

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

Hadistria Massalli¹, M. Tanzil Furqon², S.kom, M.CompSc², Budi Darma Setiawan, S.Kom, M.Cs³

1)Mahasiswa, 2,3)Dosen Pembimbing

Program Studi Informatika / Ilmu Komputer, Fakultas ilmu computer, Universitas Brawijaya

Gedung A PTIIK Lt.1, Jl.Veteran No 8, Malang, 65145, Indonesia

Telp: +62-341-577911; Fax: +62-341-577911

Email : ahadistria@gmail.com , m.tanzil.furqon@gmail.com , s.budidarma@ub.ac.id

Abstrak

Penelitian ini membahas tentang algoritma FKNN (Fuzzy K-nearest Neighbor) untuk mendignosa penderita penyakit jantung. Penelitian ini dibangun menggunakan data pasien penderita penyakit jantung(SPECT) yang diambil dari <https://archive.ics.uci.edu/ml/datasets/SPECTF+Heart>. SPECTF (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,SPECTF

1.Pendahuluan

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 keseluruh 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). Didalam tubuh, paru paru membersihkan darah yang dipompa oleh jantung ke seluruh tubuh. Fungsi jantung secara umum yaitu memompa darah ke seluruh tubuh dan menampungnya kembali setelah dibersihkan oleh paru-paru. 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)

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).

2. Single Proton Emission Computer Tomography (SPECT)

Single Proton Emission Computer Tomography (SPECT) adalah Salah satu teknik dari dunia kedokteran nuklir yang cara kerjanya melihat proses darah mengalir melalui pembuluh darah di jantung, sehingga dapat mendeteksi secara rinci pada bagian yang tersumbat serta kerja jantung secara keseluruhan. 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.

3. 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. Fk-NN menggunakan metode klasifikasi yang digunakan untuk memprediksi data uji menggunakan nilai derajat keanggotaan

data uji pada setiap kelas. Kemudian diambil kelas dengan nilai derajat keanggotaan terbesar dari data uji sebagai kelas hasil prediksi. Keuntungannya adalah nilai-nilai keanggotaan vektor seharusnya memberikan tingkat jaminan pada hasil klasifikasi.

Proses Fuzzy K-Nearest Neighbor

Tahapan proses yang dilakukan pada algoritma *Fuzzy K-Nearest Neighbor* adalah:

1. Menghitung jarak antara dua *record* menggunakan *Euclidean distance* yang ditunjukkan oleh persamaan 2.3.
2. Menghitung nilai keanggotaan $u_i(x)$ menggunakan persamaan 2.5 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.

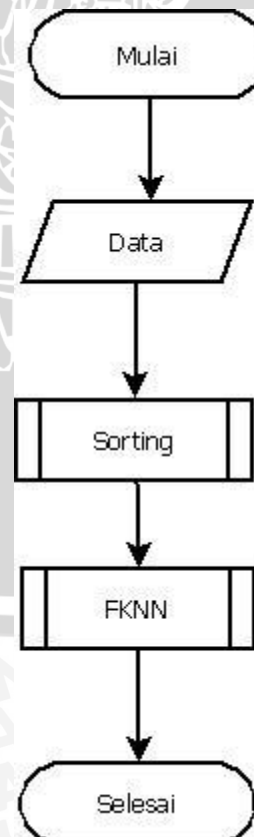
Perhitungan Akurasi

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 yaitu :

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

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).

4. Perancangan system



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

4.1. 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

DATA KELAS	F1R	F1S	F2R	F2S	F3R	F
1	59	52	70	67	73	
1	72	62	69	67	78	
1	71	62	70	64	67	
1	69	71	70	78	61	
1	70	66	61	66	61	
1	57	69	68	75	69	
1	69	66	62	75	67	
0	62	67	64	70	59	
0	62	67	68	70	65	
0	59	68	69	67	69	
0	75	75	70	77	67	
0	77	79	79	77	74	
0	68	64	74	80	76	
0	76	73	74	76	60	
1	67	68	73	78	65	
1	75	74	71	71	62	
0	70	72	70	70	65	

