6)$$

Keterangan:

- $n_j$ = jumlah anggota kelas j pada suatu dataset K
- K= total data latih yang digunakan
- j= kelas target (training/tidak training)

$$\begin{aligned} U_{1(1)} &= 0.51 + (7/14) * 0.49 \\ &= 0.51 + 0.5 * 0.49 \\ &= 0.51 + 0.245 \\ &= 0.755 \end{aligned}$$

$$\begin{aligned} U_{1(0)} &= (7/14) * 0.49 \\ &= 0.5 * 0.49 \\ &= 0.245 \end{aligned}$$

$$\begin{aligned} U_{0(0)} &= 0.51 + (7/14) * 0.49 \\ &= 0.51 + 0.5 * 0.49 \\ &= 0.51 + 0.245 \\ &= 0.755 \end{aligned}$$

$$\begin{aligned} U_{0(1)} &= (7/14) * 0.49 \\ &= 0.5 * 0.49 \\ &= 0.245 \end{aligned}$$

Setelah didapatkan membership untuk tiap kelas j dilanjutkan dengan mencari nilai keanggotaan data pada kelas dengan menggunakan persamaan 2.3

$$\begin{aligned} u_1 &= \frac{(0.245 * (43.9142^{\frac{2}{2-1}})) + (0.744 * (37.5633^{\frac{2}{2-1}})) + (0.245 * (41.4849^{\frac{2}{2-1}}))}{0.00088} \\ &= \frac{0.00088}{0.00211} \end{aligned}$$



$$= 0.41629$$

$$u_2 = \frac{0.00123}{0.00211} = 0.58371$$

Berdasarkan hasil perhitungan nilai keanggotaan didapat dua nilai keanggotaan, untuk menentukan kelas target maka dipilih nilai keanggotaan terkecil yaitu 0.00088 sehingga kelas targetnya yaitu 1 (bukan penderita jantung normal). Jadi, status diagnosa pasien tersebut adalah penderita jantung normal.

Setelah melakukan pengujian data uji langkah yang sama dilakukan terhadap data uji yang lain sehingga didapat akurasi sistem yang dihitung berdasarkan rumus (2.4)

Tabel 3.6 Perhitungan Akurasi Sistem

data uji ke-	Kelas Data	Kelas Perhitungan	Hasil	Hasil Prediksi
1	1	1	1	Benar
2	1	0	0	Salah
3	0	0	0	Benar

$$Akurasi (\%) = \frac{2}{3} \times 100\% = 66,66667 \%$$

### 3.3 Perancangan Antarmuka

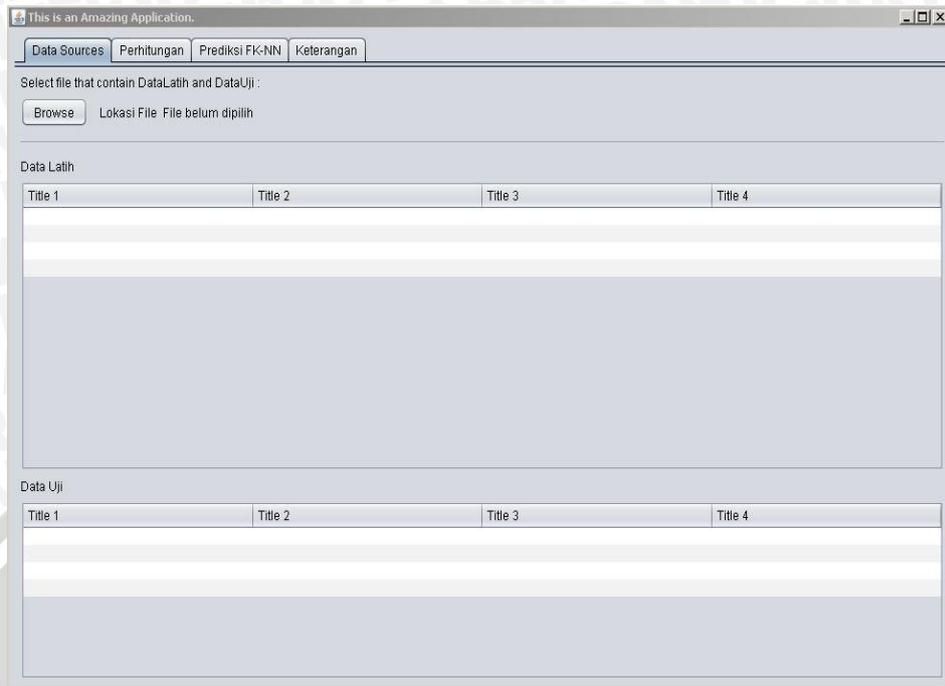
Antarmuka (*interface*) untuk menampilkan data yang digunakan sistem, terdiri dari 4 bagian, yaitu Data Source, Perhitungan, Prediksi FK-NN, dan Keterangan

Perancangan antarmuka sistem akan ditunjukkan pada gambar berikut ini.

