

BAB II

LANDASAN TEORI

2.1. Pengertian Baja

Baja adalah logam paduan dengan besi sebagai unsur dasar dan karbon sebagai unsur paduan utamanya. Kandungan karbon dalam baja berkisar antara 0.2% hingga 2.1% berat sesuai *grade*-nya. Fungsi karbon dalam baja adalah sebagai unsur penguat. Unsur paduan lain yang biasa ditambahkan selain karbon adalah mangan (*manganese*), krom (*chromium*), vanadium, dan nikel. Dengan memvariasikan kandungan karbon dan unsur paduan lainnya, berbagai jenis kualitas baja bisa didapatkan. Penambahan kandungan karbon pada baja dapat meningkatkan kekerasan (*hardness*) dan kekuatan tariknya (*tensile strength*), namun di sisi lain membuatnya menjadi getas (*brittle*) serta menurunkan keuletannya (*ductility*).

Pengaruh utama dari kandungan karbon dalam baja adalah pada kekuatan, kekerasan, dan sifat mudah dibentuk. Kandungan karbon yang besar dalam baja mengakibatkan meningkatnya kekerasan tetapi baja tersebut akan rapuh dan tidak mudah dibentuk (Davis, 1982).

2.2. Jenis Baja (secara garis besar).

Baja merupakan besi dengan kadar karbon kurang dari 2 %. Baja dapat dibentuk menjadi berbagai macam bentuk sesuai dengan keperluan. Secara garis besar ada 2 jenis baja, yaitu :

a. Baja Karbon

Baja karbon disebut juga plain karbon *steel*, mengandung terutama unsur karbon dan sedikit silikon, belerang dan fosfor. Berdasarkan kandungan karbonnya, baja karbon dibagi menjadi :

- Baja dengan kadar karbon rendah (< 0,2 % C)

Baja ini dengan komposisi karbon kurang dari 2%. Fasa dan struktur mikronya adalah ferrit dan perlit. Baja ini tidak bisa dikeraskan dengan cara perlakuan panas (*martensit*) hanya bisa dengan

pengerjaan dingin. Sifat mekaniknya lunak, lemah dan memiliki keuletan dan ketangguhan yang baik. Serta mampu mesin (*machinability*) dan mampu las nya (*weldability*) baik cocok untuk bahan bangunan konstruksi gedung, jembatan, rantai, body mobil.

- Baja dengan kadar karbon sedang (0,1%-0,5 % C)

Baja karbon sedang memiliki komposisi karbon antara 0,2%-0,5% C (berat). Dapat dikeraskan dengan perlakuan panas dengan cara memanaskan hingga fasa austenit dan setelah ditahan beberapa saat didinginkan dengan cepat ke dalam air atau sering disebut *quenching* untuk memperoleh fasa ang keras yaitu martensit. Baja ini terdiri dari baja karbon sedang biasa (*plain*) dan baja mampu keras. Kandungan karbon yang relatif tinggi itu dapat meningkatkan kekerasannya. Namun tidak cocok untuk di las, dengan kata lain mampu las nya rendah. Dengan penambahan unsur lain seperti Cr, Ni, dan Mo lebih meningkatkan mampu kerasnya. Baja ini lebih kuat dari baja karbon rendah dan cocok untuk komponen mesin, roda kereta api, roda gigi (gear), poros engkol (*crankshaft*) serta komponen struktur yang memerlukan kekuatan tinggi, ketahanan aus, dan tangguh.

- Baja dengan kadar karbon tinggi (>0,5 % C)

Baja karbon tinggi memiliki komposisi antara 0,6- 1,4% C (berat). Kekerasan dan kekuatannya sangat tinggi, namun keuletannya kurang. baja ini cocok untuk baja perkakas, *dies* (cetakan), pegas, kawat kekuatan tinggi dan alat potong yang dapat dikeraskan dan ditemper dengan baik. Baja ini terdiri dari baja karbon tinggi biasa dan baja perkakas. Khusus untuk baja perkakas biasanya mengandung Cr, V, W, dan Mo. Dalam pemuaduanya unsur-unsur tersebut bersenyawa dengan karbon menjadi senyawa yang sangat keras sehingga ketahanan aus sangat baik.

