

**PENERAPAN *FUZZY K-NEAREST NEIGHBOR* UNTUK MENENTUKAN
STATUS EVALUASI KINERJA KARYAWAN**

SKRIPSI



Disusun oleh :

YANITA CAHYANINGTYAS

NIM. 0810963075

**PROGRAM STUDI INFORMATIKA / ILMU KOMPUTER
PROGRAM TEKNOLOGI INFORMASI DAN ILMU KOMPUTER**

UNIVERSITAS BRAWIJAYA

MALANG

2013

**PENERAPAN *FUZZY K-NEAREST NEIGHBOR* UNTUK MENENTUKAN
STATUS EVALUASI KINERJA KARYAWAN**

SKRIPSI

Sebagai salah satu syarat untuk memperoleh
Gelar Sarjana dalam bidang Ilmu Komputer



Disusun oleh :

YANITA CAHYANINGTYAS

NIM. 0810963075

**PROGRAM STUDI INFORMATIKA / ILMU KOMPUTER
PROGRAM TEKNOLOGI INFORMASI DAN ILMU KOMPUTER**

UNIVERSITAS BRAWIJAYA

MALANG

2013

**PENERAPAN *FUZZY K-NEAREST NEIGHBOR* UNTUK MENENTUKAN
STATUS EVALUASI KINERJA KARYAWAN**

SKRIPSI

Sebagai salah satu syarat untuk memperoleh
Gelar Sarjana dalam bidang Ilmu Komputer



Disusun oleh :

YANITA CAHYANINGTYAS

NIM. 0810963075

Telah diperiksa dan disetujui oleh:

Dosen Pembimbing I,

Dosen Pembimbing II,

Drs. Achmad Ridok, M.Kom

NIP. 19680825 199403 1 002

Candra Dewi, S.Kom., MSc

NIP. 19771114 200312 2 001

**LEMBAR PENGESAHAN SKRIPSI
PENERAPAN *FUZZY K-NEAREST NEIGHBOR* UNTUK MENENTUKAN
STATUS EVALUASI KINERJA KARYAWAN**

SKRIPSI

Sebagai salah satu syarat untuk memperoleh
Gelar Sarjana dalam bidang Ilmu Komputer

Disusun oleh:

**YANITA CAHYANINGTYAS
NIM. 0810963075**

**Setelah dipertahankan di depan Majelis Penguji
pada tanggal 18 Januari 2013
dan dinyatakan memenuhi syarat untuk memperoleh
gelar Sarjana dalam bidang Ilmu Komputer**

Penguji,

Penguji,

**Lailil Muflikhah, S.Kom., M.Sc
NIP. 19741113 200501 2 001**

**Budi Darma Setiawan, S.Kom., M.Cs
NIK. 841015 06 1 1 0090**

Penguji,

**Dian Eka Ratnawati, S.Si., M.Kom
NIP. 19730619 200212 2 001**

**Mengetahui
Ketua Program Studi Teknik Informatika**

**Drs. Marji, M.T.
NIP. 19670801 199203 1 001**

PERNYATAAN ORISINALITAS SKRIPSI

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 dikutip dalam naskah ini dan disebutkan dalam sumber kutipan dan 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, Januari 2013

Mahasiswa,

Yanita Cahyaningtyas

NIM. 0810963075

KATA PENGANTAR

Puji syukur kehadiran Allah SWT, hanya dengan rahmat dan karunia yang telah diberikan kepada penulis, sehingga dapat menyelesaikan skripsi yang berjudul “Penerapan *Fuzzy K-Nearest Neighbor* Untuk Menentukan Status Evaluasi Kinerja Karyawan”.

Skripsi ini merupakan salah satu syarat untuk memenuhi persyaratan akademis untuk menyelesaikan studi di program Sarjana Ilmu Komputer Universitas Brawijaya.

Pada kesempatan ini penulis mengucapkan banyak terima kasih atas segala bantuan dan dedikasi moral maupun material dalam rangka penyusunan skripsi ini.

1. Drs. Achmad Ridok, M.Kom., selaku dosen pembimbing I yang telah membimbing dengan bijaksana dan sabar dalam membimbing dengan baik penyusunan skripsi ini.
2. Candra Dewi, S.Kom., MSc., selaku dosen pembimbing II yang telah membimbing dengan bijaksana dan sabar dalam membimbing dengan baik penyusunan skripsi ini.
3. Drs. Marji, M.T, selaku Ketua Program Studi Ilmu Komputer di Program Teknologi Informasi & Ilmu Komputer Universitas Brawijaya.
4. Wayan Firdaus Mahmudy, S.Si., M.T., selaku dosen pembimbing akademik atas nasehat, bimbingan, saran, dukungan yang diberikan selama penulis menuntut ilmu di Program Studi Ilmu Komputer.
5. Ir. Sutrisno, M.T., selaku Ketua Program Teknologi Informasi dan Ilmu Komputer Universitas Brawijaya.
6. Segenap Bapak dan Ibu dosen yang telah mendidik dan mengajarkan ilmunya kepada Penulis selama menempuh pendidikan di Program Teknologi Informasi & Ilmu Komputer Universitas Brawijaya.
7. Segenap staf dan karyawan di Program Teknik Informatika Program Teknologi Informasi & Ilmu Komputer Universitas Brawijaya yang telah banyak membantu Penulis dalam pelaksanaan penyusunan skripsi ini.

8. Setiyanto, S.E., M.M., selaku manajer HRD PT Sierad Produce,Tbk yang telah memberikan data untuk digunakan dalam skripsi ini.
9. Ayah, Ibu, dan adikku serta keluarga besarku yang tersayang, terima kasih atas dukungan dan doanya.
10. Teman saya Mbak Sari dan Tirana yang sering memberikan bantuan hingga penulis dapat menyelesaikan skripsi ini.
11. Teman saya Tika, Neny, Mutya, Rosi, Likha, Dinda, serta semua teman-teman Program Studi Ilmu Komputer 2008 Kelas A dan Kelas B atas segala bantuan, motivasi dan doanya.
12. Penghuni Bunga Andong 8A atas semua doa dan motivasinya.
13. Ervin, Winda, Eni, Nana atas dukungan dan doanya.
14. Seluruh pihak yang tidak dapat disebut secara langsung yang telah memberikan bantuan demi terselesaikannya skripsi ini.

Penulis menyadari bahwa masih banyak kekurangan dalam penulisan laporan ini yang disebabkan oleh keterbatasan kemampuan dan pengalaman. Oleh karena itu, Penulis sangat menghargai saran dan kritik yang sifatnya membangun demi perbaikan penulisan dan mutu isi penelitian ini.

Penulis berharap semoga laporan ini dapat memberikan manfaat kepada pembaca dan bisa diambil manfaatnya untuk pengembangan selanjutnya.

Malang, Januari 2013

Penulis

ABSTRAK

Yanita Cahyaningtyas. 2013. Penerapan *Fuzzy K-Nearest Neighbor* Untuk Menentukan Status Evaluasi Kinerja Karyawan.

Dosen Pembimbing: Drs. Achmad Ridok, M.Kom dan Candra Dewi, S.Kom., M.Sc.

Sumber daya manusia yang mempunyai kompetensi tinggi sangat diperlukan dalam suatu perusahaan. Dengan keahlian yang dimiliki akan menjadi pertimbangan bagi suatu perusahaan untuk melakukan evaluasi kinerja. Evaluasi kinerja karyawan dilihat dari aspek-aspek yang telah ditentukan perusahaan dimana karyawan yang dievaluasi akan mempunyai nilai evaluasi yang berbeda-beda untuk setiap aspeknya. Dengan nilai evaluasi yang berbeda-beda tersebut akan menjadi suatu masalah apabila dikaitkan dengan hasil keputusan yang akan berguna untuk mengembangkan sumber daya manusia dalam suatu perusahaan. Maka diperlukan suatu penggalan data untuk menentukan status evaluasi kinerja karyawan pada evaluasi kinerja karyawan. Dalam penelitian ini digunakan metode *Fuzzy K-Nearest Neighbor* untuk menentukan status evaluasi kinerja karyawan.

Proses yang dilakukan untuk menentukan status evaluasi kinerja karyawan meliputi dua tahap, yaitu tahap normalisasi nilai atribut dan tahap klasifikasi dengan algoritma *fuzzy k-nearest neighbor*. Tahap normalisasi dilakukan untuk menormalisasikan nilai atribut pada dataset. Setelah didapatkan nilai atribut yang ternormalisasi dilakukan tahap klasifikasi yang digunakan untuk mencari distance yang diperlukan untuk mendapatkan nilai derajat keanggotaan dimana dipilih nilai derajat keanggotaan terbesar yang menentukan didapatkannya kelas target.

Hasil penelitian menunjukkan bahwa semakin bertambahnya nilai tetangga terdekat maka akurasi yang didapatkan semakin menurun dan stabil pada nilai tetangga terdekat tertentu untuk masing-masing data latih. Sedangkan untuk nilai tetangga terdekat yang kecil dapat mendominasi dataset yang disebabkan oleh kelas yang dominan pada data tersebut.

Kata kunci : *Fuzzy K-Nearest Neighbor*, *Data Mining*, klasifikasi, evaluasi kinerja karyawan.

ABSTRACT

Yanita Cahyaningtyas. 2013. Application of *Fuzzy K-Nearest Neighbor* to Determine Employee Performance Evaluation Status.

Advisor: Drs. Achmad Ridok, M.Kom and Candra Dewi, S.Kom., M.Sc.

High competence human resources is needed in a company because of their expertise to support employee performance. It will be a consideration for a company to evaluate the performance. Employee performance evaluation is observed with the aspects of which have been defined by the company where the employee being evaluated will have a different evaluation value for each aspect. It will be an issue if it is linked with decision which will be useful to develop human resources in the company. Therefore, we need a data mining to determine the status of employee performance evaluation on employee performance evaluation. This study used the method of *Fuzzy K-Nearest Neighbor* to determine the status of employee performance evaluation

The process undertaken to determine the status of employee performance evaluation involves two phases, namely the value of attribute normalization stage and stage fuzzy classification algorithm k-nearest neighbor. Normalization phase done to normalize the attribute values in the dataset. Having obtained the attribute values are normalized performed classification stage which is used to find the distance needed to obtain the highest degree of membership which will be chosen the value that determines the degree of the largest membership class acquired target.

The results showed that the increasing value of the nearest neighbor obtained, the accuracy decreased and stabilized at a certain value of the nearest neighbors for each data. As for the value of the nearest neighbor can dominate small datasets due to the dominant class in the data.

Keywords : *Fuzzy K-Nearest Neighbor*, *Data Mining*, classification, employee performance evaluation.

DAFTAR ISI

HALAMAN SAMPUL.....	i
LEMBAR PERSETUJUAN.....	ii
LEMBAR PENGESAHAN SKRIPSI.....	iii
PERNYATAAN ORISINALITAS SKRIPSI.....	iv
KATA PENGANTAR.....	v
ABSTRAK.....	vii
ABSTRACT.....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xiii
DAFTAR SOURCECODE.....	xiv
DAFTAR LAMPIRAN.....	xv
BAB I PENDAHULUAN.....	1
1. 1 Latar Belakang.....	1
1. 2 Rumusan Masalah.....	3
1. 3 Batasan Masalah.....	3
1. 4 Tujuan Penelitian.....	3
1. 5 Manfaat Penelitian.....	3
1. 6 Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Sumber Daya Manusia.....	5
2.2 Kinerja Karyawan.....	5
2.3 Evaluasi Kinerja.....	6
2.4 Logika <i>Fuzzy</i>	6
2.5 Himpunan <i>Fuzzy</i>	7
2.6 <i>Data Mining</i>	8
2.6.1 Definisi <i>Data Mining</i>	8
2.6.2 Proses <i>Data Mining</i>	10
2.7 Klasifikasi.....	11

2.8	Algoritma <i>K-Nearest Neighbor</i>	11
2.8.1	Definisi <i>K-Nearest Neighbor</i>	11
2.8.2	Proses <i>K-Nearest Neighbor</i> (KNN)	12
2.9	<i>Fuzzy K-Nearest Neighbor</i>	13
2.9.1	Definisi <i>Fuzzy K-Nearest Neighbor</i>	13
2.9.2	Proses <i>Fuzzy K-Nearest Neighbor</i>	15
2.10	Data Evaluasi Kinerja Karyawan PT Sierad Produce, Tbk.....	15
2.11	Perhitungan Akurasi.....	15
 BAB III METODOLOGI DAN PERANCANGAN.....		16
3. 1	Studi Literatur.....	16
3. 2	Data Penelitian.....	17
3. 3	Deskripsi Umum Sistem.....	17
3. 4	Perancangan Sistem.....	17
3.4.1	Proses Klasifikasi Algoritma <i>Fuzzy K-Nearest Neighbor</i>	18
3.4.1.1	Flowchart Sistem.....	18
3.4.1.2	Proses Normalisasi Nilai Atribut.....	19
3.4.1.3	Proses Klasifikasi dengan Algoritma <i>Fuzzy K-Nearest Neighbor</i> ...	20
3.4.1.4	Proses Menentukan Maximum Membership dan Kelas Target.....	20
3. 5	Contoh Perhitungan Manual.....	22
3.5.1	Data Latih dan Data Uji pada data evaluasi kinerja karyawan.....	22
3.5.2	Melakukan proses normalisasi terhadap nilai atribut.....	24
3.5.3	Menghitung jarak antara record baru pada data uji dengan tiap record pada data latih.....	26
3.5.4	Menentukan k record terdekat.....	27
3.5.5	Menentukan maximum membership dan kelas target.....	27
3. 6	Perancangan Antarmuka.....	29
3. 7	Perancangan Uji Coba.....	33
3.7.1	Uji pengaruh nilai k dan data latih terhadap tingkat akurasi klasifikasi.....	33

BAB IV IMPLEMENTASI DAN PEMBAHASAN.....	35
4. 1 Lingkungan Implementasi.....	35
4. 1.1 Lingkungan Perangkat Keras.....	35
4. 1.2 Lingkungan Perangkat Lunak.....	35
4. 2 Implementasi Program.....	35
4. 2.1 Implementasi <i>class</i>	35
4. 2.2 Implementasi proses pelatihan.....	42
4. 2.2.1 Proses normalisasi nilai setiap atribut.....	42
4. 2.2.2 Proses klasifikasi dengan algoritma <i>fuzzy k-nearest neighbor</i>	43
4. 3 Penerapan Aplikasi.....	45
4. 3.1 Implementasi antarmuka.....	45
4. 3.1.1 Tabpage Tampil Data.....	45
4. 3.1.2 Tabpage Fuzzy Knn.....	46
4. 3.1.3 Tabpage Akurasi.....	47
4. 4 Implementasi Pengujian.....	48
4. 4.1 Hasil Uji.....	48
4. 5 Analisa Hasil.....	50
 BAB V KESIMPULAN.....	 53
5. 1 Kesimpulan.....	53
5. 2 Saran.....	53
 DAFTAR PUSTAKA.....	 54

DAFTAR GAMBAR

Gambar 3. 1 Diagram Alir Penelitian.....	16
Gambar 3. 2 Alur Proses Sistem.....	19
Gambar 3. 3 Flowchart Normalisasi Nilai Atribut.....	21
Gambar 3. 4 Flowchart Klasifikasi dengan Algoritma <i>Fuzzy K-Nearest Neighbor</i>	21
Gambar 3. 5 Flowchart Proses Menentukan Maximum Membership dan Kelas Target.....	22
Gambar 3. 6 Tampilan Rancangan Antarmuka untuk Menampilkan Data.....	30
Gambar 3. 7 Tampilan Rancangan Antarmuka Fuzzy Knn.....	32
Gambar 3. 8 Tampilan Rancangan Antarmuka Akurasi.....	33
Gambar 4. 1 Tampilan rancangan antarmuka tampil data.....	46
Gambar 4. 2 Tampilan rancangan antarmuka Fuzzy K-NN.....	46
Gambar 4. 3 Tampilan rancangan antarmuka Akurasi.....	47
Gambar 4. 4 Pengaruh nilai k terhadap tingkat akurasi.....	51