#### 1. Tampilan awal Program :

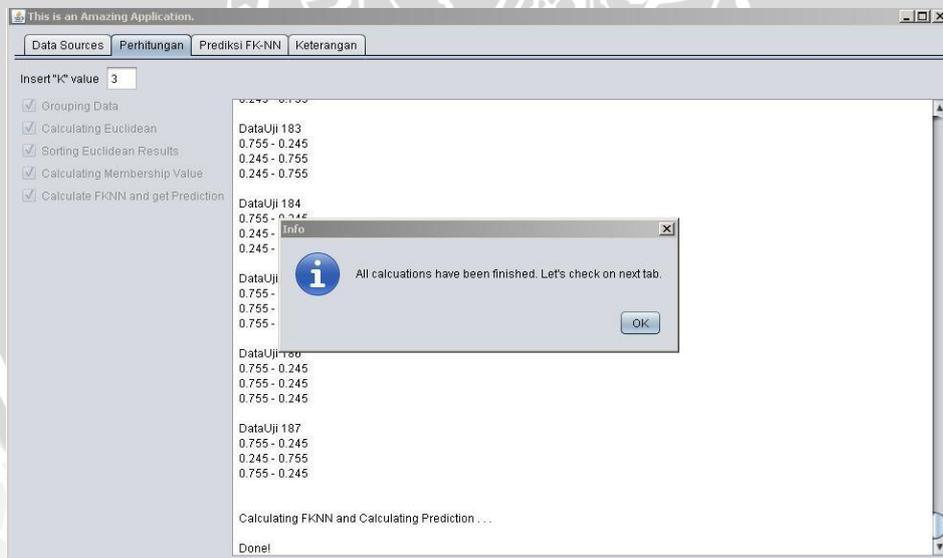
Didalam Data sources, tombol Browse digunakan untuk menginput data yang ingin dimasukan yaitu data latihan dan data uji





Gambar 3.5 Tampil Kolom Data Latih dan Data Uji

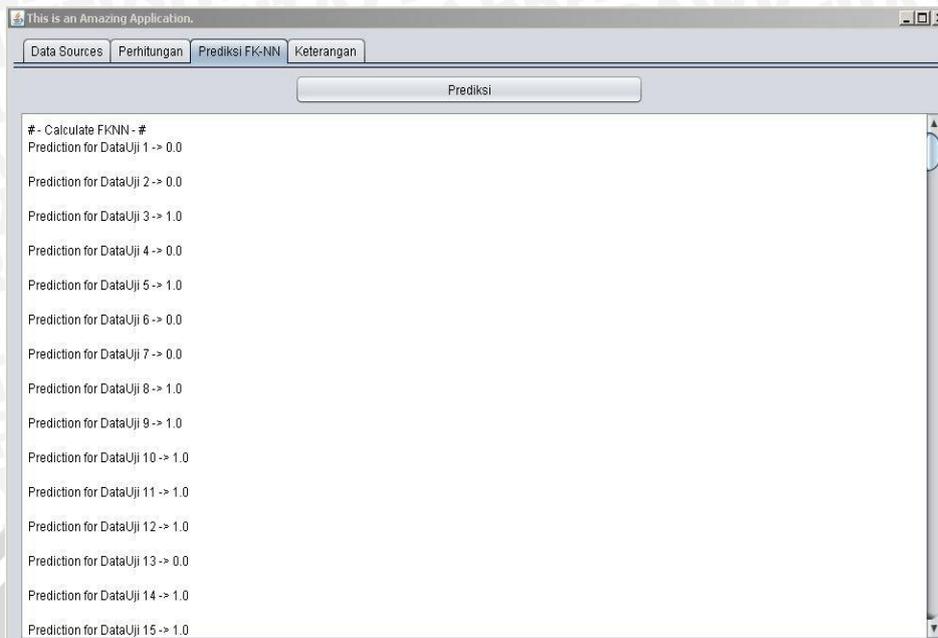
2. Masukan K didalam table perhitungan
3. Tekan tombol Hitung untuk mulai proses perhitungan