Kadar karbon yang terdapat di dalam baja akan mempengaruhi kuat tarik, kekerasan dan keuletan baja. Semakin tinggi kadar karbonnya, maka kuat tarik dan kekerasan baja semakin meningkat

tetapi keuletannya cenderung turun. Penggunaan baja di bidang teknik sipil pada umumnya berupa baja konstruksi atau baja profil, baja tulangan untuk beton dengan kadar karbon 0,10% - 0,50%. Selain itu baja karbon juga digunakan untuk baja/kawat pra tekan dengan kadar karbon s/d 0,90 %. Pada bidang teknik sipil sifat yang paling penting adalah kuat tarik dari baja itu sendiri.

b. Baja Paduan

Baja dikatakan di padu jika komposisi unsur-unsur paduannya secara khusus, bukan baja karbon biasa yang terdiri dari unsur fosfor dan mangan. Baja paduan semakin banyak di gunakan. Unsur yang paling banyak di gunakan untuk baja paduan, yaitu : Cr, Mn, Si, Ni, W, Mo, Ti, Al, Cu, Nb, Zr.

- Baja Paduan Rendah (*Low Alloy Steel*)

Baja paduan rendah merupakan baja paduan yang elemen paduannya kurang dari 2,5% wt, misalnya unsur Cr, Mn, Ni, S, Si, P dan lain-lain. Biasanya digunakan untuk membuat perkakas potong, gergaji, cetakan penarikan, pahat kayu, mata pisau, pemotong kikir, gurdi batu.

- Baja Paduan Menengah (*Medium Alloy Steel*)

Baja paduan menengah merupakan baja paduan yang elemen paduannya 2,5% - 10% wt, misalnya unsur Cr, Mn, Ni, S, Si, P dan lain-lain. Biasanya digunakan untuk membuat alat pengukur, cetakan penarikan, rol derat, mata gunting untuk plat tebal.

- Baja Paduan Tinggi (*High Alloy Steel*)

Baja paduan tinggi merupakan baja paduan yang elemen paduannya lebih dari 10% wt, misalnnnya unsur Cr, Mn, Ni, S, Si, P dan lain-lain (Amanto, 1999). Banyak digunakan untuk cetakan penarikan kawat, cetakan pengetrim, pengukur, rol derat.

Menurut (Amstead, 1993) secara umumnya, baja paduan memiliki sifat yang unggul daripada baja karbon biasa, diantaranya:

1. Keuletan yang tinggi tanpa pengurangan kekuatan tarik.
2. Tahan terhadap korosi dan keausan yang tergantung dari jenis paduannya.
3. Tahan terhadap perubahan suhu, ini berarti bahwa sifat fisiknya tidak banyak berubah.
4. Memiliki butiran halus dan homogen.

Pengaruh unsur-unsur paduan dalam baja adalah sebagai berikut:

1. Unsur karbon (C)

Karbon merupakan unsur terpenting yang dapat meningkatkan kekerasan dan kekuatan baja. Kandungan karbon di dalam baja sekitar 0,1%-1,7%, sedangkan unsur lainnya dibatasi sesuai dengan kegunaan baja. Unsur paduan yang bercampur di dalam lapisan baja adalah untuk membuat baja bereaksi terhadap pengerjaan panas dan menghasilkan sifat-sifat yang khusus. Karbon dalam baja dapat meningkatkan kekuatan dan kekerasan tetapi jika berlebihan akan menurunkan ketangguhan.

2. Unsur Mangan (Mn)

Semua baja mengandung mangan karena sangat dibutuhkan dalam proses pembuatan baja. Kandungan mangan kurang lebih 0,6% tidak mempengaruhi sifat baja, dengan kata lain mangan tidak memberikan pengaruh besar pada struktur baja dalam jumlah yang rendah. Penambahan unsur mangan dalam baja dapat menaikkan kuat tarik tanpa mengurangi atau sedikit mengurangi regangan, sehingga baja dengan penambahan mangan memiliki sifat kuat dan ulet.