Keterangan:

x_1 = data latih

x_2 = data uji

i = attribute ke - i

d = jarak

p = dimensi data

$$\begin{aligned} & ((0.571 * ((0.5 - 0.25)^2)) + (0.682 * ((0.593 - 0.556)^2))) \\ & ((0.563 * ((0.667 - 0.667)^2)) + (0.616 * ((0.875 - 0.375)^2))) \\ & + (0.435 * ((0.316 - 0)^2)) + (0.435 * ((0.208 - 0)^2)) \\ & + (0.643 * ((0 - 0)^2)) + (0.755 * ((0 - 0.667)^2)) + \\ & (0.6 * ((0.4 - 0.2)^2)) + (0.743 * ((0.5 - 0.7)^2)) \\ & + (0.652 * ((0.563 - 0.375)^2)) + (0.510 * ((0.143 - 0.143)^2)) \end{aligned}$$

Hasil = 54.5802

Selanjutnya dihitung lagi jarak antara record uji dengan record data latih yang lain

Jarak euclidean

KELAS	DATA UJI 1
0	42.9418
0	49.5681
0	37.5633
0	50.1398
0	51.3225
0	66.4455
0	55.3353
1	54.5802
1	52.2015
1	85.4985
1	41.4849
1	85.6738
1	34.9142
1	50.2295

Keterangan :

Kelas : Pasien Jantung normal= 0; Pasien Jantung abnormal = 1

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.

Dari data diatas kemudian dihitung Euclidean dengan rumus :

Setelah hasil perhitungan euclidean didapatkan maka di ambil nilai yang

Terkecil. Hasil perhitungan jarak setelah diurutkan dari yang terkecil

4.2 Menentukan k record terdekat

Berdasarkan hasil perhitungan euclidean pada tabel 3.8, kemudian dilakukan pengurutan terhadap euclidean yang ditunjukkan pada tabel 3.9. kemudian diambil k record dengan berat terbesar.

Apabila ditentukan k = 3

Menentukan k record terdekat dengan seleksi 3 data dengan nilai terkecil

Nilai	kelas	Urutan Terkecil
34.9142	1	1
37.5633	0	2
41.4849	1	3

Nilai	kelas	Urutan terkecil
38	0	1
40	1	2
40.43513	0	3

Nilai	kelas	Urutan terkecil
33.2114	0	1
34.6699	0	2
36.1249	1	3

4.3 Menentukan maximum membership dan kelas target

Proses menentukan maximum membership dan kelas target dengan mencari nilai membership untuk tiap kelas j dengan menggunakan persamaan 2.5. $K=14$ (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

mencari nilai Membership				
K=14				
N0=7	N1=7	N0=NORMAL	N1=ABNORMAL	
Data Uji 1		0	1	
34.9142	1	0.245	0.755	
37.5633	0	0.755	0.245	
41.4849	1	0.245	0.755	

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

Setelah didapatkan membership untuk tiap kelas, dilanjutkan dengan mencari nilai keanggotaan sebuah data pada masing – masing kelas

$$u_1 = \left(0.245 * \left(43.9142 \cdot \frac{2}{2-1} \right) + \left(0.744 * \left(37.5633 \cdot \frac{2}{2-1} \right) + \left(0.245 * \left(41.4849 \cdot \frac{2}{2-1} \right) \right) \right)$$

Hasilnya adalah u_0 : 0.00088

$$u_2 = \left(0.755 * \left(34.9142 \cdot \frac{2}{2-1} \right) + \left(0.245 * \left(37.5633 \cdot \frac{2}{2-1} \right) + \left(0.755 * \left(41.4849 \cdot \frac{2}{2-1} \right) \right) \right)$$

Hasilnya adalah u_1 : 0.00123

Berikut ini adalah hasil dari semua kelas :

RUMUS FKNN		
PEMBILANG		
$\sum_{j=1}^k \mu_{ij} \left(\frac{1}{ x - x_j ^{2/m-1}} \right)$		
Data Uji 1	u(0)	0.00088
	u(1)	0.00123
DATA UJI 2	u(0)	0.00115
	u(1)	0.00079
DATA UJI 3	u(0)	0.0015
	u(1)	0.001