DAFTAR TABEL

Tabel 3. 1 Data Latih dan Data Uji.....	23
Tabel 3. 2 Hasil perhitungan normalisasi terhadap nilai atribut.....	25
Tabel 3. 3 Hasil Perhitungan jarak.....	26
Tabel 3. 4 Hasil Perhitungan Jarak Setelah diurutkan.....	27
Tabel 3. 5 Uji Pengaruh nilai k dan data latih terhadap tingkat akurasi.....	34
Tabel 4. 1 Deskripsi metode/fungsi pada class OlahData.....	36
Tabel 4. 2 Deskripsi metode/fungsi pada class FuzzyKnn.....	38
Tabel 4. 3 Deskripsi metode/fungsi pada class Akurasi.....	41
Tabel 4. 4 Hasil pengujian dataset evaluasi kinerja karyawan.....	49

DAFTAR SOURCECODE

<i>Source Code</i> 4. 1 Fungsi untuk menghitung hasil normalisasi.....	43
<i>Source Code</i> 4. 2 Fungsi untuk menghitung distance.....	43
<i>Source Code</i> 4. 3 Fungsi untuk mengurutkan hasil distance.....	44
<i>Source Code</i> 4. 4 Fungsi untuk mengambil hasil distance berdasarkan nilai k.....	44
<i>Source Code</i> 4. 5 Fungsi untuk menghitung nilai keanggotaan tiap kelas.....	44
<i>Source Code</i> 4. 6 Fungsi untuk mencari nilai maksimum membership.....	44
<i>Source Code</i> 4. 7 Fungsi untuk menentukan kelas.....	45

DAFTAR LAMPIRAN

LAMPIRAN I Pengujian Data Latih dengan komposisi jumlah kelas yang sama.....	57
---	-----------



BAB I PENDAHULUAN

1.1 Latar Belakang

Sumber daya manusia yang mempunyai kompetensi tinggi sangat diperlukan karena keahlian atau kompetensi akan dapat mendukung peningkatan prestasi kerja karyawan. Dalam suatu perusahaan diperlukan sumber daya manusia yang berkualitas dimana menjadi salah satu faktor yang digunakan untuk meningkatkan kinerja [WIB-11]. Kinerja karyawan merupakan prestasi kerja atau hasil kerja yang dicapai sumber daya manusia persatuan periode waktu dalam melaksanakan tugas kerjanya sesuai dengan tanggung jawab yang diberikan kepadanya [MAN-00]. Suatu kinerja individu dapat ditingkatkan apabila ada kesesuaian antara pekerjaan dan kemampuan. Kinerja individu sendiri dipengaruhi oleh kepuasan kerja. Kepuasan kerja itu sendiri adalah perasaan individu terhadap pekerjaannya dimana perasaan itu berupa hasil penilaian mengenai seberapa jauh pekerjaannya secara keseluruhan. Dalam hal ini dibutuhkan suatu evaluasi, yang kemudian dikenal dengan penilaian kinerja.

Penilaian kinerja harus dilakukan untuk mengetahui prestasi yang dapat dicapai setiap karyawan. Proses evaluasi kinerja penting bagi setiap karyawan yang harus dilakukan secara adil, jujur, relevan dan valid yang sesuai dengan pekerjaan yang dilakukan serta bebas dari hal diskriminasi karena hal ini berguna bagi perusahaan untuk menetapkan tindakan kebijaksanaan selanjutnya. Evaluasi kinerja karyawan dilihat dari aspek-aspek yang telah ditentukan oleh perusahaan dimana karyawan yang dievaluasi akan mempunyai nilai evaluasi yang berbeda-beda untuk setiap aspeknya. Dengan nilai evaluasi yang berbeda-beda tersebut akan menjadi suatu masalah apabila dikaitkan dengan hasil keputusan yang akan berguna untuk mengembangkan sumber daya manusia dalam suatu perusahaan. Aspek-aspek yang dinilai pada evaluasi kinerja karyawan ialah kooperatif, kehadiran dan ketepatan waktu, keandalan, inisiatif, sikap, *judgement*, komunikasi, hubungan antar manusia dan keahlian profesional.

Salah satu proses yang dapat digunakan untuk mengevaluasi kinerja karyawan dengan melakukan proses *data mining* yaitu klasifikasi [MOR-09].

Menurut [KHU-07], klasifikasi adalah suatu fungsionalitas *data mining* yang akan menghasilkan model untuk memprediksi kelas atau kategori dari objek-objek di dalam baris data. Dalam pengevaluasian kinerja karyawan dilakukan pengklasifikasian berdasarkan *behavior* karyawan. Ada beberapa metode yang digunakan dalam pengklasifikasian, diantaranya ialah: *neural network*, *K-Nearest Neighbor* (K-NN), *naïve bayes classifier* (LAI-06]). Kelebihan algoritma KNN itu sendiri ialah sederhana dan efektif dalam penggunaannya [SAR-00]. Metode pengklasifikasian yang digunakan dalam penelitian ini ialah *K-Nearest Neighbor* (KNN). Menurut [RAK-07], Algoritma *K-Nearest Neighbor* (k-NN atau KNN) adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Beberapa permasalahan yang dihadapi dalam menggunakan *K-Nearest Neighbor Classifier* yaitu pertama ialah masing-masing vektor sampel dianggap sama pentingnya di dalam label kelas untuk vektor *input* sehingga sering menyebabkan kesulitan dalam pemrosesan vektor *input*. Permasalahan kedua yaitu setelah suatu vektor *input* dimasukkan ke dalam suatu kelas dan diproses ternyata vektor tersebut tidak memiliki indikasi keanggotaan yang kuat dalam kelas tersebut. Dari dua permasalahan tersebut diselesaikan dengan cara menggabungkan teori *fuzzy* pada aturan *KNN* [KEL-85].

Metode *Fuzzy K-Nearest Neighbor* adalah salah satu varian dari algoritma *K-Nearest Neighbor*. Metode *Fuzzy K-Nearest Neighbor* merupakan metode klasifikasi yang menggabungkan teknik *fuzzy* dan *K-Nearest Neighbor Classifier*. Menurut [BEZ-81], dalam penelitiannya menunjukkan bahwa algoritma *Fuzzy K-Nearest Neighbor* didapat dari hasil alokasi keanggotaan kelas *fuzzy* untuk vektor masukan, sehingga menghasilkan keputusan *fuzzy* berdasarkan label *fuzzy*. Penelitian ini berkaitan dengan menggabungkan metode himpunan *fuzzy* ke dalam aturan keputusan *KNN*. Khususnya, algoritma *Fuzzy K-Nearest Neighbor* dikembangkan memanfaatkan keanggotaan kelas *fuzzy* set sampel sehingga menghasilkan aturan klasifikasi *fuzzy*. Hasil dari kedua *crisp* dan aturan *Fuzzy KNN* dibandingkan pada dua *dataset*, dan algoritma *fuzzy* terbukti mendominasi *crisp* dengan memiliki tingkat kesalahan lebih rendah dan dengan memproduksi nilai-nilai keanggotaan yang berfungsi sebagai alat ukur dalam klasifikasi.

Berdasarkan latar belakang yang telah diuraikan, maka skripsi ini diberi judul **“Penerapan *Fuzzy K-Nearest Neighbor* (Fuzzy K-NN) Untuk Menentukan Status Evaluasi Kinerja Karyawan”**.

1.2 Rumusan Masalah

Berdasarkan uraian pada latar belakang, maka dalam skripsi ini dapat dirumuskan permasalahan sebagai berikut:

- a. Bagaimana menerapkan metode *Fuzzy K-Nearest Neighbor* (Fuzzy K-NN) pada kasus evaluasi kinerja karyawan.
- b. Seberapa besar akurasi klasifikasi yang dipengaruhi sejumlah nilai k pada *dataset* evaluasi kinerja karyawan menggunakan metode *Fuzzy K-Nearest Neighbor* (Fuzzy K-NN).

1.3 Batasan Masalah

Untuk membatasi masalah yang dikaji, maka penulis membatasi dan memfokuskan permasalahan yaitu data yang digunakan dalam penelitian ini adalah data penilaian kinerja karyawan berdasarkan *behavior* karyawan PT. Sierad Produce, Tbk.

1.4 Tujuan Penelitian

Tujuan penelitian dalam skripsi ini adalah:

- a. Menerapkan metode *Fuzzy K-Nearest Neighbor* (Fuzzy K-NN) pada kasus evaluasi kinerja karyawan.
- b. Menghitung akurasi klasifikasi yang dipengaruhi sejumlah nilai k pada *dataset* evaluasi kinerja karyawan menggunakan metode *Fuzzy K-Nearest Neighbor* (Fuzzy K-NN).

1.5 Manfaat Penelitian

Manfaat yang bisa diambil dari skripsi ini adalah untuk menentukan status evaluasi kinerja karyawan pada PT. Sierad Produce, Tbk. dan diharapkan dapat membantu penyelesaian masalah perusahaan tersebut dengan menggunakan waktu seefisien mungkin dalam pengambilan keputusan kinerja karyawan.

1.6 Sistematika Penulisan

Skripsi ini disusun berdasarkan sistematika penulisan sebagai berikut:

BAB I PENDAHULUAN

Bab ini menjelaskan pentingnya algoritma *Fuzzy K-Nearest Neighbor* yang termuat dalam latar belakang, perumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, metodologi penelitian, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab ini menguraikan teori-teori yang berhubungan dengan penilaian kinerja karyawan, *data mining*, metode *Fuzzy K-Nearest Neighbor*, dan akurasi.

BAB III METODOLOGI DAN PERANCANGAN SISTEM

Bab ini menjelaskan tentang metode yang digunakan dan tahapan-tahapan perhitungan algoritma *Fuzzy K-Nearest Neighbor*.

BAB IV IMPLEMENTASI DAN PEMBAHASAN

Bab ini menjelaskan mengenai implementasi program, uji coba dan analisisnya.

BAB V PENUTUP

Bab ini berisi kesimpulan dari seluruh rangkaian penelitian serta saran untuk kemungkinan pengembangan lebih lanjut.

BAB II

TINJAUAN PUSTAKA

2.1 Sumber Daya Manusia

Sumber daya manusia dalam sebuah organisasi perlu dikelola secara profesional agar terwujud keseimbangan antara kebutuhan pegawai dengan tuntutan dan kemampuan organisasi perusahaan. Keseimbangan tersebut merupakan kunci utama perusahaan agar dapat berkembang secara produktif dan wajar [MAN-04].

Meningkatnya peranan manajemen dalam suatu perusahaan mengakibatkan bertambahnya perhatian terhadap pentingnya faktor sumber daya manusia dalam perusahaan. Perhatian perusahaan yang pada mulanya lebih besar ditekankan pada bidang mekanis dan modal, kini telah mengalami perubahan. Perusahaan kini memberikan perhatian yang lebih besar terhadap masalah yang berhubungan dengan faktor sumber daya manusia. Sumber daya manusia selalu dibutuhkan oleh setiap perusahaan, baik itu perusahaan kecil maupun besar. Meskipun sumber daya manusia merupakan salah satu faktor produksi yang ada pada perusahaan akan tetapi peranannya dalam pencapaian tujuan sangat besar.

Sumber daya manusia yang mempunyai kompetensi tinggi sangat diperlukan karena keahlian atau kompetensi akan dapat mendukung peningkatan prestasi kerja karyawan. Dalam suatu perusahaan diperlukan sumber daya manusia yang berkualitas dimana menjadi salah satu faktor yang digunakan untuk meningkatkan kinerja [WIB-11].

2.2 Kinerja Karyawan

Secara etimologi, kinerja berasal dari kata prestasi kerja (*performance*). Istilah kinerja berasal dari kata *job performance* atau *actual performance* (prestasi kerja atau prestasi sesungguhnya yang dicapai seseorang) yaitu hasil kerja secara kualitas dan kuantitas yang dicapai oleh seorang pegawai dalam melaksanakan tugasnya sesuai dengan tanggung jawab yang diberikan kepadanya [MAN-05].

Kinerja atau prestasi kerja adalah hasil atau prestasi yang telah dicapai oleh seseorang dalam melaksanakan pekerjaannya sesuai dengan *standard* dan *criteria* yang ditetapkan untuk pekerjaan itu. Atau dari sisi lain prestasi kerja atau kinerja didefinisikan sebagai tindakan-tindakan atau pelaksanaan tugas yang telah diselesaikan oleh seseorang atau kelompok orang dalam kurun waktu tertentu dan dapat diukur.

2.3 Evaluasi Kinerja

Istilah Evaluasi dapat disamakan dengan penaksiran (*appraisal*), pemberian angka (*rating*) dan penilaian (*assesment*). Evaluasi kinerja sangat penting untuk menilai akuntabilitas organisasi dalam menghasilkan pelayanan publik. Akuntabilitas bukan sekedar kemampuan menunjukkan bagaimana uang publik dibelanjakan, akan tetapi meliputi apakah uang tersebut dibelanjakan secara ekonomis, efektif, dan efisien.

Evaluasi kinerja membantu pimpinan untuk mengambil keputusan dalam suatu kebijakan, nilai yang dihasilkan dari evaluasi membuat suatu kebijakan bermanfaat bagi pelayanan publik. Evaluasi kinerja merupakan suatu proses penilaian kinerja aparatur yang dilakukan untuk melihat tanggung jawab pekerjaannya setiap hari apakah terjadi peningkatan atau penurunan sehingga pemimpin bisa memberikan suatu motivasi penunjang untuk melihat kinerja aparatur kedepannya. Evaluasi harus sering dilakukan agar masalah yang di hadapi dapat diketahui dan dicari jalan keluar yang baik [MAN-05].

Pada skripsi ini atribut yang digunakan dalam data evaluasi kinerja karyawan ialah kooperatif, kehadiran dan ketepatan waktu, keandalan, inisiatif, sikap, judgement, komunikasi, hubungan antar manusia, dan keahlian profesional.

2.4 Logika Fuzzy

Logika *fuzzy* adalah suatu cara yang tepat untuk memetakan ruang *input* kedalam suatu ruang *output* [KUS-03]. Konsep ini diperkenalkan dan dipublikasikan pertama kali oleh Lotfi A. Zadeh, seorang profesor dari University of California di Berkeley pada tahun 1965. Logika *fuzzy* menggunakan ungkapan bahasa untuk menggambarkan nilai variabel. Logika *fuzzy* bekerja dengan

menggunakan derajat keanggotaan dari sebuah nilai yang kemudian digunakan untuk menentukan hasil yang ingin dihasilkan berdasarkan atas spesifikasi yang telah ditentukan. Telah disebutkan sebelumnya bahwa logika *fuzzy* memetakan ruang *input* ke ruang *output*. Antara *input* dan *output* ada suatu kotak hitam yang harus memetakan *input* ke *output* yang sesuai. Alasan mengapa orang menggunakan logika *fuzzy*, yaitu [KUS-10]:

1. Konsep logika *fuzzy* mudah dimengerti karena menggunakan konsep matematis yang sederhana dan mudah dimengerti.
2. Logika *fuzzy* sangat fleksibel.
3. Logika *fuzzy* memiliki toleransi terhadap data yang tidak tepat.
4. Logika *fuzzy* mampu memodelkan fungsi-fungsi nonlinier yang sangat kompleks.
5. Logika *fuzzy* dapat membangun dan mengaplikasikan pengalaman-pengalaman para pakar secara langsung tanpa harus melalui proses pelatihan.
6. Logika *fuzzy* dapat bekerjasama dengan teknik-teknik kendali secara konvensional.
7. Logika *fuzzy* didasarkan pada bahasa alami.

2.5 Himpunan *Fuzzy*

Himpunan tegas (*crisp*) A didefinisikan oleh item-item yang ada pada himpunan itu. Pada himpunan tegas (*crisp*), nilai keanggotaan suatu item x dalam suatu himpunan A ($\mu_A(x)$) memiliki dua kemungkinan [KUS-10], yaitu:

1. Satu (1), yang berarti bahwa suatu item menjadi anggota dalam suatu himpunan.
2. Nol (0), yang berarti bahwa suatu item tidak menjadi anggota dalam suatu himpunan.