3. Unsur Silikon (Si)

Silikon merupakan unsur paduan yang ada pada setiap baja dengan kandungan lebih dari 0,4% yang mempunyai pengaruh untuk menaikkan tegangan tarik dan menurunkan laju pendinginan kritis. Silikon dalam baja dapat meningkatkan kekuatan, kekerasan, kekenyalan, ketahanan aus, dan ketahanan terhadap panas dan karat. Unsur silikon

menyebabkan sementit tidak stabil, sehingga memisahkan dan membentuk grafit. Unsur silikon juga merupakan pembentuk ferit, tetapi bukan pembentuk karbida, silikon juga cenderung membentuk partikel oksida sehingga memperbanyak pengintian kristal dan mengurangi pertumbuhan akibatnya struktur butir semakin halus.

4. Unsur Nikel (Ni)

Nikel mempunyai pengaruh yang sama seperti mangan, yaitu memperbaiki kekuatan tarik dan menaikkan sifat ulet, tahan panas, jika pada baja paduan terdapat unsur nikel sekitar 25% maka baja dapat tahan terhadap korosi. Unsur nikel yang bertindak sebagai tahan karat (korosi) disebabkan nikel bertindak sebagai lapisan penghalang yang melindungi permukaan baja.

5. Unsur Kromium (Cr)

Sifat unsur kromium dapat menurunkan laju pendinginan kritis (kromium sejumlah 1,5% cukup meningkatkan kekerasan dalam minyak). Penambahan kromium pada baja menghasilkan struktur yang lebih halus dan membuat sifat baja dikeraskan lebih baik karena kromium dan karbon dapat membentuk karbida. Kromium dapat menambah kekuatan tarik dan keplastisan serta berguna juga dalam membentuk lapisan pasif untuk melindungi baja dari korosi serta tahan terhadap suhu tinggi.

2.3. Definisi Sistem

Sistem secara fisik adalah kumpulan dari elemen-elemen yang beroperasi bersama-sama untuk menyelesaikan suatu sasaran (Gordon, 1991). Menurut (Jogianto, 2005) Sistem adalah kumpulan dari elemen-elemen yang berinteraksi untuk mencapai suatu tujuan tertentu.

2.4. Klasifikasi

Klasifikasi adalah proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan

untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui. Dalam mencapai tujuan tersebut, proses klasifikasi membentuk suatu model yang mampu membedakan data kedalam kelas-kelas yang berbeda berdasarkan aturan atau fungsi tertentu. Model itu sendiri bisa berupa aturan "Jika-maka", berupa pohon keputusan, atau formula matematis.

Tahapan dari klasifikasi dalam data mining menurut (Han dan Kamber, 2006) terdiri dari :

- **Pembangunan Model**

Pada tahapan ini dibuat sebuah model untuk menyelesaikan masalah klasifikasi class atau atribut dalam data. Tahap ini merupakan fase pelatihan, dimana data latih dianalisis menggunakan algoritma klasifikasi, sehingga model pembelajaran direpresentasikan dalam bentuk aturan klasifikasi.

- **Penerapan Model**

Pada tahapan ini model yang sudah dibangun sebelumnya digunakan untuk menentukan atribut/class dari sebuah data baru yang atribut/classnya belum diketahui sebelumnya. Tahap ini digunakan untuk memperkirakan keakuratan aturan klasifikasi terhadap data uji. Jika model dapat diterima, maka aturan dapat diterapkan terhadap klasifikasi data baru.

2.5. Data Mining

Data mining adalah serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data. Informasi yang dihasilkan diperoleh dengan cara mengekstraksi dan mengenali pola yang penting atau menarik dari data yang terdapat dalam basis data. Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi

dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar.