PENYEBUT		
$\sum_{j=1}^k \left(\frac{1}{ x - x_j ^{2/m-1}} \right)$		
	u(0)	0.00211
	u(1)	0.00211
	u(0)	0.00194
	u(1)	0.00194
	u(0)	0.0025
	u(1)	0.0025

HASIL		PREDIKSI
$\mu_i(x) = \frac{\sum_{j=1}^k \mu_{ij} \left(\frac{1}{ x - x_j ^{2/m-1}} \right)}{\sum_{j=1}^k \left(\frac{1}{ x - x_j ^{2/m-1}} \right)}$		
u(0)	0.41629	1
u(1)	0.58371	
u(0)	0.59239	0
u(1)	0.40761	
u(0)	0.59898	0
u(1)	0.40102	

Berdasarkan hasil perhitungan nilai keanggotaan didapat dua nilai keanggotaan, untuk menentukan kelas target maka dipilih nilai keanggotaan terbesar yaitu 0.58371 sehingga termasuk kedalam kelas 1

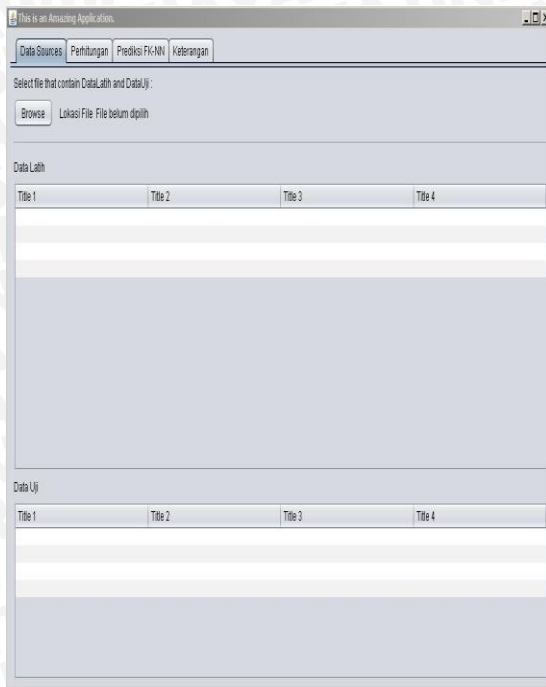
HASIL DENGAN RUMUS FKNN			
DATA UJI	KELAS	HASIL PERHITUNGAN	HASIL PREDIKSI
1	1	1	BENAR
2	1	0	SALAH
3	0	0	BENAR
TINGKAT AKURASI		66.6667	

Jadi hasil akurasi yang didapat sebesar 66,66%

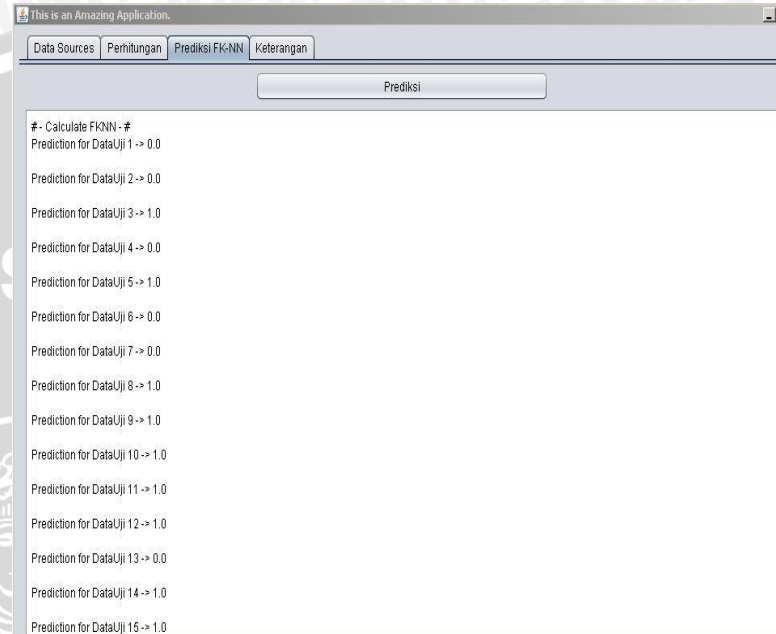
5. Perancangan Antarmuka

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.