Gambar 3.6 Tabel Perhitungan

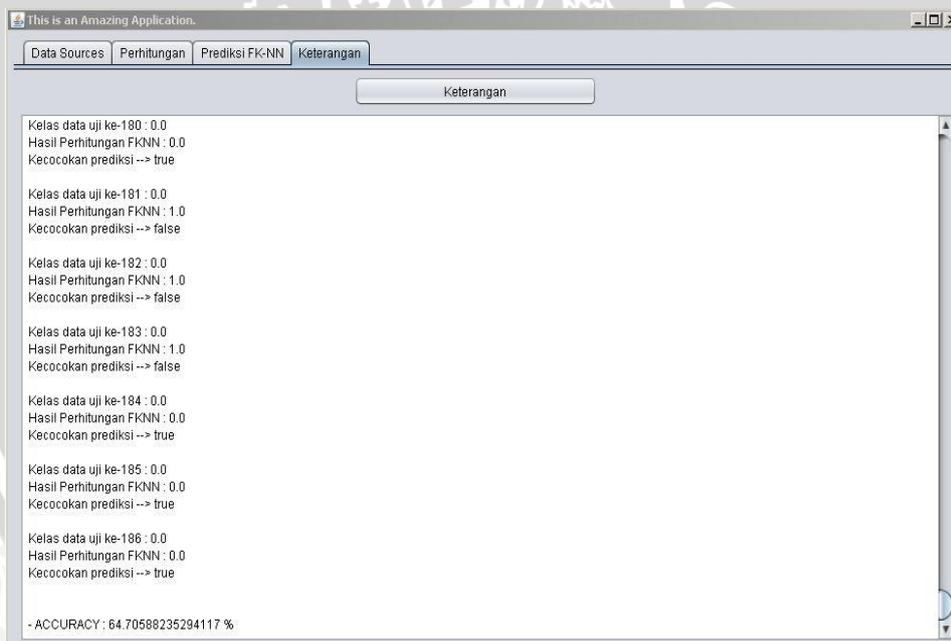
4. Tekan tombol Prediksi untuk menghitung FK-NN





Gambar 3.7 Tabel prediksi FKNN

5. Tekan tombol Keterangan untuk mengetahui Akurasi yang didapat.



Gambar 3.8 Tabel Keterangan

### 3.4 Perancangan Uji Coba

Setelah sistem selesai dibuat, langkah selanjutnya adalah melakukan pengujian terhadap sistem tersebut. Pengujian dilakukan untuk mengetahui tingkat akurasi dari hasil klasifikasi pada *SPECTF heart dataset* dengan menggunakan metode *Fuzzy k-Nearest Neighbor*.

Terdapat 1 macam pengujian yang dilakukan dalam penelitian ini, yaitu: pengujian untuk mengetahui pengaruh nilai  $k$  (jumlah tetangga terdekat) dan data latih terhadap tingkat akurasi klasifikasi.

### 3.5 Uji pengaruh nilai $k$ dan data latih terhadap tingkat akurasi klasifikasi

Uji pengaruh nilai  $k$  dan jumlah data latih terhadap tingkat akurasi klasifikasi dilakukan pada sejumlah data uji yang sama yaitu 20 data uji dengan beberapa data latih yang berbeda-beda. Dengan beberapa data latih yang berbeda-beda, tentunya akan mempengaruhi keakuratan hasil keputusan yang berpengaruh terhadap penentuan kelas. Dalam pengujian ini ditentukan nilai  $k = 2 \dots n$ . Dari data latih diambil sejumlah record, meliputi 20, 40, 60, dan 80 record. Dari pengujian ini, diperoleh tingkat akurasi terhadap nilai  $k$  dan jumlah data latih. Tabel 3.7 menampilkan rancangan tabel yang akan digunakan untuk mencatat hasil dari pengujian ini.

Tabel 3.7 Uji Pengaruh nilai  $k$  dan data latih terhadap tingkat akurasi

K	Jumlah Data Latih	Akurasi (%)
2		
...	...	
n	N	

Keterangan :

$K$  : nilai  $k$  yang akan diuji.

Jumlah Data Latih : jumlah data yang digunakan sebagai data latih.

Akurasi : tingkat akurasi dalam persen, yang dihitung dengan persamaan 2.6.

## BAB IV

### IMPLEMENTASI

#### 4.1 Lingkungan Implementasi

Dalam implementasi metode *Fuzzy K-nearest Neighbor (FK-NN)* untuk diagnosa penyakit jantung dibutuhkan beberapa aspek yang perlu diperhatikan yaitu segi perangkat keras dan perangkat lunak.