Menurut Gartner Group data mining adalah suatu proses menemukan hubungan yang berarti, pola, dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika. Data mining bukanlah suatu bidang yang sama sekali baru. Salah satu kesulitan untuk mendefinisikan data mining adalah kenyataan bahwa data mining mewarisi banyak aspek dan teknik dari bidang-bidang ilmu yang sudah mapan terlebih dulu. Berawal dari beberapa disiplin ilmu, data mining bertujuan untuk memperbaiki teknik tradisional sehingga bisa menangani:

- Jumlah data yang sangat besar
- Dimensi data yang tinggi
- Data yang heterogen dan berbeda sifat

Menurut para ahli, data mining merupakan sebuah analisa dari observasi data dalam jumlah besar untuk menemukan hubungan yang tidak diketahui sebelumnya dan metode baru untuk meringkas data agar mudah dipahami serta kegunaannya untuk pemilik data.

Data-data yang ada, tidak dapat langsung diolah dengan menggunakan sistem data mining. Data-data tersebut harus dipersiapkan terlebih dahulu agar hasil yang diperoleh dapat lebih maksimal, dan waktu komputasinya lebih minimal. Proses persiapan data ini sendiri dapat mencapai 60 % dari keseluruhan proses dalam data mining. Karena itu *data mining* sebenarnya memiliki akar yang panjang dari bidang ilmu seperti kecerdasan buatan (*artificial intelligent*), *machine learning*, statistik dan database. Beberapa metode yang sering disebut-sebut dalam literatur *data mining* antara lain *clustering*, *yesification*, *association rules mining*, *neural network*, *genetic algorithm* dan lain-lain (Pramudiono, 2007).

Sebagai suatu rangkaian proses, *data mining* dapat dibagi menjadi beberapa tahap yang diilustrasikan di Gambar, Tahap-tahap tersebut bersifat interaktif, pemakai terlibat langsung dengan perantaraan *knowledge base*.

2.5.1. Tahap-tahap Data Mining

Tahap-tahap *data mining* ada 6 yaitu :

1. Pembersihan data (*data cleaning*).

Pembersihan data merupakan proses menghilangkan *noise* dan data yang tidak konsisten atau data yang tidak relevan. Pada umumnya data diperoleh, baik dari *database* suatu perusahaan maupun hasil eksperimen, memiliki isian-isian yang tidak sempurna seperti data yang hilang, data yang tidak valid atau juga hanya sekedar salah ketik. Selain itu, ada juga atribut-atribut data yang tidak relevan dengan hipotesa *data mining* yang dimiliki. Data-data yang tidak relevan itu juga lebih baik dibuang. Pembersihan data juga akan mempengaruhi performansi dari teknik *data mining* karena data yang ditangani akan berkurang jumlah dan kompleksitasnya.

2. Integrasi Data (*data integraton*)

Integrasi data merupakan penggabungan data dari berbagai *database* ke dalam satu *database* baru. Tidak jarang data yang diperlukan untuk *data mining* tidak hanya berasal dari satu *database* atau file teks. Integrasi data dilakukan pada atribut-atribut yang mengidentifikasi entitas-entitas yang unik seperti atribut nama, jenis produk, nomor pelanggan dan lainnya. Integrasi data perlu dilakukan secara cermat karena kesalahan pada integrasi data bisa menghasilkan hasil yang menyimpang dan bahkan menyesatkan pengambilan aksi nantinya. Sebagai contoh bila integrasi data berdasarkan jenis produk dari kategori yang berbeda maka akan didapatkan korelasi antar produk yang sebenarnya tidak ada.

3. Seleksi Data (*Data Selection*)

Data yang ada pada *database* sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari *database*. Sebagai contoh, sebuah kasus yang meneliti faktor kecenderungan orang membeli dalam kasus *market analysis*, tidak perlu mengambil nama pelanggan, cukup dengan id pelanggan saja.

4. Transformasi Data (*data transformation*)

Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam *data mining*. Beberapa metode *data mining* membutuhkan format data yang khusus sebelum bisa diaplikasikan. Sebagai contoh beberapa metode standar seperti analisis asosiasi dan *clustering* hanya bisa menerima input data kategorikal, Karenanya data berupa angka numerik yang berlanjut perlu dibagi-bagi menjadi beberapa interval. Proses ini sering disebut transformasi data.

5. Proses *Mining*

Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data.