Himpunan *Fuzzy* didasarkan pada gagasan untuk memperluas jangkauan fungsi karakteristik sedemikian hingga fungsi tersebut akan mencakup bilangan real pada interval $[0,1]$. Nilai keanggotaannya menunjukkan bahwa suatu item dalam semesta pembicaraan tidak hanya berada pada 0 atau 1, namun juga nilai yang terletak diantaranya. Dengan kata lain, nilai kebenaran suatu item tidak hanya benar atau salah. Nilai 0 menunjukkan salah, nilai 1 menunjukkan benar,

dan masih ada nilai-nilai yang terletak antara benar dan salah. Himpunan *fuzzy* memiliki 2 atribut, yaitu [KUS-03]:

- a. Linguistik, yaitu penamaan suatu grup yang mewakili suatu keadaan atau kondisi tertentu dengan menggunakan bahasa alami.
- b. Numeris, yaitu suatu nilai (angka) yang menunjukkan ukuran dari suatu variabel.

Ada beberapa hal yang perlu diketahui dalam memahami sistem *fuzzy*, yaitu:

- a. Variabel *Fuzzy*

Variabel *fuzzy* merupakan variabel yang hendak dibahas dalam suatu sistem *fuzzy*.

- b. Himpunan *Fuzzy*

Himpunan *fuzzy* merupakan suatu grup yang mewakili suatu kondisi atau keadaan tertentu dalam suatu variabel.

- c. Semesta Pembicaraan

Semesta pembicaraan adalah keseluruhan nilai yang diperbolehkan untuk dioperasikan dalam suatu variabel *fuzzy*. Semesta pembicaraan merupakan himpunan bilangan real yang senantiasa naik (bertambah) secara monoton dari kiri ke kanan. Nilai semesta pembicaraan dapat berupa bilangan positif maupun negatif. Adakalanya nilai semesta pembicaraan ini tidak dibatasi batas atasnya.

- d. Domain

Domain himpunan *fuzzy* adalah keseluruhan nilai yang diijinkan dalam semesta pembicaraan dan boleh dioperasikan dalam suatu himpunan *fuzzy*. Seperti halnya semesta pembicaraan, domain merupakan himpunan bilangan real yang senantiasa naik (bertambah) secara monoton dari kiri ke kanan. Nilai domain dapat berupa bilangan positif maupun negatif.

2.6 Data Mining

2.6.1 Definisi Data Mining

Pengertian *data mining* dari beberapa referensi adalah sebagai berikut:

1. Secara sederhana dapat didefinisikan bahwa *data mining* adalah ekstraksi informasi atau pola yang penting atau menarik dari data yang ada di

database yang besar sehingga menjadi informasi yang sangat berharga [GIR-03].

2. *Data Mining* merupakan proses penemuan yang efisien sebuah pola terbaik yang dapat menghasilkan sesuatu yang bernilai dari suatu koleksi data yang sangat besar [THE-00].
3. *Data Mining* adalah satu pola yang menguntungkan dalam melakukan *search* pada sebuah *database* yang terdapat pada sebuah model. Proses ini dilakukan berulang-ulang (*iterasi*) hingga didapat satu set pola yang memuaskan yang dapat berfungsi sesuai yang diharapkan [CHU-98].
4. *Data Mining* adalah sebuah *class* dari suatu aplikasi *database* yang mencari pola-pola yang tersembunyi di dalam sebuah kumpulan data yang dapat digunakan untuk memprediksi perilaku yang akan datang [KHU-07].

Berdasarkan beberapa pengertian tersebut, dapat ditarik kesimpulan bahwa *data mining* adalah suatu algoritma di dalam menggali informasi berharga yang terpendam atau tersembunyi pada suatu koleksi data (*database*) yang sangat besar sehingga ditemukan satu pola yang menarik yang sebelumnya tidak diketahui.

Menurut [ZAI-99], metode *data mining* dibagi ke dalam dua macam, yaitu [ZAI-99]:

1. *Supervised Learning* (proses belajar terawasi)

Pada pembelajaran ini target tersedia. Dengan kata lain, label atau nama dan nomor kelas pada *data training* (supervisi) diketahui, kemudian data baru diklasifikasikan berdasarkan *data mining* tersebut. Contoh: Klasifikasi.

2. *Unsupervised Learning* (proses belajar tak terawasi)

Pada proses pembelajaran ini, tidak membutuhkan target untuk keluarannya. Dengan kata lain, tidak terdapat label atau nama kelas pada *data training*, atau bahkan tidak diketahui nomor kelasnya. *Data training* dikelompokkan berdasarkan ukuran kesamaan (*similarity*) oleh karena itu tidak ada perbandingan yang dilakukan dengan respon ideal yang ditetapkan sebelumnya. Rangkaian pelatihan hanya berisi vektor masukan saja. Contoh: *Clustering*.

Kebanyakan tugas *data mining* merupakan *supervised data mining*.

2.6.2 Proses Data Mining

Prosedur yang umum digunakan untuk permasalahan data mining meliputi tahap-tahap sebagai berikut: [KAN-03]

1. Menentukan permasalahan dan merumuskan hipotesis

Pada tahap ini, ditentukan variabel-variabel dan hipotesis awal.

2. Mengumpulkan data

Tahap ini berkaitan dengan bagaimana data dihasilkan dan dikumpulkan.

3. *Preprocessing* data

Dilakukan pembersihan terhadap *outlier*, penanganan *missing value* maupun transformasi data.

4. Memperkirakan model

Pemilihan dan penerapan teknik data mining yang sesuai adalah tugas utama dalam tahap ini.

5. Menafsirkan model dan menarik kesimpulan

Pada tahap ini, dilakukan penafsiran model untuk membantu dalam 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. Oleh karena itu, perlu dilakukan transformasi data berupa normalisasi terhadap nilai untuk membakukan skala pengaruh yang ada pada atribut, terhadap hasil. Ada beberapa teknik normalisasi data seperti normalisasi *min-max* dan *Z-score* [MOR-09].

Dalam skripsi digunakan normalisasi *min-max*. Normalisasi *min-max* dihitung dengan persamaan 2.1 [MOR-09]. Menurut [JAY-11] normalisasi *min-max* memiliki keunggulan yaitu menjaga relasi pada data. Serta mempunyai fungsi yaitu menyatukan satuan dari berbagai atribut.

$$V' = \frac{V - \min_A}{\max_A - \min_A} \quad (2.1)$$

Dimana,

V' : hasil normalisasi yang nilainya berkisar antara 0 dan 1

V : nilai atribut A yang akan dinormalisasi

\min_A : nilai minimum dari suatu atribut, A

\max_A : nilai maksimum dari suatu atribut, A.

2.7 Klasifikasi

Menurut [KHU-07], 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.

2.8 Algoritma *K-Nearest Neighbor*

2.8.1 Definisi *K-Nearest Neighbor*

Algoritma *K-Nearest Neighbor* (KNN) adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Data pembelajaran diproyeksikan ke ruang berdimensi banyak, dimana masing-masing dimensi merepresentasikan

fitur dari data. Ruang ini dibagi menjadi bagian-bagian berdasarkan klasifikasi data pembelajaran. Sebuah titik pada ruang ini ditandai kelas c jika kelas c merupakan klasifikasi yang paling banyak ditemui pada k buah tetangga terdekat titik tersebut. Dekat atau jauhnya tetangga biasanya dihitung berdasarkan jarak *euclidean*.

Beberapa keuntungan dari metode *K-Nearest Neighbor* (KNN) adalah sebagai berikut:

- a. Sederhana dalam penggunaannya.
- b. Dapat menangani data training yang mengandung noise.
- c. Efektif jika data training besar.

Meskipun memiliki beberapa keuntungan metode *K-Nearest Neighbor* (KNN) juga memiliki beberapa kelemahan seperti berikut ini [HAM-08]:

- a. *Computation cost* cukup tinggi karena perlu untuk menghitung jarak setiap data training.
- b. Membutuhkan memori yang besar.
- c. Rendahnya tingkat akurasi pada dataset multidimensi.
- d. Perlu untuk menentukan nilai K parameter, jumlah tetangga terdekat.
- e. Menggunakan perhitungan jarak, yang belum diketahui pasti jenis jarak yang digunakan.
- f. Belum diketahui atribut yang lebih baik untuk menghasilkan hasil terbaik.

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 [SAR-00]. Menurut [MOR-09], *K-Nearest Neighbor* juga merupakan algoritma yang sering digunakan untuk klasifikasi pada teknik *data mining*.

2.8.2 Proses *K-Nearest Neighbor* (KNN)

Menurut Agusta (2007), bahwa prinsip kerja *K-Nearest Neighbor* (KNN) adalah mencari jarak terdekat antara data yang dievaluasi dengan k tetangga terdekatnya dalam data pelatihan. Persamaan perhitungan untuk mencari *euclidean* ditunjukkan oleh persamaan 2.2.

$$d_i = \sqrt{\sum_{i=1}^p (x_{2i} - x_{1i})^2} \quad (2.2)$$

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.

2.9 Fuzzy K-Nearest Neighbor

2.9.1 Definisi Fuzzy K-Nearest Neighbor

Fuzzy K-NN merupakan metode klasifikasi yang menggabungkan teknik *fuzzy* dengan *k-Nearest Neighbor classifier*. Algoritma *fuzzy K-nearest neighbor* memberikan nilai keanggotaan kelas pada vektor sampel dan bukan menempatkan vektor pada kelas tertentu. FK-NN merupakan 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. Sebagai contoh, jika vektor diberikan nilai keanggotaan 0.9 pada kelas pertama dan 0.05 pada dua kelas lainnya peneliti dapat cukup yakin bahwa kelas dengan nilai keanggotaan 0.9 adalah kelas milik vektor tersebut. Di sisi lain, jika sebuah vektor diberi nilai keanggotaan 0.55 pada kelas pertama, 0.44 pada kelas kedua, dan 0.01 pada kelas ketiga, maka peneliti harus ragu-ragu untuk menetapkan vektor berdasarkan hasil ini. Namun, dapat diyakinkan bahwa vektor tersebut bukan milik kelas ketiga. Dalam kasus seperti ini, perlu diperiksa lebih lanjut untuk menentukan klasifikasinya, karena memiliki derajat keanggotaan

yang tinggi pada dua kelas yaitu satu dan dua. Pemberian nilai keanggotaan oleh algoritma ini jelas berguna dalam proses klasifikasi.

Sebuah data memiliki nilai keanggotaan pada setiap kelas yang berbeda dengan nilai derajat keanggotaan dalam interval $[0,1]$. Teori himpunan fuzzy menggeneralisasi teori K-NN klasik dengan mendefinisikan nilai keanggotaan sebuah data pada masing-masing kelas. Formula yang digunakan ditunjukkan oleh persamaan 2.3.

$$u_i(x) = \frac{\sum_{j=1}^k u_{ij} (\|x-x_j\|^{-2/(m-1)})}{\sum_{j=1}^k (\|x-x_j\|^{-2/(m-1)})} \quad (2.3)$$

Keterangan:

$u_i(x)$ = nilai keanggotaan data x ke kelas c_i

k = jumlah tetangga terdekat yang digunakan

u_{ij} = nilai keanggotaan kelas i pada vector j

$x-x_j$ = selisih jarak dari data x ke data x_j dalam k tetangga terdekat

m = bobot pangkat (*weight exponent*) yang besarnya $m > 1$.

Nilai u_{ij} pada $u_i(x)$ terlebih dahulu diproses dengan menggunakan persamaan 2.4.

$$u_{ij} = \begin{cases} 0.51 + (n_j/K) * 0.49, & \text{jika } j=i \\ (n_j/K) * 0.49, & \text{jika } j \neq i \end{cases} \quad (2.4)$$

Keterangan:

n_j = jumlah anggota kelas j pada suatu dataset K

K = total data latih yang digunakan

j = kelas target (training/tidak training)

2.9.2 Proses Fuzzy K-Nearest Neighbor

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

1. Melakukan normalisasi terhadap nilai-nilai atribut menggunakan normalisasi *min-max* yang ditunjukkan oleh persamaan 2.1.
2. Menghitung jarak antara dua *record* menggunakan *Euclidean distance* yang ditunjukkan oleh persamaan 2.2.
3. Menghitung nilai keanggotaan $u_i(x)$ menggunakan persamaan 2.3 untuk setiap i , dimana $1 \leq i \leq C$.
4. Mengambil nilai terbesar dari proses nomer 3 untuk semua $1 \leq i \leq C$, C adalah jumlah kelas.
5. Memberikan label kelas baru pada proses nomer 4.

2.10 Data Evaluasi Kinerja Karyawan PT Sierad Produce, Tbk

Data Evaluasi Kinerja Karyawan PT Sierad Produce, Tbk merupakan data hasil penilaian kinerja karyawan yang dilakukan oleh departemen *Human Resources Development* (HRD) PT Sierad Produce, Tbk. Data ini terdiri dari 9 atribut, yaitu kooperatif, kehadiran dan ketepatan waktu, keandalan, inisiatif, sikap, judgement, komunikasi, hubungan antar manusia, dan keahlian profesional.

2.11 Perhitungan Akurasi

Akurasi merupakan seberapa dekat suatu angka hasil pengukuran terhadap angka sebenarnya (*true value* atau *reference value*). Dalam penelitian ini akurasi diagnosis dihitung dari jumlah diagnosis yang tepat dibagi dengan jumlah data. Tingkat akurasi diperoleh dengan perhitungan sesuai dengan persamaan 2.5 [NUG-06].

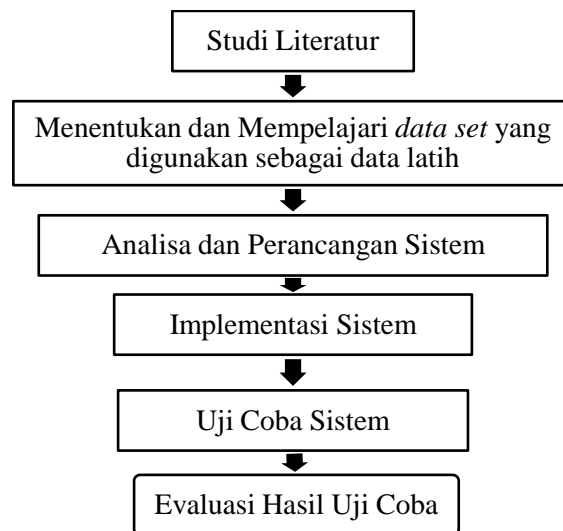
$$\text{Akurasi} = \frac{\sum \text{data uji benar}}{\sum \text{jumlah total data uji}} \times 100\% \quad (2.5)$$

BAB III METODOLOGI DAN PERANCANGAN

Pada bab metodologi dan perancangan ini akan dibahas metode, perancangan yang digunakan, dan langkah-langkah yang dilakukan dalam penelitian tentang *fuzzy k-nearest neighbor* untuk menentukan status evaluasi kinerja karyawan. Penelitian dilakukan dengan tahapan-tahapan sebagai berikut.

1. Mempelajari literatur yang berhubungan dengan evaluasi kinerja karyawan dan algoritma *fuzzy k-nearest neighbor*.
2. Menentukan dan mempelajari *data set* yang digunakan sebagai data latih.
3. Menganalisa dan melakukan perancangan sistem dengan algoritma *fuzzy k-nearest neighbor*.
4. Membangun perangkat lunak berdasarkan analisis dan perancangan yang telah dilakukan (implementasi).
5. Melakukan uji coba terhadap perangkat lunak.
6. Mengevaluasi hasil yang diperoleh dari uji coba perangkat lunak.

Adapun tahapan-tahapan penelitian dapat digambarkan dalam bentuk diagram alir yang ditunjukkan pada gambar 3.1.



Gambar 3. 1 Diagram Alir Penelitian

3.1 Studi Literatur

Dalam penelitian ini dibutuhkan studi literatur untuk merealisasikan tujuan dan penyelesaian masalah. Teori-teori mengenai evaluasi kinerja karyawan,

algoritma *fuzzy k-nearest neighbor* digunakan sebagai dasar penelitian yang diperoleh dari buku, jurnal dan browsing dari internet. Kemudian data yang diperoleh diubah sehingga dapat digunakan untuk analisis. Setelah dianalisis maka dapat diimplementasikan ke dalam program.

3.2 Data Penelitian

Dataset yang digunakan dalam skripsi ini adalah kumpulan data tentang evaluasi kinerja karyawan pada PT. Sierad Produce, Tbk yang merupakan data hasil penilaian kinerja karyawan yang dilakukan oleh departemen *Human Resources Development (HRD)*. Pada *dataset* evaluasi kinerja karyawan, atribut yang digunakan ialah kooperatif, kehadiran dan ketepatan waktu, keandalan, inisiatif, sikap, *judgement*, komunikasi, hubungan antara manusia, dan keahlian profesional. Sedangkan kelas output yaitu:

- 1 = Mengikuti training/pelatihan.
- 2 = Tidak mengikuti training/pelatihan.