##### 4.1.1 Lingkungan Implementasi Perangkat Keras

Dalam mengembangkan sistem dan penerapan metode penelitian ini digunakan beberapa komponen perangkat keras sebagai berikut :

1. Processor : Intel® Core™ 2 Duo CPU @2.10GHz
2. Memory : 4.00 GB
3. Hard disk : 250 GB

##### 4.1.2 Lingkungan Implementasi Perangkat Lunak

Pengembangan sistem dan penelitian ini dibutuhkan beberapa perangkat lunak yang digunakan sebagai berikut ini :

1. Sistem Operasi yang digunakan Windows 7 Ultimate 64bit
2. Aplikasi pembangunan GUI dan code menggunakan NetBeans IDE 8.0.2
3. Bahasa pemrograman yang dipakai yaitu bahasa pemrograman java.
4. Komponen java yang digunakan yaitu JDK 7.5

#### 4.2 Implementasi Program

Berdasarkan metode penelitian dan perancangan proses yang terdapat dalam bab 3, maka pada sub bab ini akan dijelaskan implementasi proses-proses tersebut.

##### 4.2.1 Proses Baca File

Proses baca file digunakan untuk membaca file data yang akan digunakan dalam pelatihan dan pengujian. Tahapan proses baca file ditunjukkan pada source code 4.1

###### Proses Baca File

```
private void dataReader(){
    File file = new File (dataPath);
    try {

        Workbook workbook = jxl.Workbook.getWorkbook(file);
        Sheet []sheet = workbook.getSheets();

        int rowDataLatih = sheet[0].getRows();
        int columnDataLatih = sheet[0].getColumns();
    }
}
```

```

dls = new DataLatihStorage(rowDataLatih, columnDataLatih);

for(int i = 0; i<rowDataLatih; i++){
    for(int j=0; j<columnDataLatih; j++){
        dls.setValue(i, j, Double.parseDouble(sheet[0].getCell(j,
i).getContents()));
    }
}

int rowDataUji = sheet[1].getRows();
int columnDataUji = sheet[1].getColumns();

dus = new DataUjiStorage(rowDataUji, columnDataUji);

for(int i = 0; i<rowDataUji; i++){
    for(int j=0; j<columnDataUji; j++){
        dus.setValue(i, j, Double.parseDouble(sheet[1].getCell(j,
i).getContents()));
    }
}

catch (IOException | BiffException | NumberFormatException ex){
    System.out.println("Error while reading file. " + ex.getMessage());
}
}
}

```

Source Code 4.1 Baca File

#### 4.2.2 Proses FKNN

Proses FKNN merupakan proses perhitungan nilai membership .Source code FKNN dapat dilihat pada source code 4.5

```

Proses FKNN

protected void calculateEuclidean(){
    double value;
    double tmp;

    for(int i=0; i<dus.countRow(); i++){
        for(int j=0; j<dls.countRow(); j++){
            value = 0;
            for(int k=1; k<dls.countColumn(); k++){
                tmp = Math.pow(dus.getValue(i, k)-dls.getValue(j, k), 2);
                // System.out.print("(" + dus.getValue(i, k) + "-"
                "+dls.getValue(j, k) + ")^2 + ");
                value = value + tmp;
            }
        }
    }
}

```

```

    }
    euclideanData[i*2][j] = dls.getValue(j, 0); //insert "kelas"
value to euclideanData table
    euclideanData[(i*2)+1][j] = Math.sqrt(value); //insert
"euclidean" value to euclideanData table
    }
}

euclidean = euclidean.concat("##### Printing euclideanData
table #####\n");
for(int i=0; i<dus.countRow(); i++){
    for(int j=0; j<dls.countRow(); j++){
        euclidean = euclidean.concat("Kelas "+euclideanData[i*2][j]+"
- with data "+euclideanData[(i*2)+1][j]+" \n");
    }
    euclidean = euclidean.concat("\n");
}
}

////////// CONVERT OK! ////////////
protected void sortingEuclidean(){
    double minValue;
    double kelas;
    int maxIndex=0;
    boolean swap;

    //sort here
    for(int i=0; i<dus.countRow(); i++){
        for(int idx=0; idx<dls.countRow(); idx++){

            swap=false;

            minValue=euclideanData[(i*2)+1][idx];
            kelas=euclideanData[i*2][idx];

            for(int j=idx; j<dls.countRow(); j++){
                if(euclideanData[(i*2)+1][j]<minValue){
                    minValue=euclideanData[(i*2)+1][j];
                    kelas=euclideanData[i*2][j];
                    maxIndex=j;
                    swap=true;
                }
            }
        }

        if(swap){
            //swap Data

            euclideanData[(i*2)+1][maxIndex]=euclideanData[(i*2)+1][idx];
            euclideanData[(i*2)+1][idx]=minValue;