6. Evaluasi Pola (*pattern evaluation*).

Untuk mengidentifikasi pola-pola menarik kedalam *knowledge based* yang ditemukan. Dalam tahap ini hasil dari teknik *data mining* berupa pola-pola yang khas maupun model prediksi dievaluasi untuk menilai apakah hipotesa yang ada tercapai. Bila ternyata hasil yang diperoleh tidak sesuai hipotesa ada beberapa alternatif yang dapat diambil seperti menjadikannya umpan balik untuk memperbaiki proses *data mining*, mencoba metode *data mining* yang lebih sesuai, atau menerima hasil ini sebagai suatu hal yang diluar dugaan yang mungkin bermanfaat.

2.5.2. Pengelompokan Data Mining

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu (Daniel T. Larose, 2005):

1. Deskripsi

Deskripsi adalah menggambarkan pola dan kecenderungan yang terdapat dalam data secara sederhana. Deskripsi dari pola dan kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.

2. Klasifikasi

Suatu teknik dengan melihat pada kelakuan dan atribut dari kelompok yang telah didefinisikan. Teknik ini dapat memberikan klasifikasi pada data baru dengan memanipulasi data yang telah diklasifikasi dan dengan menggunakan hasilnya untuk memberikan sejumlah aturan. Klasifikasi menggunakan *supervised learning*.

3. Estimasi

Estimasi hampir sama dengan klasifikasi, perbedaannya adalah variabel target estimasi lebih ke arah numerik daripada ke arah kategori. Model dibangun dengan menggunakan *record* lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi.

4. Prediksi

Prediksi memiliki kesamaan dengan klasifikasi dan estimasi, perbedaannya adalah hasil dari prediksi akan ada dimasa mendatang. Beberapa teknik yang digunakan dalam klasifikasi dan estimasi dapat juga digunakan (untuk keadaan yang tepat) untuk prediksi.

5. Klastering

Klastering merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan dengan *record-record* dalam kluster lain. Klastering menggunakan *unsupervised learning*.

6. Asosiasi

Tugas asosiasi atau sering disebut juga sebagai *market basket analysis* dalam data mining adalah menemukan relasi atau korelasi diantara himpunan item-item dan menemukan atribut yang muncul

dalam satu waktu. Asosiasi menggunakan *unsupervised learning*. Penting tidaknya suatu aturan asosiatif dapat diketahui dengan dua parameter, *support* dan *confidence*.

2.6. *K-Nearest Neighbor*

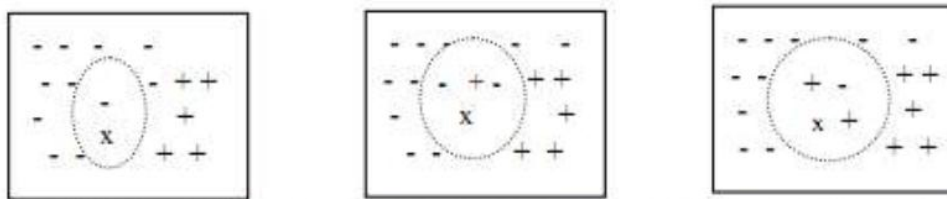
K-Nearest Neighbor sangat sering digunakan dalam klasifikasi dengan tujuan dari algoritma ini adalah untuk mengklasifikasi objek baru berdasarkan atribut dan *training samples* (Larose, 2002). Algoritma *K-nearest neighbor* (KNN) adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Teknik ini sangat sederhana dan mudah diimplementasikan. Tujuan dari algoritma ini adalah mengklasifikasikan objek baru berdasarkan atribut dan training sample.

KNN memiliki beberapa kelebihan yaitu ketagguhan terhadap training data yang memiliki banyak noise dan efektif apabila training datanya besar. Sedangkan kelemahan KNN adalah KNN perlunya menentukan nilai parameter K (jumlah dari tetangga terdekat), training berdasarkan jarak tidak jelas mengenai jenis jarak apa yang harus digunakan untuk mendapatkan hasil terbaik, dan biaya komputasi sangat tinggi karena diperlukan perhitungan jarak dari tiap *query instance* pada keseluruhan *training sample*.