3.3 Deskripsi Umum Sistem

Secara umum sistem yang dibangun adalah suatu perangkat lunak untuk mengklasifikasi status evaluasi kinerja karyawan yang mengimplementasikan metode *Fuzzy K-Nearest Neighbor (Fuzzy K-NN)*. Sistem ini bertujuan untuk menentukan apakah seorang karyawan dapat mengikuti *training* (pelatihan) atau tidak berdasarkan beberapa atribut yang digunakan.

Perangkat lunak ini akan menguji keakuratan hasil klasifikasi dataset evaluasi kinerja karyawan terhadap data sebenarnya. Parameter uji yang berkaitan dengan nilai k (tetangga) dan data latih yang berpengaruh terhadap tingkat akurasi.

3.4 Perancangan Sistem

Pada perancangan sistem ini akan dijelaskan mengenai proses-proses dalam membangun sebuah sistem. Sistem ini bertujuan untuk memberikan informasi pada karyawan tentang status evaluasi kinerja karyawan, sistem akan melakukan klasifikasi sesuai dengan kriteria yang dimasukkan, klasifikasi

dilakukan dengan menggunakan metode *Fuzzy K-Nearest Neighbor (Fuzzy K-NN)*.

3.4.1 Proses Klasifikasi Algoritma *Fuzzy K-Nearest Neighbor*

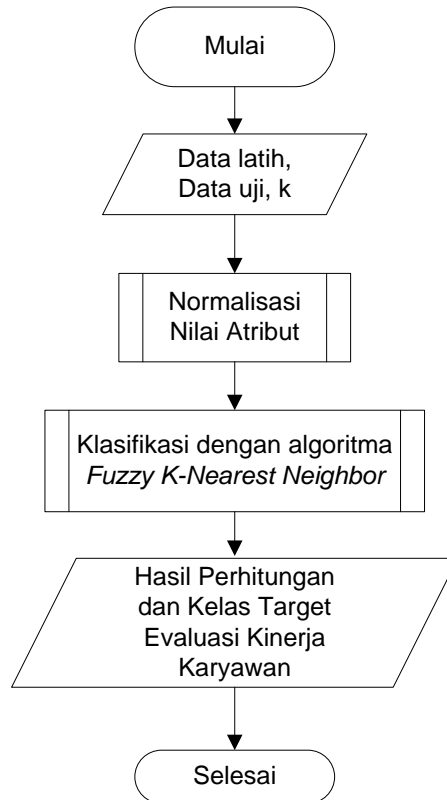
Pada proses ini sistem akan melakukan klasifikasi pada data latih dengan menggunakan algoritma *Fuzzy K-Nearest Neighbor (Fuzzy K-NN)*. Tahapan proses adalah sebagai berikut:

1. Proses global adalah sebuah proses yang mencakup alur dari proses secara garis besar.
2. Proses *request* data adalah sebuah proses yang melakukan pemanggilan data yang disimpan pada sebuah *database*.
3. Proses klasifikasi adalah sebuah proses yang menunjukkan alur perhitungan dari perhitungan awal sampai ditemukan kelas evaluasi kinerja karyawan.
4. Menghitung normalisasi *min-max*.
5. Menghitung *euclidean*.
6. Menghitung nilai derajat keanggotaan dan mengambil nilai terbesar dari proses tersebut dan ditentukan kelas targetnya.

3.4.1.1 Flowchart Sistem

Langkah-langkah dalam proses ini antara lain:

1. Melakukan input data latih dan data uji evaluasi kinerja karyawan dan k .
 2. Melakukan proses normalisasi atribut dengan menggunakan normalisasi *min-max*.
 3. Melakukan klasifikasi data evaluasi kinerja karyawan dengan algoritma *Fuzzy K-Nearest Neighbor (FK-NN)*.
 4. Output berupa hasil perhitungan dan kelas target evaluasi kinerja karyawan.
- Untuk lebih jelasnya dapat dilihat pada gambar 3.2 alur proses sistem.



Gambar 3. 2 Alur Proses Sistem

Alur proses sistem terdiri atas 2 proses, yaitu proses normalisasi setiap nilai atribut, proses klasifikasi dengan algoritma *fuzzy k-nearest neighbor*. Input proses sistem berupa data latih, data uji evaluasi kinerja karyawan dan k , sedangkan output proses berupa hasil perhitungan dan kelas target evaluasi kinerja karyawan.

Proses normalisasi nilai atribut merupakan proses transformasi berupa normalisasi terhadap nilai-nilai atribut yang berguna untuk membakukan skala pengaruh yang ada pada atribut terhadap hasil. Proses klasifikasi dengan algoritma *Fuzzy K-Nearest Neighbor* yang berguna untuk mengklasifikasikan hasil perhitungan jarak antara dua record menggunakan *Euclidean distance* dan menghitung nilai derajat keanggotaan serta memilih nilai derajat keanggotaan yang terbesar untuk ditentukannya kelas target.

3.4.1.2 Proses Normalisasi Nilai Atribut

Langkah-langkah dalam proses normalisasi nilai atribut yaitu:

1. Masukan (input) berupa data latih dan data uji evaluasi kinerja karyawan.

2. Menghitung iterasi untuk mencari nilai maksimum (max) dan nilai minimum (min) untuk tiap-tiap atribut.
3. Melakukan perhitungan normalisasi min-max dari hasil pada langkah ke-2 dengan menggunakan persamaan 2.4 dan didapatkan hasil nilai atribut pada data latih dan data uji yang ternormalisasi.

Flowchart untuk proses normalisasi nilai atribut ditampilkan pada gambar 3.3.

3.4.1.3 Proses Klasifikasi dengan Algoritma *Fuzzy K-Nearest Neighbor*

Langkah-langkah dalam proses klasifikasi dengan algoritma fuzzy k-nearest neighbor ialah:

1. Masukan (input) berupa data latih dan data uji evaluasi kinerja karyawan yang telah dinormalisasi serta nilai k yang dimasukkan oleh pengguna.
2. Dilakukan iterasi untuk menghitung distance (jarak) antara record baru dengan tiap record data latih menggunakan persamaan 2.2.
3. Menentukan maximum membership dan kelas target evaluasi kinerja karyawan.
4. Hasil keluaran berupa kelas target evaluasi kinerja karyawan dari data uji.

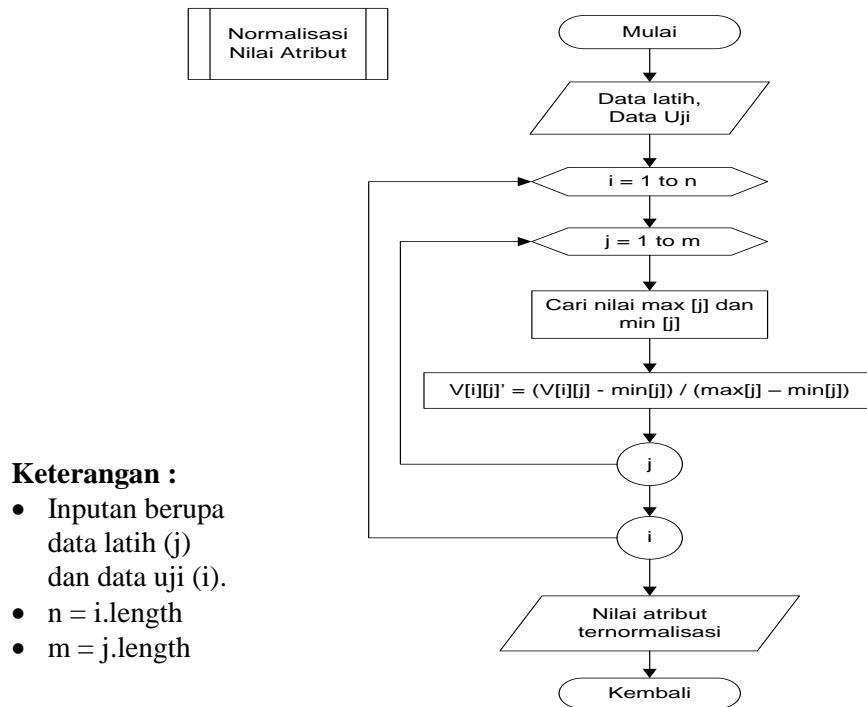
Flowchart untuk proses klasifikasi dengan algoritma fuzzy k-nearest neighbor ditunjukkan pada gambar 3.4.

3.4.1.4 Proses Menentukan Maximum Membership dan Kelas Target

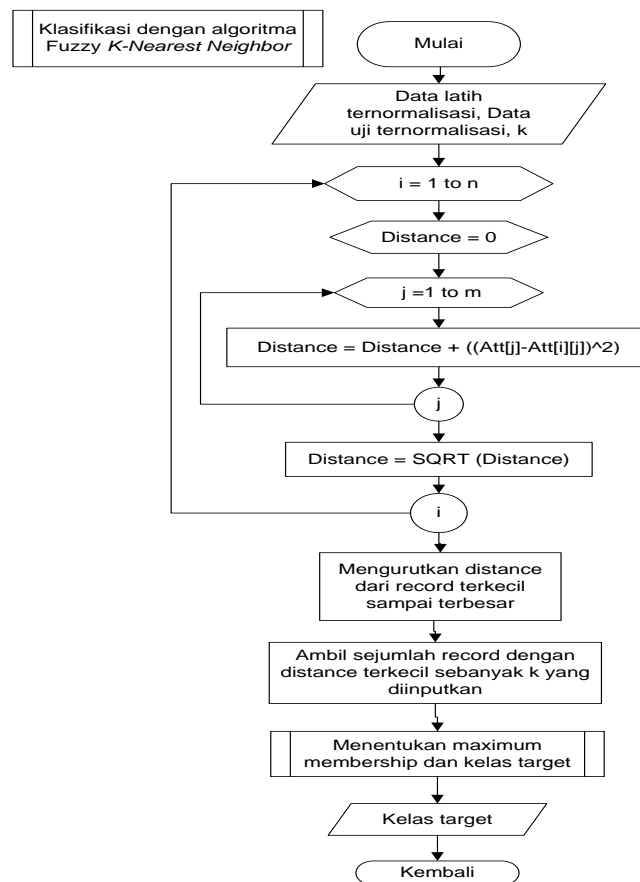
Langkah-langkah dalam proses menentukan maximum membership dan kelas target evaluasi kinerja karyawan ialah:

1. Masukan (input) berupa record dengan distance terkecil sejumlah k ($Distance_k[]$), m (bobot pangkat yang besarnya $m > 1$), dan nilai k.
2. Dilakukan proses untuk mencari nilai membership pada masing-masing kelas dan diambil nilai membership yang terbesar digunakan untuk menentukan kelas target yang baru.
3. Hasil keluaran berupa maksimum membership dan kelas target dari data uji.

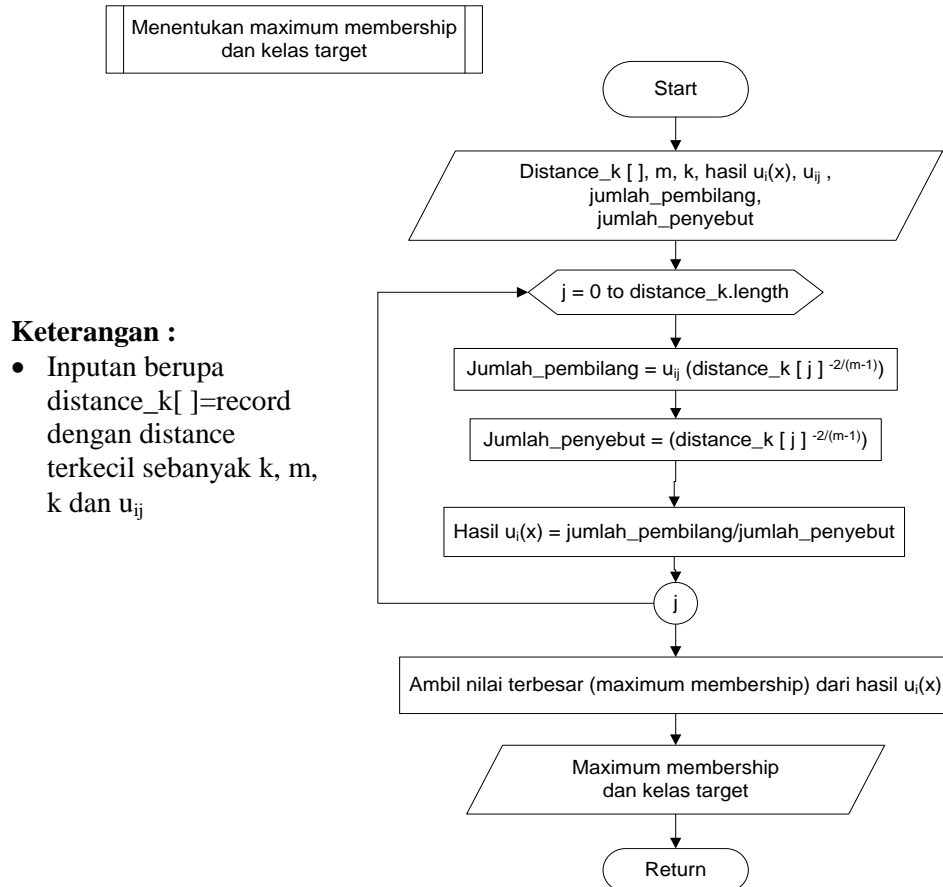
Flowchart untuk proses penentuan kelas target ditunjukkan pada gambar 3.5.



Gambar 3. 3 Flowchart Normalisasi Nilai Atribut



Gambar 3. 4 Flowchart Klasifikasi dengan Algoritma Fuzzy K-Nearest Neighbor



Gambar 3. 5 Flowchart Proses Menentukan Maximum Membership dan Kelas Target

3.5 Contoh Perhitungan Manual

Perhitungan manual ini menggunakan sampel data yang diambil dari *dataset* evaluasi kinerja karyawan dan digunakan seluruh atribut *dataset* tersebut yang berjumlah 10 atribut. Diambil sebanyak 16 record dengan rincian record ke-1 sampai 15 merupakan data latih dan record ke-16 merupakan data uji.

3.5.1 Data Latih dan Data Uji pada data evaluasi kinerja karyawan

Pada contoh perhitungan kali ini digunakan 15 data latih dan menggunakan satu data uji. Data latih dan data uji yang dipakai seperti yang ditunjukkan pada tabel 3.1.

Tabel 3. 1 Data Latih dan Data Uji

No Rec .	Koop eratif (1)	Kehadiran dan Ketepatan Waktu (2)	Keandalan (3)	Inisiatif (4)	Sikap (5)	Judgement (6)	Komunikasi (7)	Hubungan antar manusia (8)	Keahlian Profesional (9)	Target (10)
1	81.4	80.4	74.6	75.8	80.3	76.8	81.9	74.7	53.3	2
2	97.3	91.6	74.5	86.6	77.1	82.5	93.1	79.7	76.3	2
3	88.2	90	85.2	78.4	84.2	81.8	72.8	82	83.6	2
4	85.6	76	83.7	83.9	76.1	74.8	76.2	75.1	75.7	2
5	80.9	90	76	83.7	83.7	76	87.9	80.9	73.6	2
6	81.6	86.5	76	77.4	76	87.2	83.7	85.8	76	2
7	70.3	81.2	83.3	75.5	72.8	83.3	71.9	75.6	80.5	2
8	75.1	90	84.7	83.1	79	79.1	70.7	86.3	73	2
9	77.5	75.9	79	74.2	79.9	83.1	89.5	77.2	85.9	2
10	79.3	78.5	81	86.1	72.8	83.9	90.4	75.4	89.4	2
11	83.4	98.1	83.9	92.7	81	75.9	84.8	82.9	89.6	2
12	73.6	66.8	91.5	89.3	73.1	76.3	93.4	84.6	91.4	2
13	71	95.5	85.5	73.7	83.5	70.6	58.3	83.1	72.5	1
14	64	80.2	62.5	73.8	86.5	69.3	71.6	69.2	73.2	1
15	69.8	89.6	65.1	80.2	75.2	66.8	76.2	87.5	80.9	1
16	61.9	71.7	65.9	52.8	82.9	57.5	54.6	75	64.5	?