```

```

//swap Kelas
euclideanData[i*2][maxIndex]=euclideanData[i*2][idx];
euclideanData[i*2][idx]=kelas;
    }
    }
}

protected void setK_Value(int val){
    this.K_VALUE = val;
    MEMBERSHIP_TABLE = new double[2*dus.countRow()][K_VALUE];
}

////////// CONVERT OK! //////////
protected void membershipValue(){
    for(int i=0; i<dus.countRow(); i++){ // horizontal reading
        for(int j=0; j<K_VALUE; j++){ //vertical reading
            if(euclideanData[i*2][j]==0){

MEMBERSHIP_TABLE[i*2][j]=0.51+((jumlahDataNormal/(jumlahDataNormal+jumlahDataAbnormal))*0.49);

MEMBERSHIP_TABLE[(i*2)+1][j]=(jumlahDataNormal/(jumlahDataNormal+jumlahDataAbnormal))*0.49;
            }
            else{

MEMBERSHIP_TABLE[(i*2)+1][j]=0.51+((jumlahDataAbnormal/(jumlahDataNormal+jumlahDataAbnormal))*0.49);

MEMBERSHIP_TABLE[i*2][j]=(jumlahDataAbnormal/(jumlahDataNormal+jumlahDataAbnormal))*0.49;
            }
        }
    }

    membership = membership.concat("##### Printing MEMBERSHIP_TABLE #####\n");
    for(int i=0; i<dus.countRow(); i++){
        membership = membership.concat("DataUji "+(i+1)+"\n");
        for(int j=0; j<K_VALUE; j++){
            membership = membership.concat(MEMBERSHIP_TABLE[i*2][j]+" "+MEMBERSHIP_TABLE[(i*2)+1][j]+" \n");
        }
        membership = membership.concat("\n");
    }
}

```

```

}

////////// CONVERT OK! //////////

protected void calculateFKNN(){
    prediction = prediction.concat("# - Calculate FKNN -#\n");
    double tmpNormalValue, tmpAbnormalValue;
    for(int i=0; i<dus.countRow(); i++){

        tmpNormalValue = tmpAbnormalValue = 0;

        for(int j=0; j<K_VALUE; j++){ //calculate pembilang
            tmpNormalValue =
            tmpNormalValue+(MEMBERSHIP_TABLE[i*2][j]*(1/Math.pow(euclideanData[(i*2)+1][j], 2)));
            tmpAbnormalValue =
            tmpAbnormalValue+(MEMBERSHIP_TABLE[(i*2)+1][j]*(1/Math.pow(euclideanData[(i*2)+1][j], 2)));
        }
        FKNN_TABLE[0][i]=tmpNormalValue;
        FKNN_TABLE[1][i]=tmpAbnormalValue;

        tmpNormalValue = tmpAbnormalValue = 0;

        for(int j=0; j<K_VALUE; j++){ //calculate penyebut
            tmpNormalValue =
            tmpNormalValue+(1/Math.pow(euclideanData[(i*2)+1][j], 2));
            tmpAbnormalValue =
            tmpAbnormalValue+(1/Math.pow(euclideanData[(i*2)+1][j], 2));
        }
        FKNN_TABLE[2][i]=tmpNormalValue;
        FKNN_TABLE[3][i]=tmpAbnormalValue;

        // calculate FKNN result
        FKNN_TABLE[4][i]=FKNN_TABLE[0][i]/FKNN_TABLE[2][i];
        FKNN_TABLE[5][i]=FKNN_TABLE[1][i]/FKNN_TABLE[3][i];

        // get prediction result
        if(FKNN_TABLE[4][i]>FKNN_TABLE[5][i]){
            FKNN_TABLE[6][i]=0;
        }
        else{
            FKNN_TABLE[6][i]=1;
        }
        prediction = prediction.concat("Prediction for DataUji "+(i+1)+" -
> "+FKNN_TABLE[6][i]+" \n\n");
    }
}

```

Source code 4.5 FKNN

## BAB V PENGUJIAN DAN ANALISIS

Pada bab ini akan dilakukan pengujian dan analisis dari Penerapan *fuzzy k-nearest neighbor* (FK-NN) untuk mengklasifikasi penyakit jantung berdasarkan *spectf heart dataset*.