Ketepatan algoritma KNN sangat dipengaruhi oleh ada atau tidak adanya fitur-fitur yang tidak relevan atau jika bobot fitur tersebut setara dengan relevansinya terhadap klasifikasi. Riset terhadap algoritma ini sebagian besar membahas bagaimana memilih dan memberi bobot terhadap fitur agar performa klasifikasi menjadi lebih baik.

Classifier tidak menggunakan model apapun untuk dicocokkan dan hanya berdasarkan pada memori. Diberikan titik *query*, akan ditemukan sejumlah K obyek atau titik training yang paling dekat dengan titik *query*. Klasifikasi menggunakan *voting* terbanyak diantara klasifikasi dari K obyek algoritma KNN menggunakan klasifikasi ketetanggaan sebagai nilai prediksi dari *query instance* yang baru. Mirip dengan teknik clustering,

pengelompokkan suatu data baru berdasarkan jarak data baru itu ke beberapa data/tetangga (*neighbor*) terdekat. Dalam hal ini jumlah data/tetangga terdekat ditentukan oleh *user* yang dinyatakan dengan *K*. Misalkan ditentukan $K=5$, maka setiap data testing dihitung jaraknya terhadap data training dan dipilih 5 data *training* yang jaraknya paling dekat ke data *testing*. Lalu periksa *output* atau labelnya masing-masing, kemudian tentukan *output* mana yang frekuensinya paling banyak. Lalu masukkan suatu data *testing* ke kelompok dengan *output* paling banyak. Misalkan dalam kasus klasifikasi dengan 3 kelas, lima data tadi terbagi atas tiga data dengan *output* kelas 1, satu data dengan *output* kelas 2 dan satu data dengan *output* kelas 3, maka dapat disimpulkan bahwa *output* dengan label kelas 1 adalah yang paling banyak. Maka data baru tadi dapat dikelompokkan ke dalam kelas 1. Prosedur ini dilakukan untuk semua data *testing* (Santosa, 2007). Gambar berikut ini adalah bentuk representasi K-NN dengan 1, 2 dan 3 tetangga data terhadap data baru *x*.



a) 1-nearest neighbor

b) 2-nearest neighbor

c) 3-nearest neighbor

Gambar 2.1 Ilustrasi 1-, 2-, 3-nearest neighbor terhadap data baru (*x*)

Untuk mendefinisikan jarak antara dua titik yaitu titik pada data training (*x*) dan titik pada data testing (*y*) maka digunakan rumus *Euclidean*:

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

Dengan:

x_1 = sampel data

x_2 = data uji

i = variabel data

d = jarak

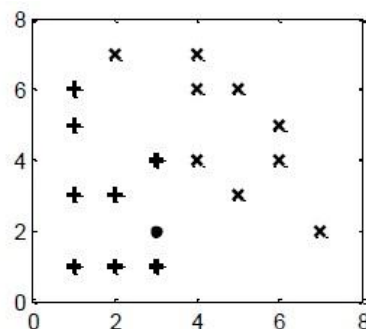
p = dimensi data

Langkah-langkah untuk menghitung metode *K-Nearest Neighbor* :

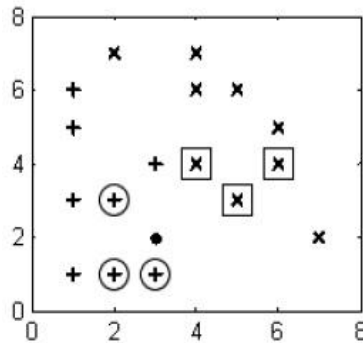
1. Menentukan parameter K (jumlah tetangga paling dekat).
2. Menghitung kuadrat jarak *Euclid* (*query instance*) masing-masing obyek terhadap data sampel yang diberikan.
3. Kemudian mengurutkan objek-objek tersebut kedalam kelompok yang mempunyai jarak *Euclid* terkecil.
4. Mengumpulkan kategori Y (Klasifikasi *nearest neighbor*).
5. Dengan menggunakan kategori *nearest neighbor* yang paling mayoritas maka dapat dipredisikan nilai *query instance* yang telah dihitung.