Keterangan:



= data latih



= data uji

3.5.2 Melakukan proses normalisasi terhadap nilai atribut

Proses perhitungan normalisasi terhadap nilai atribut ialah dengan menggunakan normalisasi min-max. Sebelum melakukan normalisasi terhadap nilai-nilai atribut, terlebih dahulu dicari nilai minimum dan nilai maksimum dari setiap atribut. Berikut ini adalah contoh perhitungan normalisasi nilai record pertama untuk atribut ke-1 (kooperatif).

$$V = 81.4 \text{ (nilai record pertama atribut ke-1)}$$

$$\min_A = 61.9$$

$$\max_A = 97.3$$

Setelah diketahui nilai minimum dan maksimum, kemudian dilakukan perhitungan menggunakan persamaan 2.1.


$$\begin{aligned} V' &= \frac{V - \min_A}{\max_A - \min_A} \\ &= \frac{81.4 - 61.9}{97.3 - 61.9} \\ &= 0.550847458 \end{aligned}$$


Dari perhitungan ini diperoleh nilai normalisasi untuk record pertama pada atribut ke-1, yaitu 0.550847458. Proses yang sama juga dilakukan untuk nilai record pada seluruh atribut. Hasil perhitungan normalisasi ditampilkan pada tabel 3.2.

Tabel 3. 2 Hasil perhitungan normalisasi terhadap nilai atribut

No Rec	Kooperatif (1)	Kehadiran dan Ketepatan Waktu (2)	Keandalan (3)	Inisiatif (4)	Sikap (5)	Judgement (6)	Komunikasi (7)	Hubungan antar manusia (8)	Keahlian Profesional (9)	Target (10)
1	0.550847458	0.434504792	0.417241	0.5764	0.55	0.6498316	0.703608247	0.300546448	0	1
2	1	0.792332268	0.413793	0.8471	0.31	0.8417508	0.992268041	0.573770492	0.603674541	1
3	0.742937853	0.741214058	0.782759	0.6416	0.83	0.8181818	0.469072165	0.699453552	0.795275591	1
4	0.669491525	0.293929712	0.731034	0.7794	0.24	0.5824916	0.556701031	0.322404372	0.587926509	2
5	0.536723164	0.741214058	0.465517	0.7744	0.8	0.6228956	0.858247423	0.639344262	0.532808399	1
6	0.556497175	0.629392971	0.465517	0.6165	0.23	1	0.75	0.907103825	0.595800525	1
7	0.237288136	0.460063898	0.717241	0.5689	0	0.8686869	0.445876289	0.349726776	0.713910761	2
8	0.372881356	0.741214058	0.765517	0.7594	0.45	0.7272727	0.414948454	0.93442623	0.517060367	1
9	0.440677966	0.290734824	0.568966	0.5363	0.52	0.8619529	0.899484536	0.43715847	0.855643045	1
10	0.491525424	0.373801917	0.637931	0.8346	0	0.8888889	0.922680412	0.338797814	0.947506562	2
11	0.607344633	1	0.737931	1	0.6	0.6195286	0.778350515	0.74863388	0.952755906	1
12	0.330508475	0	1	0.9148	0.02	0.6329966	1	0.841530055	1	1
13	0.257062147	0.916932907	0.793103	0.5238	0.78	0.4410774	0.095360825	0.759562842	0.503937008	2
14	0.059322034	0.428115016	0	0.5263	1	0.3973064	0.43814433	0	0.522309711	2
15	0.223163842	0.728434505	0.089655	0.6867	0.18	0.3131313	0.556701031	1	0.724409449	1
16	0	0.156549521	0.117241	0	0.74	0	0	0.316939891	0.293963255	?

Keterangan:

 = data latih

 = data uji

3.5.3 Menghitung jarak antara record baru pada data uji dengan tiap record pada data latih

Proses perhitungan jarak antara record pertama data latih dengan record uji (record ke-16) menggunakan persamaan 2.2 adalah sebagai berikut.

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

$$\text{Jarak} = \sqrt{\begin{aligned} &(0 - 0.550847458)^2 + (0.156549521 - 0.434504792)^2 + \\ &(0.117241 - 0.417241)^2 + (0 - 0.5764)^2 + \\ &(0.74 - 0.55)^2 + (0 - 0.6498316)^2 + \\ &(0 - 0.703608247)^2 + (0.316939891 - 0.300546448)^2 + \\ &(0.293963255 - 0)^2 \end{aligned}}$$

$$= 1.3576$$

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

Tabel 3. 3 Hasil Perhitungan jarak

No record	Jarak	Rank minimum distance	Target
1	1.3576	3	2
2	2.0601	14	2
3	1.7448	10	2
4	1.5599	5	2
5	1.6248	8	2
6	1.8139	11	2
7	1.5832	6	2
8	1.6404	9	2
9	1.6229	7	2
10	1.967	12	2
11	2.0236	13	2
12	2.1062	15	2
13	1.3494	2	1
14	0.9693	1	1
15	1.4916	4	1

Setelah hasil perhitungan jarak didapatkan proses selanjutnya ialah mengurutkan jarak dari nilai minimum ke jarak dengan nilai yang maksimum (*rank minimum distance*). Hasil perhitungan jarak setelah diurutkan ditunjukkan pada tabel 3.4.

Tabel 3. 4 Hasil Perhitungan Jarak Setelah diurutkan

No record	Jarak	Rank minimum distance	Target
14	0.9693	1	1
13	1.3494	2	1
1	1.3576	3	2
15	1.4916	4	1
4	1.5599	5	2
7	1.5832	6	2
9	1.6229	7	2
5	1.6248	8	2
8	1.6404	9	2
3	1.7448	10	2
6	1.8139	11	2
10	1.967	12	2
11	2.0236	13	2
2	2.0601	14	2
12	2.1062	15	2

3.5.4 Menentukan k record terdekat

Berdasarkan hasil perhitungan jarak pada tabel 3.3, kemudian dilakukan pengurutan terhadap jarak *record* yang ditunjukkan pada tabel 3.4. Kemudian diambil k record dengan jarak terkecil. Apabila ditentukan k = 3, maka record yang terpilih adalah record ke-14, 13 dan 1.

3.5.5 Menentukan maximum membership dan kelas target

Proses menentukan maximum membership dan kelas target dengan mencari nilai membership untuk tiap kelas j dengan menggunakan persamaan 2.4. $K=15$ (jumlah data latih), $n_T = 3$, $n_{TT} = 12$, $_{TT}$ =tidak training, $_T$ =training

$$\begin{aligned}
 u_{T(T)} &= 0.51 + (3/15) * 0.49 \\
 &= 0.51 + 0.2 * 0.49 \\
 &= 0.51 + 0.098 \\
 &= 0.608
 \end{aligned}$$

$$\begin{aligned}
 u_{T(TT)} &= (3/15)*0.49 \\
 &= 0.2*0.49 \\
 &= 0.098
 \end{aligned}$$

$$\begin{aligned}
 u_{TT(TT)} &= 0.51+(12/15)*0.49 \\
 &= 0.51+0.8*0.49 \\
 &= 0.51+0.392 \\
 &= 0.902
 \end{aligned}$$

$$\begin{aligned}
 u_{TT(T)} &= (12/15)*0.49 \\
 &= 0.8*0.49 \\
 &= 0.392
 \end{aligned}$$

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

$$\begin{aligned}
 u_T(x) &= \frac{\left(0.608*\left(0.9693^{-\frac{2}{2-1}}\right)\right)+\left(0.608*\left(1.3494^{-\frac{2}{2-1}}\right)\right)+\left(0.098*\left(1.3576^{-\frac{2}{2-1}}\right)\right)}{\left(0.9693^{-\frac{2}{2-1}}\right)+\left(1.3494^{-\frac{2}{2-1}}\right)+\left(1.3576^{-\frac{2}{2-1}}\right)} \\
 &= \frac{0.64+0.33+0.05}{1.06+0.55+0.54} \\
 &= \frac{1.02}{2.15} \\
 &= 0.47
 \end{aligned}$$

$$\begin{aligned}
 u_{TT}(x) &= \frac{\left(0.392*\left(0.9693^{-\frac{2}{2-1}}\right)\right)+\left(0.392*\left(1.3494^{-\frac{2}{2-1}}\right)\right)+\left(0.902*\left(1.3576^{-\frac{2}{2-1}}\right)\right)}{\left(0.9693^{-\frac{2}{2-1}}\right)+\left(1.3494^{-\frac{2}{2-1}}\right)+\left(1.3576^{-\frac{2}{2-1}}\right)} \\
 &= \frac{0.42+0.22+0.49}{1.06+0.55+0.54} \\
 &= \frac{1.13}{2.15} \\
 &= 0.53
 \end{aligned}$$

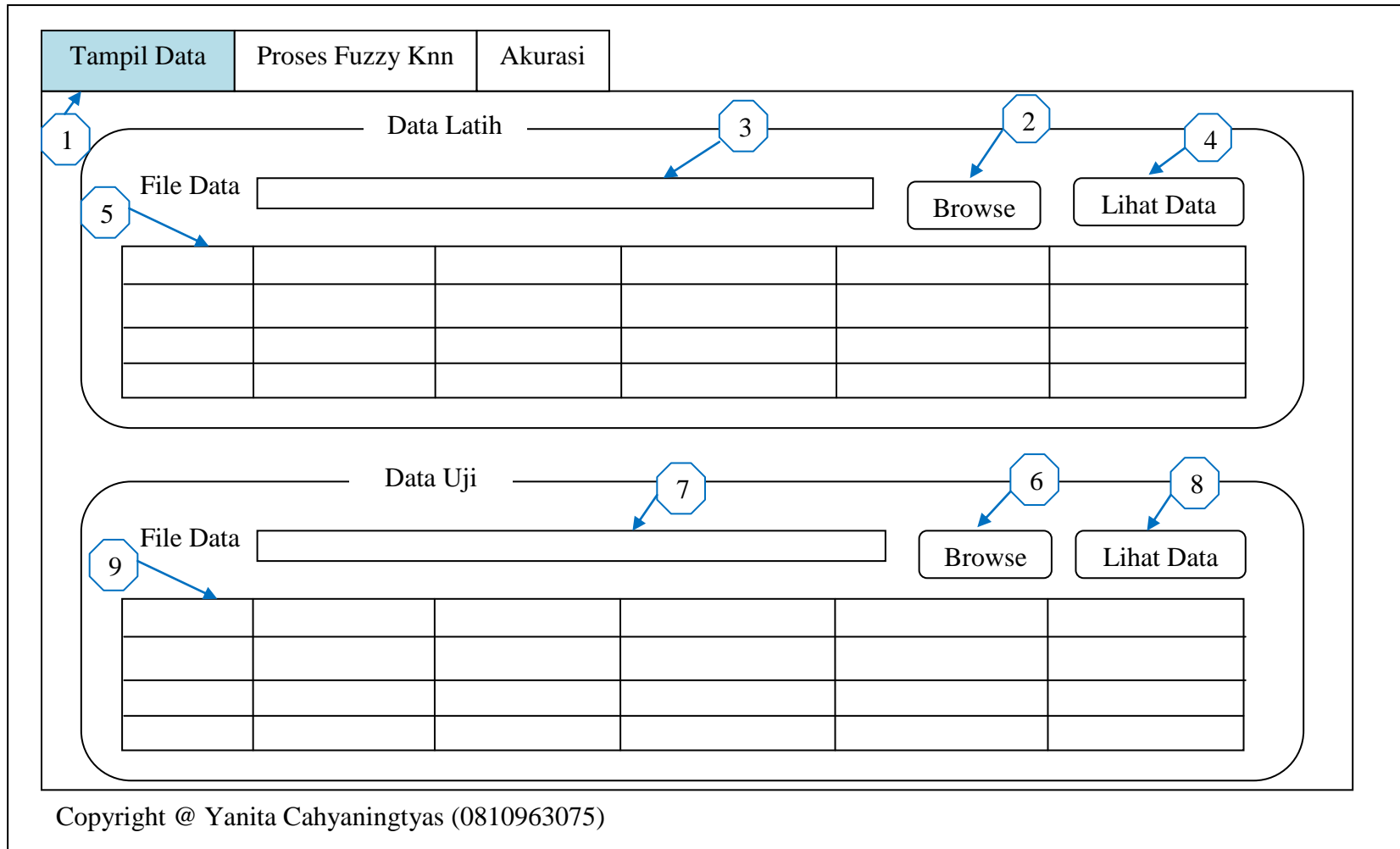
Berdasarkan hasil perhitungan nilai keanggotaan didapat dua nilai keanggotaan, untuk menentukan kelas target maka dipilih nilai keanggotaan terbesar yaitu 0.53 sehingga kelas targetnya yaitu 2 (tidak training). Jadi, status evaluasi kinerja karyawan yang didapat ialah karyawan tersebut tidak mengikuti training atau pelatihan.

3.6 Perancangan Antarmuka

Pada bagian perancangan antarmuka ini akan dijelaskan antarmuka sistem yang akan dibuat. Tampilan rancangan antarmuka untuk menampilkan data yang digunakan ditunjukkan pada gambar 3.6.

Berikut adalah penjelasan gambar 3.6 berdasarkan nomor:

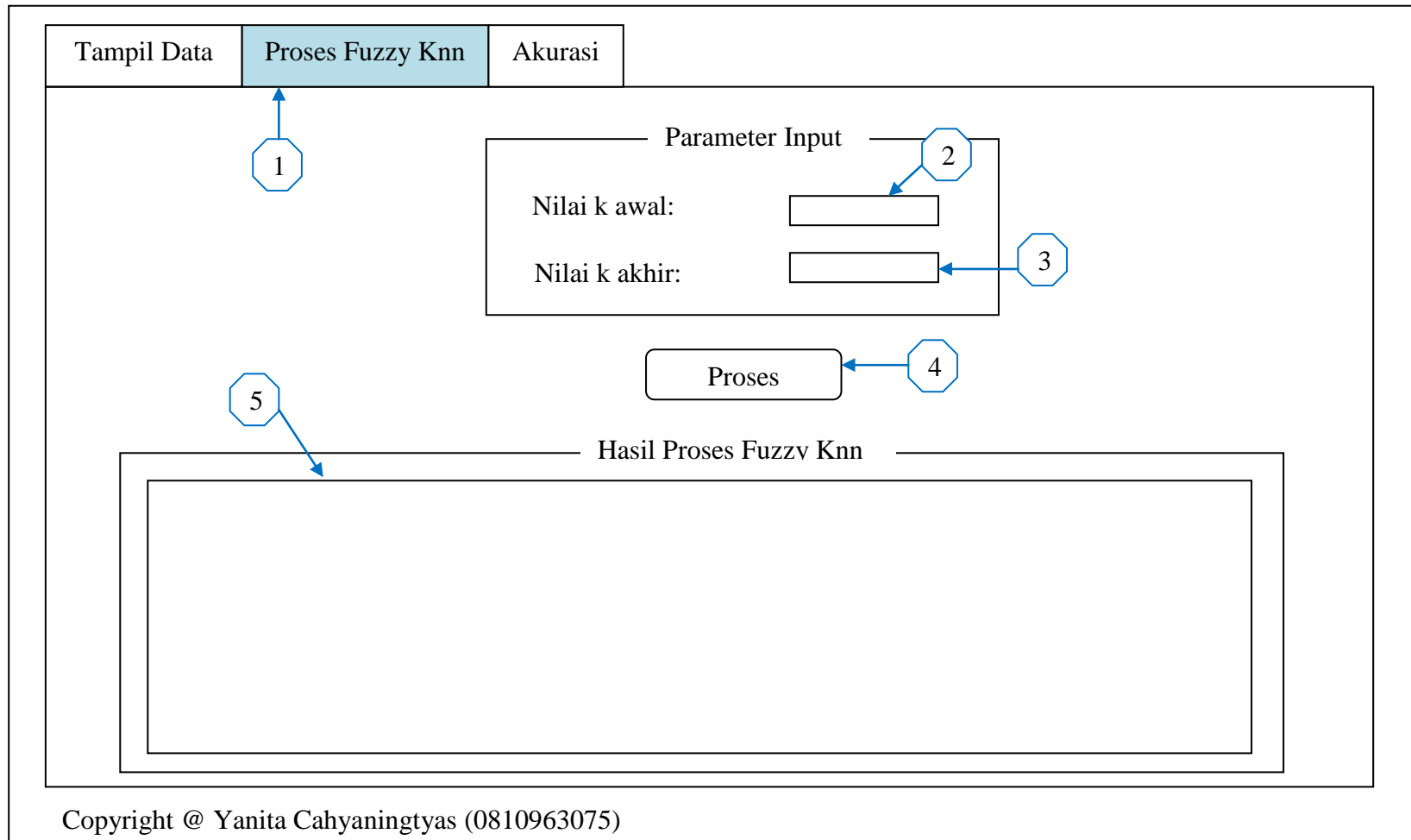
1. Tab Tampil Data digunakan untuk mengarahkan pada halaman tampil data.
2. Tombol Browse pada data latih digunakan untuk mengambil data latih evaluasi kinerja karyawan.
3. Textbox pada data latih untuk menampilkan nama data latih yang dicari.
4. Tombol Lihat Data pada data latih digunakan untuk proses menampilkan data latih yang telah di cari.
5. Bagian ini menampilkan data latih.
6. Tombol Browse pada data uji digunakan untuk mengambil data uji evaluasi kinerja karyawan.
7. Textbox pada data uji untuk menampilkan nama data uji yang dicari.
8. Tombol Lihat Data pada data uji digunakan untuk proses menampilkan data uji yang telah di cari.
9. Bagian ini menampilkan data uji.