### 5.1 Pengujian

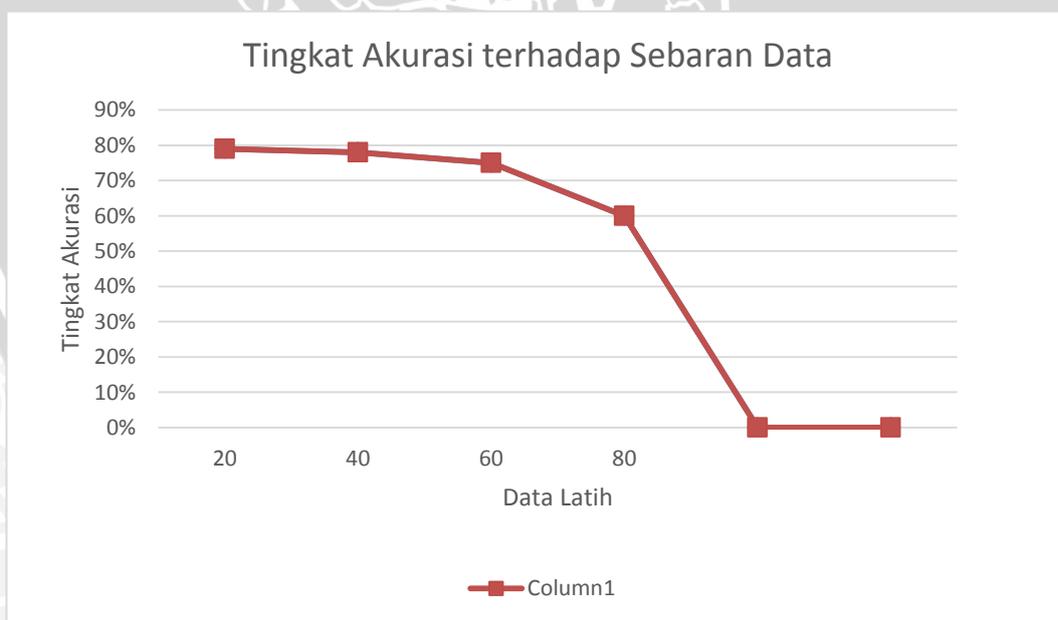
Pengujian dilakukan untuk mengetahui hasil akurasi dari implementasi yang dilakukan. Pengujian dilakukan dengan 3 macam jenis pengujian yaitu pengujian tingkat akurasi terhadap data latih, pengujian tingkat akurasi terhadap banyaknya nilai *k*, dan pengujian tingkat akurasi terhadap sebaran kelas.

#### 5.1.1 Pengujian tingkat akurasi terhadap data latih

Pada pengujian tingkat akurasi terhadap data latih, data latih yang digunakan meliputi 20,40,60 dan 80 data. Tabel 5.1 berikut ini menunjukkan pengaruh data latih terhadap tingkat akurasi, dengan nilai *k*=3.

Tabel 5.1 Hasil uji terhadap jumlah data latih

Data Latih	FKNN Akurasi (%)
20	79%
40	78%
60	75%
80	60%



Gambar 5.1 Grafik tingkat akurasi Data latih

Dari hasil yang terdapat pada tabel 5.1 dapat disimpulkan bahwa adanya fluktuatif di data latih membuat hasil akurasi tidak baik.