2.7. *Fuzzy K-Nearest Neighbor in Every Class*

Metode Fuzzy *K-Nearest Neighbor in Every Class* (FK-NNC) menggunakan sejumlah K tetangga terdekat pada setiap kelas dari sebuah data uji, bukan K tetangga terdekat seperti pada *K-NN* dan *FK-NN*. Gambar 1 memberikan gambaran K tetangga terdekat dari setiap kelas pada sebuah data uji pada metode *FK-NNC*. Pada gambar tersebut, untuk $K=3$, tiga tetangga terdekat dikelas $+$ dan tiga tetangga dikelas x yang ditemukan oleh *FK-NNC*.



a. Tanda dot hitam (solid) adalah data uji



b. Tiga tetangga dikelas + dan tiga tetangga dikelas x

Gambar 2.2 Konsep K tetangga terdekat dari setiap kelas dalam FK-NNC untuk K=3

Kerangka kerja FK-NNC menggunakan FK-NN sebagai basis kerangka kerja, dimana sebuah data uji mempunyai nilai keanggotaan pada setiap kelas dalam interval $[0,1]$. Jumlah nilai keanggotaan sebuah data pada semua kelas sama dengan 1, seperti pada persamaan (3).

$$\sum_{j=1}^c u_{ij} = 1, 0 \leq u_{ij} \leq 1 \quad (2.4)$$

Dimana u_{ij} adalah nilai keanggotaan data uji x_i ke kelas ke- j . Setiap data uji x_i , harus dicarikan K tetangga terdekat pada setiap kelas menggunakan formula (2), sehingga untuk 2 kelas akan ada $2 \times K$ tetangga yang didapat, untuk 3 kelas akan ada $3 \times K$ tetangga yang didapat, begitu seterusnya. Selanjutnya jarak data uji x_i ke semua K tetangga dari setiap kelas ke- j dijumlahkan, formula yang digunakan:

$$S_{ij} = \sum_{r=1}^K d(x_i, x_r)^{\frac{-2}{m-1}} \quad (2.5)$$

Nilai d sebagai akumulasi jarak data uji x_i ke K tetangga dalam kelas ke- j dilakukan sebanyak C kelas. Selanjutnya, akumulasi jarak data uji x_i ke setiap kelas digabungkan, disimbolkan D , formula yang digunakan:

$$D_i = \sum_{j=1}^c (S_{ij}) \quad (2.6)$$

Nilai m disini merupakan pangkat bobot (*weight exponent*) seperti pada FK-NN, nilai $m > 1$. Untuk mendapatkan nilai keanggotaan data uji x_i pada setiap kelas ke- j (ada C kelas), menggunakan formula:

$$u_{ij} = \frac{S_{ij}}{D_i} \quad (2.7)$$

Untuk menentukan kelas hasil prediksi data uji x_i , dipilih kelas dengan nilai keanggotaan terbesar dari data x_i . Formula yang digunakan:

$$y' = \underset{j = 1}{\overset{C}{arg\ max}} (u_{ij}) \quad (2.8)$$

Algoritma FK-NNC untuk sebuah data uji, disajikan sebagai berikut:

1. Cari K tetangga terdekat pada setiap kelas.
2. Hitung S sebagai akumulasi jarak K tetangga pada setiap kelas.
3. Hitung D sebagai akumulasi semua jarak dari $C \times K$ tetangga.
4. Hitung u sebagai nilai keanggotaan data pada setiap kelas.
5. Pilih nilai keanggotaan terbesar kelas dengan nilai keanggotaan terbesar menjadi kelas hasil prediksi untuk data uji tersebut.

Berikut adalah contoh perhitungan dari FK-NNC terhadap data set buatan sebanyak 10 data sebagai data latih sedangkan data uji adalah $X = 3$ dan $Y = 4$ dengan $K = 3$, sedangkan perhitungan jarak yang digunakan adalah *Euclidean* dan bobot $m = 2$.

Tabel 2.1 Tabel data set buatan

No.	Variabel X	Variabel Y	Kelas Asli
1	2	3	0
2	3	2	0
3	1	3	0
4	1	5	0
5	1	6	0
6	4	4	1
7	5	3	1

8	6	6	1
9	5	6	1
10	6	4	1

Langkah pertama adalah mencari jarak setiap data latih dengan data uji dengan perhitungan *euclidean*, kemudian memilih data dengan jarak terdekat sebanyak $K=3$ tiap masing-masing kelas.