Gambar 3. 6 Tampilan Rancangan Antarmuka untuk Menampilkan Data

Antarmuka proses Fuzzy Knn yang akan menghasilkan klasifikasi ditunjukkan pada gambar 3.7. Penjelasan gambar 3.7 berdasarkan nomor:

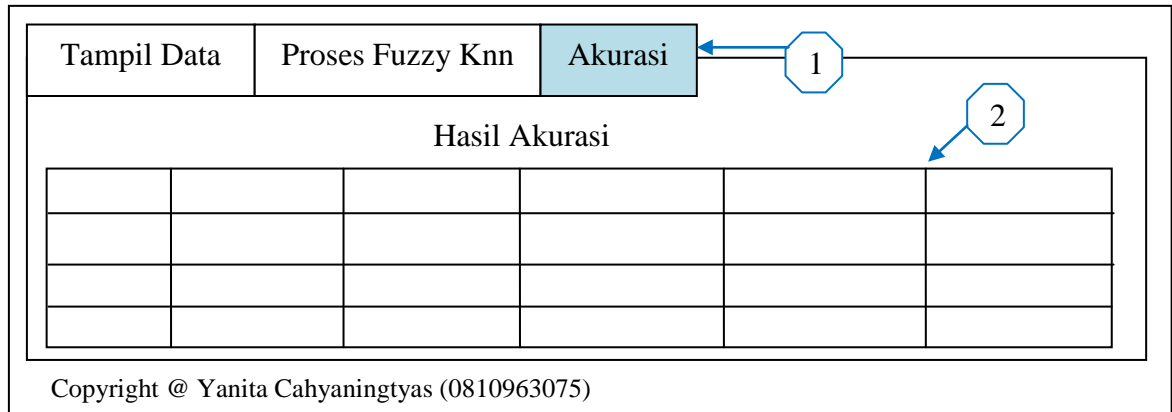
1. Tab Proses Fuzzy Knn digunakan untuk mengarahkan pada halaman Fuzzy Knn.
2. Textbox untuk menampilkan nilai k awal yang dimasukkan.
3. Textbox untuk menampilkan nilai k akhir yang dimasukkan.
4. Tombol Proses digunakan untuk melakukan proses Fuzzy Knn.
5. Bagian ini menampilkan hasil dari proses Fuzzy Knn.



Gambar 3. 7 Tampilan Rancangan Antarmuka Fuzzy Knn

Antarmuka akurasi ditunjukkan pada gambar 3.8, penjelasan gambar 3.8 berdasarkan nomor:

1. Tab akurasi digunakan untuk mengarahkan pada halaman akurasi.
2. Bagian ini menampilkan hasil akurasi dari proses Fuzzy Knn.



Gambar 3. 8 Tampilan Rancangan Antarmuka Akurasi

3.7 Perancangan Uji Coba

Setelah sistem selesai dibuat, langkah selanjutnya adalah melakukan pengujian terhadap sistem tersebut. Pengujian dilakukan untuk mengetahui tingkat akurasi dari hasil klasifikasi evaluasi kinerja karyawan 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.

3.7.1 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 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 80, 110, 140, 170, 200 dan 230 record.

Dari pengujian ini, diperoleh tingkat akurasi terhadap nilai k dan jumlah data latih. Tabel 3.5 menampilkan rancangan tabel yang akan digunakan untuk mencatat hasil dari pengujian ini.

Tabel 3. 5 Uji Pengaruh nilai k dan data latih terhadap tingkat akurasi

k	Jumlah Data Latih	Akurasi (%)
2	80	
...	...	
n	n	

Keterangan:

k : nilai k yang akan diuji.

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

Akurasi : tingkat akurasi dalam persen, yang dihitung dengan persamaan 2.5.

BAB IV

IMPLEMENTASI DAN PEMBAHASAN

Bab ini berisi implementasi system serta pembahasan dan analisa mengenai hasil uji coba terhadap sistem. Implementasi sistem merupakan penerapan dari rancangan sistem yang sudah dibahas pada bab sebelumnya.

4.1 Lingkungan Implementasi

Lingkungan implementasi yang dijelaskan dalam sub bab ini adalah lingkungan implementasi perangkat keras dan perangkat lunak.

4.1.1 Lingkungan Perangkat Keras

Perangkat keras yang digunakan dalam pengembangan fuzzy k-nearest neighbor (fuzzy k-nn) adalah:

1. Prosesor Intel® Core™ Duo T6600 CPU @ 2.2 GHz.
2. Memori 1 Gb.
3. Harddisk 320 Gb.

4.1.2 Lingkungan Perangkat Lunak

Perangkat lunak yang digunakan dalam pengembangan fuzzy k-nearest neighbor (fuzzy k-nn) adalah:

1. Sistem operasi yang digunakan yaitu Windows 7 Ultimate.
2. Aplikasi dibangun dengan bahasa pemrograman Java menggunakan Java Development Kit (JDK) 1.6 dan ditulis melalui editor Netbeans IDE 7.1.1.
3. Microsoft Office Excel 2007.

4.2 Implementasi Program

Berdasarkan analisa dan perancangan proses yang telah dipaparkan pada BAB III, maka pada bab ini akan dijelaskan proses-proses implementasinya.

4.2.1 Implementasi *class*

Pada implementasi program, program dibuat secara modular dalam bentuk kelas-kelas yang dikemas dalam suatu package. Program

diimplementasikan dalam 1 package yaitu `programfuzzyknn`. Package `programfuzzyknn` berisi 8 kelas yaitu `class OlahData`, `class FuzzyKnn`, `class DataSort`, `class MapValueSort`, `class Demo`, `class KodeGUI`, `class Akurasi` dan `class ProgramFuzzyKnn`.

1. Class OlahData

Class `OlahData` merupakan kelas public yang berisi metode/fungsi untuk menangani akses ke data yang akan digunakan dan proses normalisasi dari data pelatihan. Deskripsi singkat metode/fungsi yang dimiliki class `OlahData` dapat dilihat pada tabel 4.1.

Tabel 4. 1 Deskripsi metode/fungsi pada class `OlahData`

Tipe	Deskripsi metode/fungsi
String []	<u>getNamaKolom</u> () mendapatkan nama kolom tabel.
double[][]	<u>getData</u> (int s) mendapatkan data hasil pembacaan database. parameter: s - letak sheet yang akan digunakan
int[]	<u>getKelas</u> (int sheet) mendapatkan kelas dari data pelatihan. parameter: sheet - nama sheet yang akan digunakan data[][] - data yang digunakan
int[][]	<u>getIndex</u> (int[] kelas) mendapatkan nomor indeks dan kelas pada data pelatihan. parameter: kelas - nama kelas indeks[][] - nomor indeks data
double[][]	<u>getIndexDistance</u> (double[] distance)

	<p>mendapatkan nomor indeks dan nilai jarak.</p> <p>parameter:</p> <p>distance – nilai jarak</p> <p>indeks[][] – nomor indeks data</p>
double	<p><u>getMax</u>(double[][]inputdata, int kolom)</p> <p>mendapatkan nilai maksimum dari setiap kolom pada data pelatihan.</p> <p>parameter:</p> <p>inputdata – data yang dimasukkan</p> <p>kolom – nama kolom</p>
double []	<p><u>getMaxNormalisasi</u> (double[][]inputdata)</p> <p>mendapatkan nilai maksimum dari setiap kolom pada data pelatihan untuk proses normalisasi</p> <p>parameter:</p> <p>inputdata – data yang dimasukkan</p> <p>nilaimax[] - nilai maksimum</p>
double	<p><u>getMin</u> (double[][]inputdata, int kolom)</p> <p>mendapatkan nilai minimum dari setiap kolom pada data pelatihan.</p> <p>parameter:</p> <p>inputdata – data yang dimasukkan</p> <p>kolom – nama kolom</p>
double []	<p><u>getMinNormalisasi</u> (double[][]inputdata)</p> <p>mendapatkan nilai minimum dari setiap kolom pada data pelatihan untuk proses normalisasi</p> <p>parameter:</p> <p>inputdata – data yang dimasukkan</p> <p>nilaimin[] - nilai minimum</p>
double []	<p><u>getRange</u> (double[][]inputdata)</p> <p>mendapatkan nilai range dari nilai maksimum dan minimum</p>

	<p>yang sudah didapatkan sebelumnya.</p> <p>parameter:</p> <p>inputdata - data yang dimasukkan</p> <p>max[] - nilai maksimum</p> <p>min[] - nilai minimum</p>
double[][]	<p><u>getNormalisasi</u>(double[][]inputdata)</p> <p>mendapatkan hasil normalisasi data</p> <p>parameter:</p> <p>inputdata - data yang dimasukkan</p> <p>min[] - nilai minimum</p> <p>range[] - nilai range</p> <p>normalisasi[][] - proses normalisasi</p>

2. Class FuzzyKnn

`class Knn` merupakan kelas public yang berisi fungsi/metode untuk melakukan proses fuzzy knn pada data. Deskripsi singkat mengenai metode/fungsi yang ada di dalam `class FuzzyKnn` dapat dilihat pada tabel 4.2.

Tabel 4. 2 Deskripsi metode/fungsi pada class FuzzyKnn

Tipe	Deskripsi metode/fungsi
double[][]	<p><u>getDistance</u>(double[][]normalisasiuji, double[][]normalisasilatih)</p> <p>mendapatkan nilai distance.</p> <p>parameter:</p> <p>normalisasiuji - nilai normalisasi data uji</p> <p>normalisasilatih - nilai normalisasi data latih</p> <p>distance - proses distance</p>
double[]	<p><u>getdataperklpk</u>(double[][]distance,int kolom)</p> <p>mendapatkan nilai distance serta nama kolom kelas.</p> <p>parameter:</p>

	<p>distance - proses distance</p> <p>kolom - nama kolom</p> <p>dataperklpk[] - nilai distance</p>
String[]	<p><u>intToString</u> (int [] indeks)</p> <p>mengubah tipe data integer ke dalam tipe data string.</p> <p>parameter:</p> <p>indeks- nomor indeks</p>
String[]	<p><u>doubleToString</u> (int [] indeks)</p> <p>mengubah tipe data double ke dalam tipe data string.</p> <p>parameter:</p> <p>indeks- nomor indeks</p>
DataSort []	<p><u>getSort</u> (int [] indeks, double [] distance, int [] kelas)</p> <p>mendapatkan hasil pengurutan/sorting yang terdiri dari nomor indeks, nilai distance serta nomor kelas.</p> <p>parameter:</p> <p>indeks_[] - nomor indeks data dengan tipe data String</p> <p>distance - nilai jarak dengan tipe data String</p> <p>kelas - nomor kelas</p>
void	<p><u>setK</u> (int k)</p> <p>mengatur nilai K.</p> <p>parameter:</p> <p>k - nilai k</p>
ArrayList<DataSort>	<p><u>getK</u> (DataSort [] data)</p> <p>mendapatkan data dari nilai K yang diinputkan.</p> <p>parameter:</p> <p>data - data yang dimasukkan</p> <p>datasort - nilai datasort</p>
int	<p><u>getJlhDataLatih</u> (double [] datalatih)</p> <p>mengetahui jumlah data latih</p> <p>parameter:</p>

	datalatih-nilai datalatih
int	<p><u>getTraining</u>(double[]a, int b) mendapatkan jumlah kelas training dan tidak training. parameter: a- nilai datalatih b- nilai kelas</p>
double	<p><u>getMembershipJI</u>(int nj, int K) mendapatkan nilai membership $j=i$. parameter: nj - jumlah kelas training/tidak training K - jumlah datalatih</p>
double	<p><u>getMembershipJtdkI</u>(int nj, int K) mendapatkan nilai membership $j \neq i$ parameter: nj - jumlah kelas training/tidak training K - jumlah datalatih</p>
double	<p><u>getMembership</u>(double[]membership, double[]distance) mendapatkan membership akhir penentu kelas parameter: membership-nilai membership $j=i/j \neq i$ distance - nilai distance</p>
double	<p><u>getMaxMembership</u>(double memT, double memTT) mendapatkan nilai maksimum membership untuk dijadikan penentu kelas parameter: memT - membership Training memTT - membership Tidak Training</p>
int	<p><u>getKelas</u>(double memT, double memTT) mendapatkan kelas target baru parameter:</p>

	<p>memT – membership Training</p> <p>memTT – membership Tidak Training</p>
--	--

3. Class DataSort

`class DataSort` merupakan kelas public yang berisi fungsi/metode yang berisi tipe data yang dibuat sendiri yaitu `public int record`, `public int kelas` dan `public double distance`.

4. Class MapValueSort

`class MapValueSort` merupakan kelas public yang diambil dari fungsi java yang sudah disediakan.

5. Class Demo

`class Demo` merupakan kelas public yang berisi metode/fungsi yang digunakan untuk mengeluarkan hasil proses dari kelas-kelas yang sudah dijabarkan.

6. Class Akurasi

`class Akurasi` merupakan kelas public yang berisi fungsi/metode untuk melakukan proses akurasi data. Deskripsi singkat mengenai metode/fungsi yang ada di dalam `class Akurasi` dapat dilihat pada tabel 4.3.

Tabel 4. 3 Deskripsi metode/fungsi pada class Akurasi

Tipe	Deskripsi metode/fungsi
Double	<p><u>akurasi</u> (<code>ArrayList datasort</code>, <code>Arraylist datalatih</code>) mendapatkan nilai akurasi.</p> <p>parameter:</p> <p><code>datasort</code> – nilai kelas pada data distance yang sudah di sorting</p> <p><code>datalatih</code> – nilai kelas pada datalatih</p>

7. Class ProgramFuzzyKnn

`class ProgramFuzzyKnn` merupakan kelas `public` yang berisi fungsi/metode untuk menampilkan hasil proses di GUI.

4. 2.2 Implementasi proses pelatihan

Proses pelatihan terdiri dari 3 sub proses yaitu proses normalisasi nilai setiap atribut, proses klasifikasi algoritma *fuzzy k-nearest neighbor*, proses menentukan maximum dan kelas target. Proses normalisasi nilai setiap atribut bertujuan untuk membakukan skala pengaruh yang ada pada atribut terhadap hasil. Proses klasifikasi dengan algoritma *fuzzy k-nearest neighbor* yang bertujuan untuk mengklasifikasikan hasil perhitungan jarak antara dua record menggunakan *Euclidean distance*. Proses menentukan maximum membership dan kelas target digunakan untuk mencari nilai membership tiap kelas dimana dari hasil tersebut diambil nilai maksimum membershipnya untuk menjadi penentu kelas target hasil pengujian.