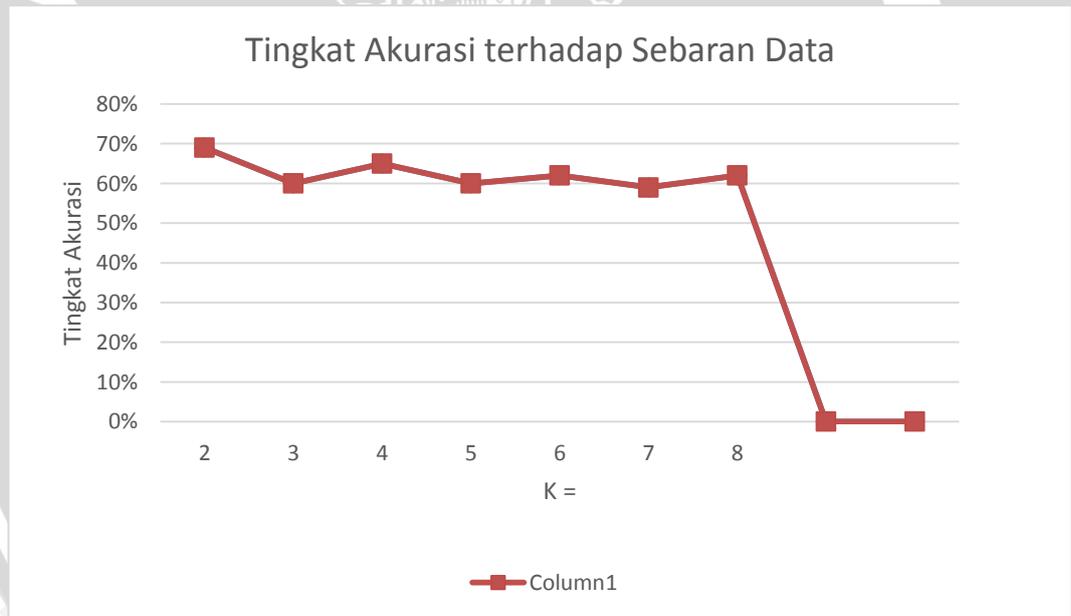


5.1.2 Pengujian tingkat akurasi terhadap nilai k

Pada pengujian tingkat akurasi terhadap pengaruh nilai k, data uji yang digunakan adalah 80 data. Sedangkan data latih yang digunakan adalah 188 data. Untuk nilai k yang digunakan adalah k = 2 sampai dengan k = n. Hasil dari pengujian tingkat akurasi terhadap nilai k dapat dilihat pada tabel 5.2.

Tabel 5.2 Hasil uji terhadap nilai k

K=	FKNN Akurasi (%)
2	69 %
3	60 %
4	65 %
5	60 %
6	62 %
7	59 %
8	62 %



Gambar 5.2 Grafik tingkat akurasi nilai K

5.1.3 Pengujian tingkat akurasi terhadap sebaran data

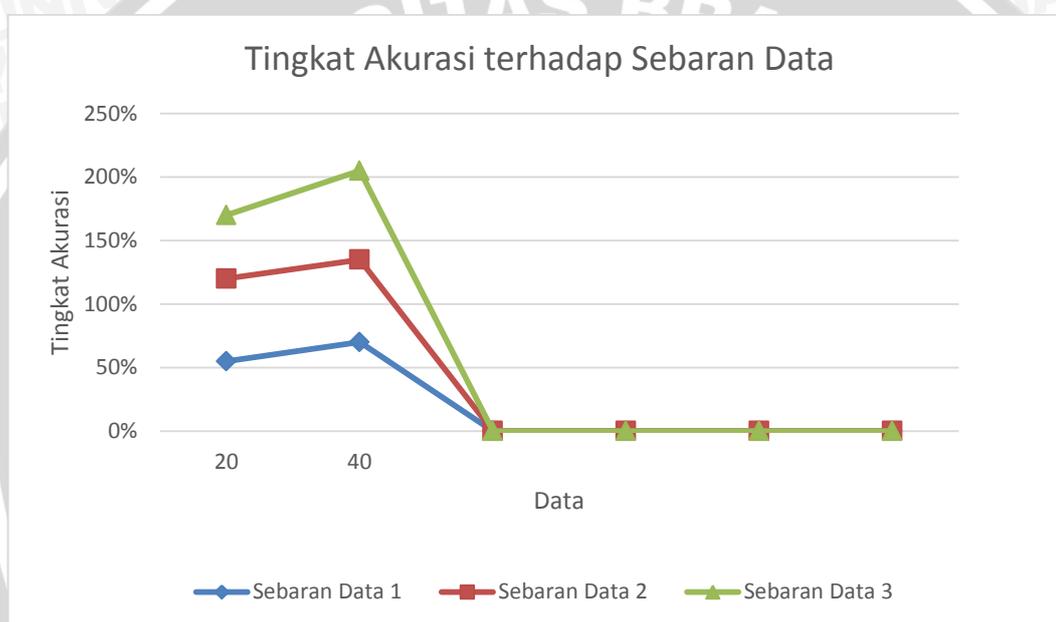
Pada pengujian tingkat akurasi terhadap sebaran data, data latih dan data uji yang digunakan meliputi 20, 40, 60, 80 data. Sebaran data 1 terdiri dari 70% kelas 1 dan 30% kelas 2, sebaran data 2 terdiri dari 30% kelas 1 dan 70% kelas 2, sedangkan sebaran data 3 terdiri dari 50% kelas 1 dan 50% kelas 2. Tabel 5.3 berikut ini menunjukkan pengaruh sebaran data terhadap tingkat akurasi dengan nilai k=3.

Tabel 5.3 Hasil uji terhadap sebaran data



Data Latih	Sebaran data 1 dengan 70% kelas 1 dan 30% kelas 2	Sebaran data 2 dengan 30% kelas 1 dan 70% kelas 2	Sebaran data 3 dengan 50% kelas 1 dan 50% kelas 2
20	55 %	65%	50%
40	70%	65%	70%

Berdasarkan pengujian tingkat akurasi terhadap data latih, nilai k dan sebaran kelas, maka dibuatlah grafik hubungan tingkat akurasi Grafik ditunjukkan pada grafik 5.3.