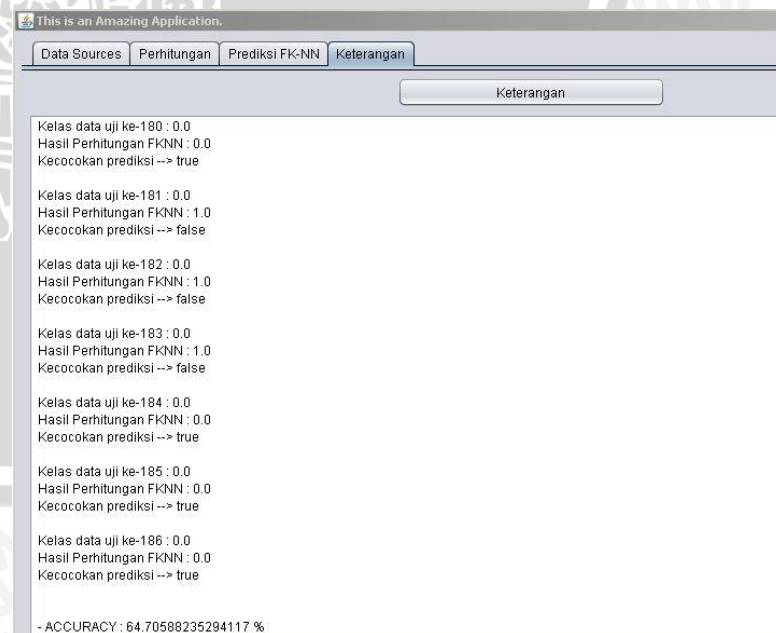
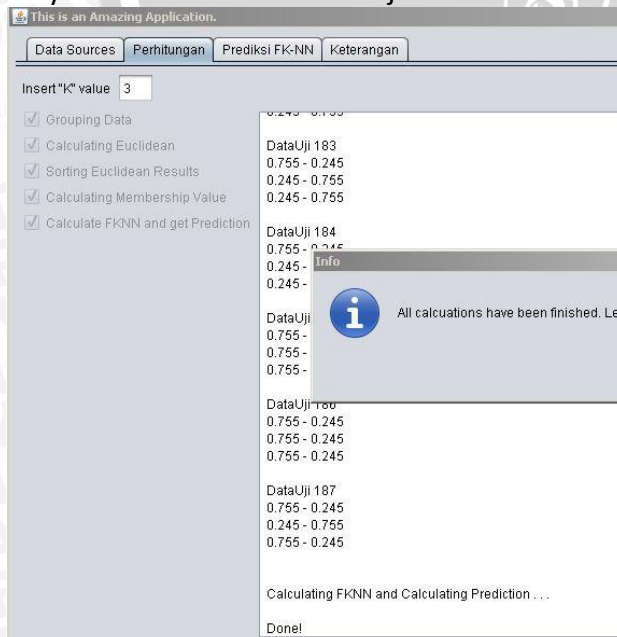


1. Masukan K didalam table perhitungan
2. Tekan tombol Hitung untuk mulai proses perhitungan



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

1. Tekan tombol Prediksi untuk menghitung FK-NN



1. Tekan tombol Keterangan untuk mengetahui Akurasi

6. 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.

6.1 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	64%
40	58%
60	49%
80	59%

Dari hasil diatas dapat disimpulkan bahwa adanya fluktuatif di data latih membuat hasil akurasi tidak baik.

6.2 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.

5.2
uji

K=	FKNN Akurasi (%)
2	60 %
3	59 %
4	59 %
5	57 %
6	57 %
7	56 %
8	58 %

Tabel
Hasil

terhadap nilai k

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*. 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 sebesar 64% dalam 184 data uji, terhadap nilai k dihasilkan akurasi sebesar 60%, terhadap sebaran data kelas 1 akurasi sebesar 49%, kelas 2 akurasi sebesar 79%, dan kelas 3 paling besar didapatkan 64%.