4. 2.2.1 Proses normalisasi nilai setiap atribut

Proses normalisasi dimulai dengan mengambil data dari Microsoft Excel yang terdiri dari nama kolom, isi kolom (data pelatihan), nomor kelas, nomor indeks. Setelah proses pengambilan data dilakukan langkah selanjutnya ialah dicari nilai maksimum dan minimum pada data latih dan data uji dan didapatkan nilai range data latih dan data uji. Setelah didapatkan range data latih dan data uji dilakukan perhitungan untuk menghitung hasil normalisasi pada data latih dan data uji. Proses normalisasi nilai setiap atribut dihitung berdasarkan persamaan 2.1. Fungsi perhitungan normalisasi ditunjukkan oleh *source code* 4.1.

```

1 public double[][] getNormalisasi(double[][] inputdata) {
2     double[] min=getMinNormalisasi(inputdata);
3     double[] range = getRange(inputdata);
4     double[][]normalisasi=new
5         double[inputdata.length][min.length];
6     for (int i = 0; i < normalisasi.length; i++) {
7         for (int j = 0; j < normalisasi[i].length;
8             j++) {
9             normalisasi[i][j] = ((inputdata[i][j] -
10                min[j]) /
11                range[j]);
12         }
13     }
14     return normalisasi;
15 }

```

Source Code 4. 1 Fungsi untuk menghitung hasil normalisasi

4. 2.2.2 Proses klasifikasi dengan algoritma *fuzzy k-nearest neighbor*

Perhitungan fuzzy knn dilakukan dengan 4 langkah yaitu langkah pertama ialah menghitung distance dari data latih dan data uji yang telah dinormalisasi.

Kode program untuk menghitung distance ditunjukkan oleh source code 4.2.

```

1 public double[][] getDistance(double[][] normalisasiuji, double[][]
2 normalisasilatih) {
3     double[][]distance=new
4         double[normalisasilatih.length][normalisasiuji.
5             length];
6     for (int i = 0; i < normalisasiuji.length; i++) {
7         for (int j = 0; j < normalisasilatih.length; j++) {
8             double tempJarak = 0;
9             for(int k = 0; k < normalisasilatih[0].length; k++)
10                {
11                    tempJarak+=Math.pow(normalisasiuji[i][k]-
12                        normalisasilatih[j][k],2);
13                }
14            distance[j][i] = Math.sqrt(tempJarak);
15        }
16    }
17    return distance;}

```

Source Code 4. 2 Fungsi untuk menghitung distance

Langkah kedua ialah mengurutkan hasil distance secara ascending (dari jarak yang paling kecil ke jarak yang paling besar). Kode program untuk mengurutkan hasil distance ditunjukkan pada source code 4.3.

```

1 public DataSort[] getSort(int[] indeks, double[] distance, int[]
2 kelas) {
3     String[] indeks_ = intToString(indeks);
4     String[] dist = doubleToString(distance);
5     DataSort[] data = valueSort.getDataSort(indeks_, dist);
6     DataSort[] dataNew = new DataSort[data.length];
7     for (int i = 0; i < dataNew.length; i++) {

```

```

8         dataNew[i] = new DataSort();
9         dataNew[i].record = data[i].record;
10        dataNew[i].distance = data[i].distance;
11        dataNew[i].kelas= kelas[dataNew[i].record];
12    }
13    return dataNew;
14 }

```

Source Code 4.3 Fungsi untuk mengurutkan hasil distance

Langkah ketiga ialah mengambil hasil distance sesuai dengan nilai k yang dimasukkan yang ditunjukkan oleh source code 4.4.

```

1 public ArrayList<DataSort> getK(DataSort[] data) {
2     ArrayList<DataSort>datasort=new ArrayList<DataSort>();
3     for (int i = 0; i < k_; i++) {
4         datasort.add(data[i]);
5     }
6     return datasort;
7 }

```

Source Code 4.4 Fungsi untuk mengambil hasil distance berdasarkan nilai k

Langkah keempat ialah menghitung nilai keanggotaan tiap kelas yang ditunjukkan oleh source code 4.5.

```

1 public double getMembership(double[] membership, double[] distance) {
2     double M = 0;
3     double pembilang = 0;
4     double penyebut = 0;
5     for (int i = 0; i < distance.length; i++) {
6         pembilang+= membership[i]*(Math.pow(distance[i], -2));
7         penyebut += (Math.pow(distance[i], -2));
8     }
9     M = pembilang / penyebut;
10    return M;
11 }

```

Source Code 4.5 Fungsi untuk menghitung nilai keanggotaan tiap kelas

Setelah didapatkan nilai keanggotaan tiap kelas maka dilakukan pencarian nilai maksimum membership yang ditunjukkan oleh source code 4.6.

```

1 public double getMaxMembership(double memT, double memTT) {
2     double maxmembership = 0;
3     if (memT > memTT) {
4         maxmembership = memT;
5     } else {
6         maxmembership = memTT;
7     }
8     return maxmembership;
9 }

```

Source Code 4.6 Fungsi untuk mencari nilai maksimum membership

Langkah selanjutnya ialah menentukan kelas dari hasil nilai maksimum membership yang didapat sebelumnya yang ditunjukkan oleh source code 4.7.

1	<code>public int getKelas(double memT, double memTT) {</code>
2	<code> int klsmembership = 0;</code>
3	<code> if (memT > memTT) {</code>
4	<code> klsmembership = 1;</code>
5	<code> } else {</code>
6	<code> klsmembership = 2;</code>
7	<code> }</code>
8	<code> return klsmembership;</code>
9	<code>}</code>

Source Code 4. 7 Fungsi untuk menentukan kelas

4.3 Penerapan Aplikasi

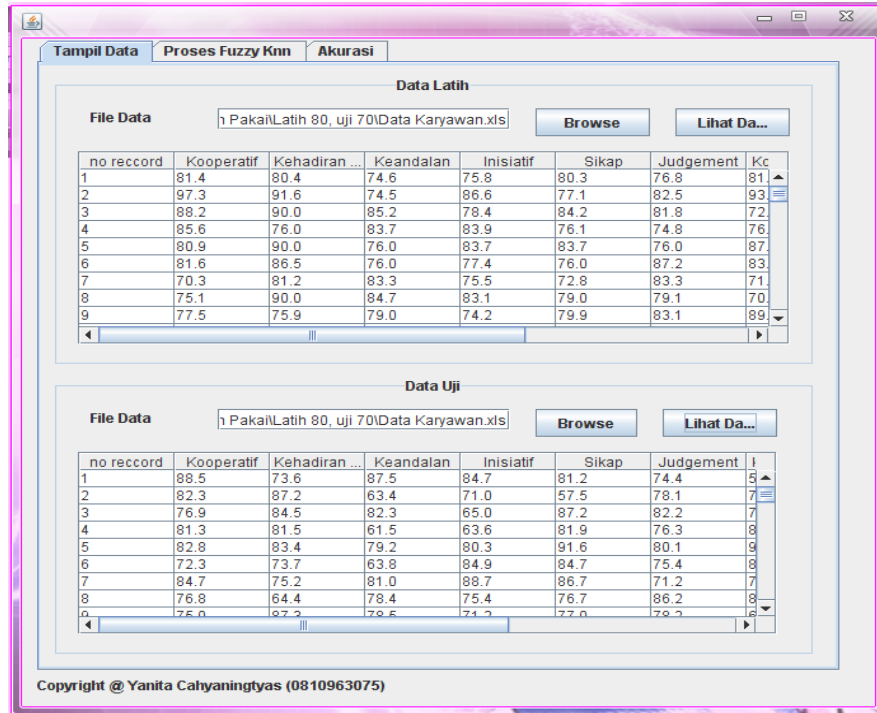
Aplikasi yang dibangun untuk melakukan berbagai proses dalam melakukan klasifikasi data evaluasi karyawan dengan menggunakan metode fuzzy knn. Berikut ini adalah beberapa tampilan (interface) aplikasi pada setiap menu berdasarkan penjelasan subbab 3.6 Perancangan Antarmuka.

4.3.1 Implementasi antarmuka

Implementasi antarmuka dari sistem terdiri dari form tampil data, proses fuzzy k-nn, dan akurasi. Form tampil data yang didalamnya berisi implementasi request data, untuk form proses fuzzy k-nn didalamnya berisi implementasi input nilai k dan penanganan proses klasifikasi data evaluasi karyawan dan form akurasi didalamnya berisi hasil akurasi dari sistem.

4.3.1.1 Tabpage Tampil Data

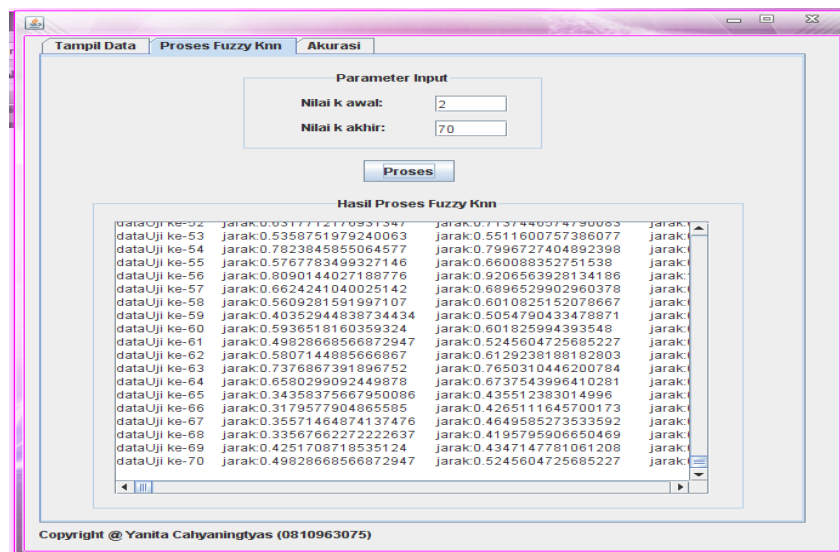
Pada tampilan rancangan antarmuka tampil data terdapat tombol browse yang berguna untuk merequest file data evaluasi karyawan dan kemudian nama data yang dipilih dapat muncul pada textarea yang sudah disediakan. Untuk menampilkan data latih dan data uji yang sudah dipilih digunakan tombol lihat data sehingga data latih dan uji dapat muncul pada listview (tabel) yang ditunjukkan pada gambar 4.1.



Gambar 4. 1 Tampilan rancangan antarmuka tampil data

4. 3.1.2 Tabpage Fuzzy Knn

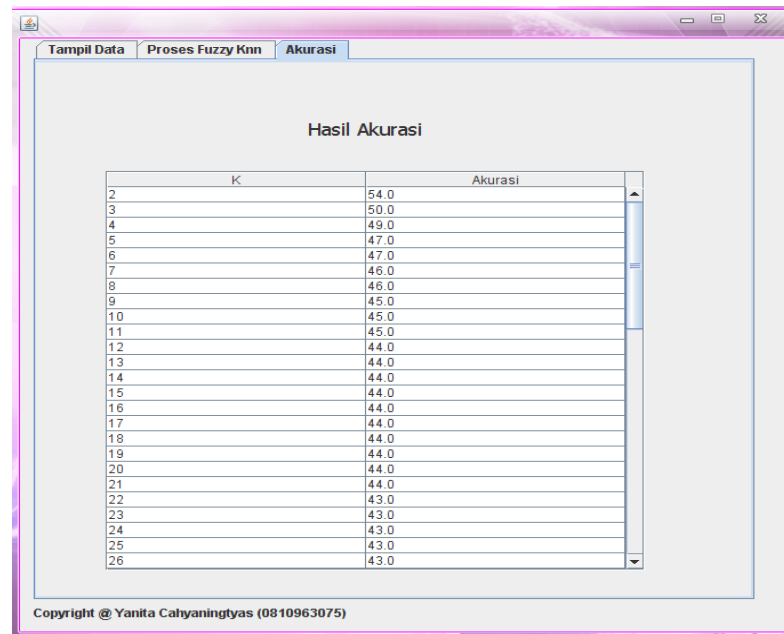
Pada tampilan rancangan antarmuka fuzzy k-nn terdapat tombol proses. Tombol proses berguna untuk melakukan proses fuzzy k-nn setelah diinputkan nilai k awal sebagai batas awal nilai k dan nilai k akhir sebagai batas akhir nilai k secara manual yang akan membantu proses pengujian data. Batas akhir nilai k disini dibatasi sebanyak data uji. Setelah menekan tombol proses, hasilnya ditampilkan pada hasil proses fuzzy k-nn yang ditunjukkan pada gambar 4.2.



Gambar 4. 2 Tampilan rancangan antarmuka Fuzzy K-NN

4.3.1.3 Tabpage Akurasi

Pada tampilan rancangan antarmuka akurasi berguna untuk menampilkan hasil akurasi dari proses fuzzy k-nn, hasil akurasi akan muncul bersamaan dengan ditekannya tombol proses pada tabpage proses fuzzy k-nn yang ditunjukkan pada gambar 4.3.



K	Akurasi
2	54.0
3	50.0
4	49.0
5	47.0
6	47.0
7	46.0
8	46.0
9	45.0
10	45.0
11	45.0
12	44.0
13	44.0
14	44.0
15	44.0
16	44.0
17	44.0
18	44.0
19	44.0
20	44.0
21	44.0
22	43.0
23	43.0
24	43.0
25	43.0
26	43.0

Copyright @ Yanita Cahyaningtyas (0810963075)

Gambar 4. 3 Tampilan rancangan antarmuka Akurasi

4.4 Implementasi Pengujian

Pada pengujian klasifikasi digunakan dataset evaluasi karyawan yang terdiri dari 300 record. Data yang digunakan dalam sistem yaitu data evaluasi karyawan yang hasil evaluasi dibagi menjadi dua kelas. Pada data ini atribut yang digunakan antara lain yaitu kooperatif, kehadiran dan ketepatan waktu, keandalan, inisiatif, sikap, judgement, komunikasi, hubungan antar manusia, keahlian profesional.

4.4.1 Hasil Uji

Seperti yang telah dibahas pada bab 3, terdapat satu macam pengujian yang dilakukan, yaitu untuk mengetahui pengaruh nilai k dan data latih terhadap tingkat akurasi.

Pada pengujian ini terdapat dua parameter antara lain menentukan jumlah data latih dan jumlah nilai k . Pada pengujian jumlah data latih terdapat 6 macam data latih yang terdiri dari 80,110,140,170, 200 dan 230 data dengan perulangan nilai $k = 2$ sampai dengan $k = 40$.

Pengujian data evaluasi kinerja karyawan ini didapatkan hasil pengujian dengan nilai akurasi maksimum pada jumlah 80 data latih adalah 54% dengan $k = 2$ dan nilai akurasi minimumnya tetap pada $k = 22$ sampai dengan $k = 40$ yaitu 43%.

Untuk 110 data latih didapatkan akurasi maksimumnya pada $k = 2$ yaitu 52% dan akurasinya mengalami penurunan dengan bertambahnya jumlah nilai k . Sedangkan akurasi minimum yang didapat ialah 43%.

Pada 140 data latih, akurasi maksimumnya ialah 53% yang didapat pada nilai $k=4$. Sedangkan untuk nilai akurasi minimumnya yaitu 44%.

Untuk pengujian 170 data latih, akurasinya mengalami penurunan dengan akurasi maksimumnya 52% dan akurasi minimumnya yaitu 44%.

Pengujian data evaluasi kinerja karyawan untuk 200 data latih akurasinya mengalami penurunan, dengan akurasi maksimum yang didapat yaitu 54% dan hasil akurasi minimumnya tetap pada $k = 24$ sampai dengan $k = 40$ yaitu 44%.

Sedangkan untuk pengujian 230 data latih akurasi maksimumnya 52% dan akurasi minimumnya yaitu 44%. Hasil pengujian dataset evaluasi kinerja karyawan ditunjukkan pada tabel 4.4.