Gambar 5.3 Grafik Tingkat Akurasi terhadap Sebaran Data

## 5.2 Analisis

Pada sub bab ini akan dilakukan analisa dari pengujian dan analisis dari Penerapan *fuzzy k-nearest neighbor* (FK-NN) untuk mengklasifikasi penyakit jantung berdasarkan *spectf heart dataset*. Analisis ini dilakukan untuk mengetahui karakteristik data dari pengujian yang telah dilakukan.

### 5.2.1 Analisis hasil pengujian tingkat akurasi terhadap data latih

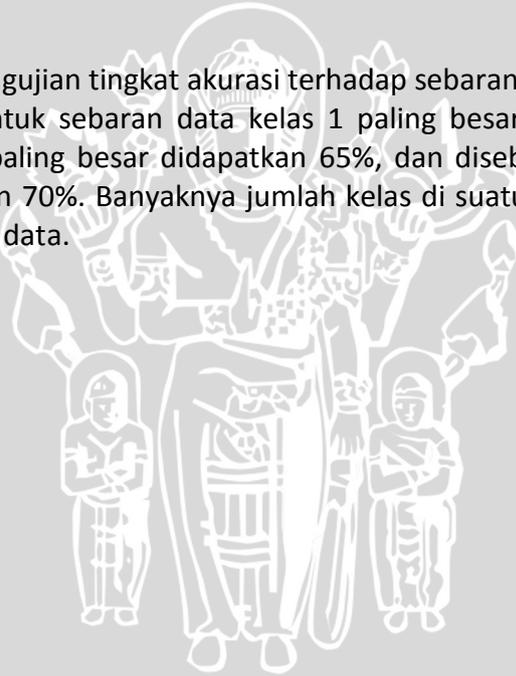
Pada pengujian tingkat akurasi terhadap data latih dapat disimpulkan bahwa hasil akurasi yang paling bagus sebesar 79% dalam data latih yang sebanyak 20 data. Dari hasil yang ada pada tabel 5.2 dapat disimpulkan bahwa adanya fluktuatif di data latih membuat hasil akurasi tidak baik.

### 5.2.2 Analisis hasil pengujian tingkat akurasi terhadap nilai k

Pada pengujian tingkat akurasi terhadap nilai k dapat disimpulkan bahwa hasil akurasi terbaik sebesar 69%. Akurasi belum tentu semakin baik dengan banyaknya K.

### 5.2.3 Analisis hasil pengujian tingkat akurasi terhadap sebaran data

Rata – rata akurasi untuk sebaran data kelas 1 paling besar didapatkan 70%, sebaran data kelas 2 paling besar didapatkan 65%, dan disebaran data kelas 3 paling besar didapatkan 70%. Banyaknya jumlah kelas di suatu data latih sangat mempengaruhi akurasi data.



## BAB VI

### PENUTUP

#### 6.1 Kesimpulan

Berdasarkan hasil penelitian untuk mengetahui tingkat akurasi dari hasil klasifikasi pada *Single Proton Emission Computed Tomography (SPECT)* dengan menggunakan metode *Fuzzy k-Nearest Neighbor (FKNN)*, dapat disimpulkan bahwa:

1. Metode *Fuzzy k-Nearest Neighbor (FK-NN)* bisa diterapkan untuk mendiagnosis penderita *Jantung* dengan menggunakan 44 parameter yang terdapat pada data *Single Proton Emission Computed Tomography (SPECT)*. Langkah pertama adalah memasukan data latih dan data uji kemudian dihitung menggunakan persamaan *euclidean distance dan membership*. Output data penderita *Jantung* yang digunakan ditransformasikan kedalam bentuk *fuzzy*. Pada akhirnya didapatkan hasil diagnose penderita *Jantung* yang dihitung menggunakan persamaan *Fuzzy k-Nearest Neighbor (FK-NN)*.
2. Berdasarkan pengujian tingkat akurasi terhadap data latih dihasilkan akurasi paling besar yaitu 79% dalam 187 data uji, terhadap nilai k dihasilkan akurasi sebesar 69%, terhadap sebaran data kelas 1 akurasi sebesar 70%, kelas 2 akurasi sebesar 65%, dan kelas 3 paling besar didapatkan 70%.

#### 6.2 Saran

saran yang dapat dijadikan bahan pertimbangan untuk penelitian selanjutnya adalah sebagai berikut:

1. Penelitian ini tidak melakukan pengujian terhadap optimasi k dan data latih. Diharapkan pada penelitian selanjutnya untuk meneliti optimasi k dan data latih.
2. Diharapkan pada penelitian berikutnya menggunakan metode lain atau FKNN yang telah dimodifikasi agar mendapat hasil yang lebih maksimal.