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.

DAFTAR PUSTAKA

- Abdul, Muhchid. 2006. *Pharmaceutical Care Untuk Pasien Penyakit Jantung Koroner : Fokus sindrom akut*.
- Adeli, A and. Mehdi, N. 2010 . A Fuzzy Expert System for Heart Disease Diagnosis. *The international Multiconference of engineers and Computer scientists 2010 vol 1, IMECS*
- Arsana, I, 2015. *Penyakit Jantung Koroner. Penyakit jantung*, [online] Tersedia di: <<http://www.penyakitjantung.net/penyakit-jantung-koroner> > [Diakses 11 november 2015]
- Dudung, 2014. *Pengertian Jantung Dan Fungsinya Pada Manusia*, [online] Tersedia di: <<http://www.dosenpendidikan.com/pengertian-jantung-dan-fungsinya-pada-manusia/>> [Diakses 11 november 2015]
- Guyton AC dan Hall JE, 2006. *Textbook of Medical Physiology 11th edition*. Elsevier Saunders.
- Hamid Parvin, 2008. MKNN : modified K-Nearest Neighbor. *Proceedings of the World Congress on Engineering and Computer Science 2008 WCECS 2008, October 22 - 24, 2008, San Francisco, USA*
- Hani Nurhayati, Fressy Nugroho2, 2012 . *Implementasi Fuzzy Expert System Untuk Diagnosis Penyakit Jantung*. UIN Maulana Malik Ibrahim Malang
- Hardika teguh, 2014. Penerapan Fuzzy K-Nearest Neighbor(FKNN) untuk mendiagnosa penderita liver berdasarkan *indian liver patient dataset (ILPD)* . Universitas Brawijaya Malang
- Hermansyah. 2012. *Aktifitas fisik dan kesehatan mental terhadap kejadian penyakit jantung koroner pada pasien rawat jalan di RSUP DR. Wahidin Sudirohusodo dan RSUD Labuang Baji Makassar*
- James, M, k, 1985. A Fuzzy K-Nearest Neighbor. *IEEE Transactions on System, Man and Cybernetics*, Vol. SMC-15 No. 4
- Keller. 1985. A Fuzzy K-Nearest Neighbor. *IEEE Transactions on System, Man and Cybernetics*, Vol. SMC-15 No. 4.
- Kusnawi. 2007. *Pengantar Solusi Data Mining* (Online). STMIK AMIKOM. Yogyakarta
- Kusumadewi, S dan Purnomo, H. 2010. *Aplikasi Logika Fuzzy untuk Pendukung Keputusan: Jilid 2*. Graha Ilmu. Yogyakarta.
- Moradian, Mehdi dan Baarani, Ahmad. 2009. *KNNBA: k-Nearest-Neighbor-Based_Association Algorithm*

(Online), <http://www.jatit.org/volume/researchpapers/Vol6No1/14Vol6No1.pdf>, diakses tanggal 27 November 2013.

Nila, Moeloek, 2014. *Penanganan penyakit jantung harus sesuai ilmu kedokteran terkini dan mengutamakan keselamatan pasien.*

Nils J. Nilsson, 1996. *Introduction to machine learning*, Robotics Laboratory department of computer science Stanford University California.

Agus Priyono, Muhammad Ridwan, Ahmad Jais Alias, Riza Atiq, O. K. Rahmat, Azmi Hassan & Mohd. Alauddin Mohd. Ali. 2005. Generation Of Fuzzy Rules With Subtractive Clustering, *Jurnal Teknologi*, 43(D) Dis. Universitas Teknologi Malaysia.

Turban, 2005. *Decision Support Systems and Intelligent Systems (Sistem Pendukung Keputusan dan Sistem Cerdas) jilid 1.* Andi Offset: Yogyakarta.

Yanita, C. 2013. Penerapan Fuzzy K-Nearest Neighbor Untuk Menentukan Status Evaluasi Kinerja Karyawan. Universitas Brawijaya.

Zhang, Juan, Yi Niu, Huabei Nie. 2009. Web Document Classification Based on Fuzzy KNN Algorithm. *International Conference on Computational Intelligence and Security.*