Tabel 4. 4 Hasil pengujian dataset evaluasi kinerja karyawan

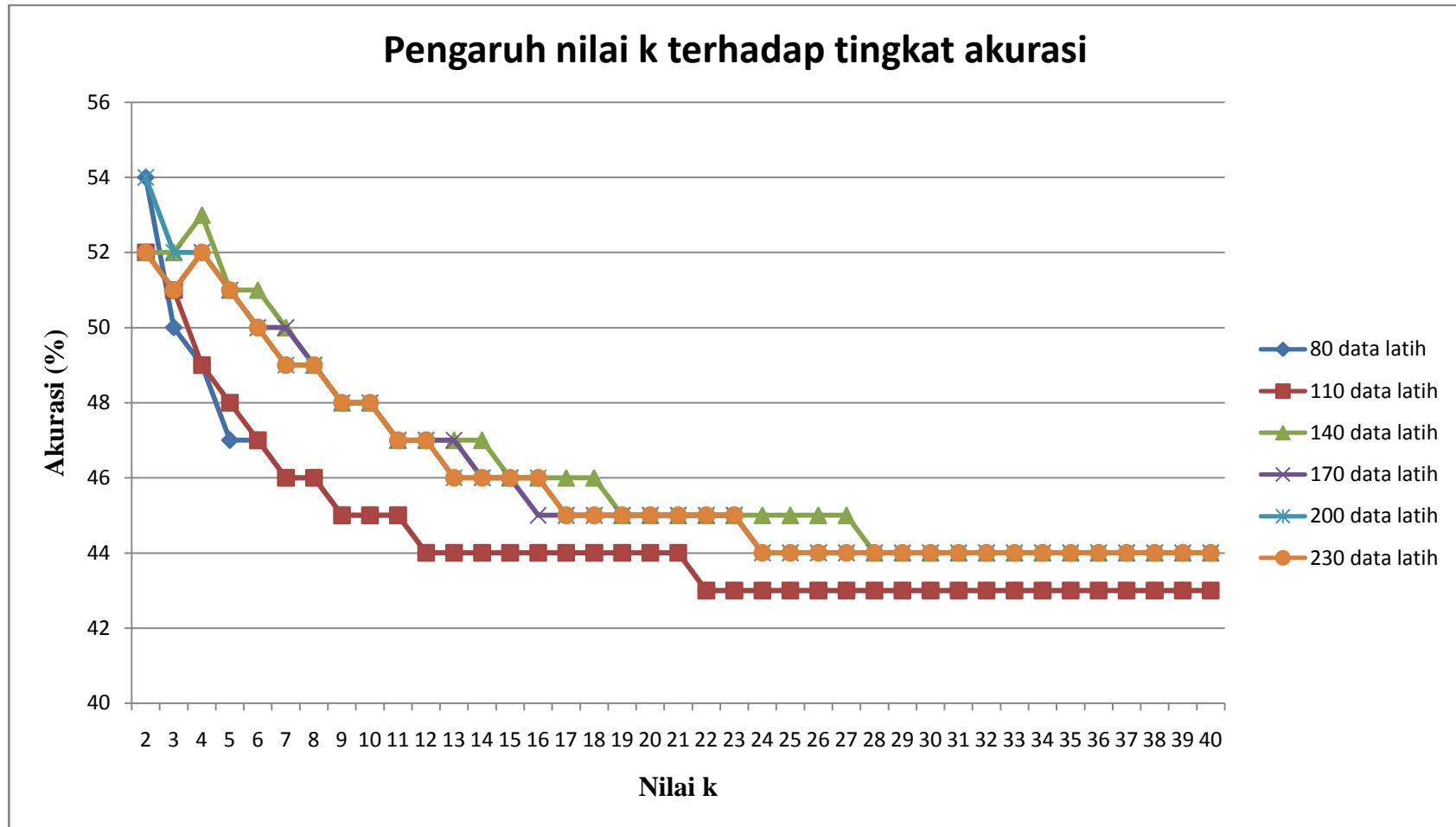
k	Akurasi (%) pada					
	80 data latih	110 data latih	140 data latih	170 data latih	200 data latih	230 data latih
2	54	52	52	52	54	52
3	50	51	52	51	52	51
4	49	49	53	52	52	52
5	47	48	51	51	51	51
6	47	47	51	50	50	50
7	46	46	50	50	49	49
8	46	46	49	49	49	49
9	45	45	49	48	48	48
10	45	45	48	48	48	48
11	45	45	48	47	47	47
12	44	44	47	47	47	47
13	44	44	47	47	46	46
14	44	44	47	46	46	46
15	44	44	46	46	46	46
16	44	44	46	45	46	46
17	44	44	46	45	45	45
18	44	44	46	45	45	45
19	44	44	45	45	45	45
20	44	44	45	45	45	45
21	44	44	45	45	45	45
22	43	43	45	45	45	45
23	43	43	45	45	45	45
24	43	43	45	44	44	44
25	43	43	45	44	44	44
26	43	43	45	44	44	44
27	43	43	45	44	44	44
28	43	43	44	44	44	44
29	43	43	44	44	44	44
30	43	43	44	44	44	44
31	43	43	44	44	44	44
32	43	43	44	44	44	44
33	43	43	44	44	44	44
34	43	43	44	44	44	44
35	43	43	44	44	44	44
36	43	43	44	44	44	44

37	43	43	44	44	44	44
38	43	43	44	44	44	44
39	43	43	44	44	44	44
40	43	43	44	44	44	44

4.5 Analisa Hasil

Pada analisa hasil ini terdapat satu proses pengujian yang didalamnya terdapat 6 jumlah data latih yang terdiri dari 80, 110, 140, 170, 200 dan 230 data latih. Sedang untuk pengujian jumlah nilai k nya ada 39 yaitu $k = 2$ sampai $k = 40$. Maka dari penelitian yang didapat akan menghasilkan akurasi yang akan digunakan untuk melakukan penelitian.

Pada hasil pengujian, pengaruh nilai k sendiri dapat dilihat bahwa akurasi maksimum data evaluasi kinerja karyawan terdapat pada $k = 2$. Tetapi pengaruh nilai k pada akurasi minimum tidak sama pada setiap record, sehingga dapat dilihat pada grafik bahwa nilai k minimumnya tidak dominan seperti nilai maksimum. Dilihat pada gambar 4.4, jumlah 80 data latih terletak pada grafik warna biru, 110 data latih pada grafik warna merah, 140 data latih pada grafik warna hijau, 170 data latih pada grafik warna ungu, 200 data latih pada grafik warna biru muda dan data latih dengan jumlah 230 record terletak pada grafik warna kuning, mengalami penurunan yang stabil pada tiap record. Pengaruh nilai k terhadap tingkat akurasi ditunjukkan pada gambar 4.4.



Gambar 4. 4 Pengaruh nilai k terhadap tingkat akurasi

Berdasarkan hasil pengujian data evaluasi kinerja karyawan bahwa akurasi paling baik pada sebagian besar data latih adalah saat nilai $k = 2$, yaitu 52% dan 54%. Akurasi tersebut tidak mencapai lebih dari 60%, karena jumlah record untuk tiap kelas tidak seimbang, dimana record dengan kelas 2 mendominasi dataset. Selain dilakukan pengujian dengan salah satu kelas yang dominan pada data, dilakukan juga pengujian terhadap data latih dengan komposisi jumlah kelas yang sama. Didapatkan hasil pengujian dengan akurasi mengalami kenaikan sampai 66%, hasil akurasi dijabarkan pada lampiran 1.

Gambar 4.4 menunjukkan bahwa tingkat akurasi dari $k = 2$ sampai dengan $k = 40$ grafiknya mengalami penurunan sampai akurasi minimumnya yaitu 43%. Dan pada grafik yang ditampilkan oleh gambar 4.4 ditunjukkan nilai akurasi yang stagnan pada nilai $k = 22$ sampai dengan $k = 40$ yaitu 43% untuk 80 dan 110 data latih, untuk 140 data latih mengalami akurasi yang stagnan pada nilai $k = 28$ sampai dengan $k = 40$ yaitu 44%. Sedangkan untuk 170 sampai dengan 230 data latih mengalami nilai akurasi yang stagnan pada nilai $k = 24$ sampai dengan $k = 40$ yaitu sebesar 44%.

Hasil pengujian yang didapat untuk mengetahui pengaruh nilai k (tetangga terdekat) terhadap tingkat akurasi yaitu, semakin bertambahnya jumlah nilai k (tetangga terdekat) maka akurasi yang didapatkan semakin menurun dan stabil pada nilai k (tetangga terdekat) tertentu untuk masing masing data latih. Hal ini dikarenakan oleh nilai membership yang digunakan untuk penentuan kelas target pada penelitian bernilai maksimum. Kemudian akurasi k (tetangga terdekat) tertinggi cenderung pada $k = 2$, hal ini dikarenakan tetangga yang dibandingkan hanya dua tetangga saja. Untuk data yang berdekatan lebih sering memiliki kelas yang sama, sehingga mempengaruhi perhitungan membership (nilai derajat keanggotaan).

BAB V

KESIMPULAN

5.1 Kesimpulan

Dari hasil uji dan analisa yang telah dilakukan, dapat diambil kesimpulan diantaranya ialah:

1. Algoritma Fuzzy K-Nearest Neighbor (FK-NN) dapat diterapkan pada dataset evaluasi kinerja karyawan. Terdapat dua proses utama yang dilakukan yaitu proses pertama ialah proses normalisasi nilai atribut pada dataset dan proses selanjutnya yaitu mengklasifikasikan hasil normalisasi nilai atribut dengan algoritma fuzzy k-nearest neighbor yang dimana didalamnya terdapat proses yang menghasilkan nilai derajat keanggotaan. Dengan dipilihnya nilai derajat keanggotaan yang terbesar yang digunakan untuk menjadi penentu kelas target yang baru.
2. Tingkat akurasi pada metode Fuzzy K-NN dipengaruhi oleh nilai tetangga terdekat. Nilai tetangga terdekat yang terlalu tinggi menghasilkan akurasi yang baik dimana nilai akurasi terbesarnya mencapai 66% dengan jumlah komposisi kelas yang sama pada data latih, yang disebabkan oleh nilai membership yang digunakan untuk penentuan kelas target pada penelitian bernilai maksimum.

5.2 Saran

Saran untuk pengembangan penelitian lebih lanjut yang dapat diberikan oleh penulis adalah penentuan nilai k (tetangga terdekat) yang digunakan menjadi hal yang kritis sehingga perlu dilakukan analisis seperti *cross-validation* untuk memilih k (tetangga terdekat) yang terbaik agar memberikan akurasi yang tinggi.

DAFTAR PUSTAKA

- [BEZ-81] Bezdek, J.C. 1981. *Pattern Recognition with Fuzzy Objective Function Algorithm*. Plenum Press. New York.
- [CHU-98] Chung, Christina. 1998. *Applying Data Mining to Data Security*. <http://sirius.cs.ucdavis.edu/teaching/289F/>. Tanggal akses: 29 Maret 2012.
- [GIR-03] Giri, Yudho, S. 2003. *Data mining – Menggali Informasi yang Terpendam*. <http://ikc.cbn.net.id/populer/yudho/yudho-datamining.zip>. Tanggal akses: 28 Maret 2012.
- [HAM-08] Hamid Parvin, Hosein Alizadeh and Behrouz Minaei-Bidgoli. 2008. *MKNN: Modified K-Nearest Neighbor*. Proceedings of the World Congress on Engineering and Computer Science. San Francisco.
- [HAS-08] Hasibuan, Malayu S. P. 2008. *Manajemen Sumber Daya Manusia (Dasar, Dan Kunci Keberhasilan)*. PT. Bumi Aksara. Jakarta.
- [JAY-11] Jayalakshmi, T. dan Santhakumaran, A. 2011. *Statistical Normalization and Backpropagation for Classification (Online)*, (<http://www.ijcte.org/papers/288-L052.pdf>, diakses tanggal 20 Mei 2012).
- [KAN-03] Kantardzic, Mehmed. 2003. *Data Mining : Concepts, Models, Methods and Algorithm*. John Wiley & Sons. New York.
- [KEL-85] Keller, James M., Gray, Michael R., Givens, James A. 1985. *A Fuzzy K-Nearest Neighbor Algorithm*. University of Missouri, Columbia. USA.

- [KHU-07] Khusnawi. 2007. *Pengantar Solusi data Mining*. STMIK AMIKOM. Yogyakarta. <http://p3m.amikom.ac.id/p3m/56%20-%20PENGANTAR%20SOLUSI%20DATA%20MINING.pdf>.
Tanggal akses: 28 Maret 2012.
- [KUS-03] Kusumadewi, S dan Purnomo, H. 2003. *Artificial Intelligence (Teknik & Aplikasinya)*. Graha Ilmu. Jogjakarta.
- [KUS-10] Kusumadewi, S dan Purnomo, H. 2010. *Aplikasi Logika Fuzzy untuk Pendukung Keputusan: Jilid 2*. Graha Ilmu. Yogyakarta.
- [LAI-06] Lai, Kin Keung., Yu, Lean., Wang, Shouyang., and Zhou, Ligang. 2006. *Credit Risk Analysis Using a Reliability-Based Neural Network Ensemble Model*. University of Hong Kong. China.
- [MAN-00] Mangkunegara, Anwar Prabu. 2000. *Manajemen Sumber Daya Manusia*. Penerbit Remaja Rosdakarya. Bandung.
- [MAN-04] Mangkunegara, Anwar Prabu. 2004. *Manajemen Sumber Daya Manusia Perusahaan*. PT. Remaja Rosdakarya. Bandung.
- [MAN-05] Mangkunegara, Anwar Prabu. 2005. *Evaluasi Kinerja SDM*. Penerbit PT. Refika Aditama. Bandung.
- [MOR-09] Moradian, Mehdi dan Baarani, Ahmad. 2009. *KNNBA: k-Nearest-Neighbor-Based-Association Algorithm (Online)*, (<http://www.jatit.org/volumes/research-papers/Vol6No1/14Vol6No1.pdf>, diakses tanggal 20 Mei 2012).
- [NUG-06] Nugraha, Dany, dkk. 2006. *Diagnosis Gangguan Sistem Urinari pada Anjing dan Kucing Menggunakan VFI 5*. Institut Pertanian Bogor.

- [RAK-07] Rakhman, Juniar P., Ramadijanti, Nana., Satriyanto, Edi. 2007. *Translasi Bahasa Isyarat*. Jurusan Teknik Informatika, PENS-ITS. Surabaya.
- [SIM-05] Simanjuntak, Payaman J. 2005. *Manajemen Dan Evaluasi Kinerja*. Penerbit Lembaga Fakultas Ekonomi Universitas Indonesia. Jakarta.
- [THE-00] Thearling, Kurt. 2000. *Data Mining and Customer Relationship*. <http://www.thearling.com/index.htm>. Tanggal akses: 28 maret 2012.
- [WIB-11] Wibowo, Luqman Putra. 2011. *Rancang Bangun Human Resource Information System (HRIS) Sebagai Aplikasi Penilaian Kinerja Karyawan Perusahaan Asuransi*. Institut Teknologi Negeri Surabaya. Surabaya.
- [ZAI-99] Zaiane, Osmar R. 1999. *Principles of Knowledge Discovery in Databases*. University of Alberta. [http://www.cs.ualberta.ca/~zaiane/courses/cmput690/slides/Chapter 8](http://www.cs.ualberta.ca/~zaiane/courses/cmput690/slides/Chapter8). Tanggal akses: 28 Maret 2012.

LAMPIRAN I

Pengujian Data Latih dengan komposisi jumlah kelas yang sama

Hasil pengujian 6 macam data latih evaluasi kinerja karyawan dengan komposisi jumlah kelas yang sama:

k	Akurasi (%) pada					
	80 data latih	110 data latih	140 data latih	170 data latih	200 data latih	230 data latih
2	57	64	54	51	61	60
3	61	65	61	55	62	61
4	63	66	61	57	62	60
5	64	66	60	56	62	61
6	63	66	58	56	62	61
7	62	65	58	55	60	60
8	60	64	57	54	59	58
9	59	62	56	53	57	57
10	57	60	54	52	55	55
11	56	59	53	51	53	53
12	54	57	52	50	52	51
13	53	55	50	49	51	50
14	52	54	49	48	50	49
15	51	54	48	47	49	49
16	50	53	47	46	49	49
17	50	52	46	46	48	48
18	49	51	46	45	47	47
19	48	51	45	44	47	46
20	48	50	44	44	46	46
21	47	50	44	43	46	45
22	47	49	43	42	45	44
23	46	49	42	42	44	44
24	45	48	42	41	44	43
25	45	48	41	40	43	42
26	44	47	41	40	43	42
27	44	47	41	40	42	42
28	44	46	40	40	42	41
29	43	46	40	39	42	41
30	43	45	40	39	41	41
31	43	45	39	39	41	40
32	42	44	39	39	41	40
33	42	44	39	39	40	40
34	42	44	39	38	40	40
35	42	43	39	38	40	40

Pengaruh nilai k terhadap tingkat akurasi