Berikut perhitungan jarak *euclidean* data ke-1 dengan data uji $X = 3$ dan $Y = 4$:

$$d_1 = \sqrt{(3 - 2)^2 + (4 - 3)^2} = \sqrt{2} = 1.414$$

Tabel 2.2 Tabel hasil perhitungan jarak *euclidean*

No.	Variabel X	Variabel Y	Kelas Asli	d
1	2	3	0	1.414214
2	3	2	0	2
3	1	3	0	2.236068
4	1	5	0	2.236068
5	1	6	0	2.828427
6	4	4	1	1
7	5	3	1	2.236068
8	6	6	1	3.605551
9	5	6	1	2.828427
10	6	4	1	3

Setelah melakukan perhitungan jarak *euclidean* maka didapatkan 3 data yang terdekat untuk kelas 0 yakni data nomor 1, 2 dan 3 sedangkan 3 data yang terdekat kelas 1 yakni nomor 6, 7 dan 9. Kemudian dilakukan perhitungan jarak S seperti rumus 2.5 dengan $m=2$ dan dihitung akumulasi jarak keseluruhan data D seperti rumus 2.6 sebagai berikut :

$$\begin{aligned}
 S_0 &= (d_1^{-2}) + (d_2^{-2}) + (d_3^{-2}) \\
 &= (1.414^{-2}) + (2^{-2}) + (2.236^{-2}) \\
 &= 0.5 + 0.25 + 0.2 = 0.95
 \end{aligned}$$

$$\begin{aligned}
 S_1 &= (d_6^{-2}) + (d_7^{-2}) + (d_9^{-2}) \\
 &= (1^{-2}) + (2.236^{-2}) + (2.828^{-2}) \\
 &= 1 + 0.2 + 0.125 = 0.95 = 1.325
 \end{aligned}$$

$$\begin{aligned}
 D &= S_0 + S_1 \\
 &= 0.95 + 1.325 \\
 &= 2.275
 \end{aligned}$$

Tabel 2.3 Tabel hasil perhitungan jarak S tiap kelas dan D akumulasi jarak keseluruhan kelas

No.	Variabel X	Variabel Y	Kelas Asli	D	d^{-2}
1	2	3	0	1.414214	0.5
2	3	2	0	2	0.25
3	1	3	0	2.236068	0.2
6	4	4	1	1	1
7	5	3	1	2.236068	0.2
9	5	6	1	2.828427	0.125
S_0					0.95
S_1					1.325
D					2.275

Langkah selanjutnya adalah menghitung nilai keanggotaan tiap kelas dengan rumus 2.7 sebagai berikut.

. Nilai keanggotaan kelas 0 adalah :

$$\begin{aligned}
 u_0 &= \frac{S_0}{D} \\
 &= \frac{0.95}{2.275} = 0.417
 \end{aligned}$$

Nilai keanggotaan kelas 1 adalah :

$$\begin{aligned} u_1 &= \frac{S_1}{D} \\ &= \frac{1.325}{2.275} = 0.583 \end{aligned}$$

Untuk tahap terakhir yakni penentuan kelas dengan memilih nilai keanggotaan terbesar. Hasil yang didapat dari perhitungan nilai keanggotaan masing-masing kelas untuk data uji $X = 3$ dan $Y = 4$ menunjukkan bahwa data uji termasuk kelas 1 dengan nilai keanggotaan 0.583.

2.8. Penelitian Sebelumnya

Mohammad Sholikhuddin. 2015, Universitas Muhammadiyah Gresik. dengan judul “*Sistem Aplikasi Diagnosa Jenis Baja Berdasarkan Komposisi Kimia Dengan Menggunakan Metode FK-NN (Fuzzy K-Nearest Neighbor)*”. Dengan menggunakan Metode Fuzzy KNN Sistem dapat memprediksi jenis baja yang baik sesuai kebutuhan karena selama ini hanya mementingkan pemenuhan order tanpa memperhitungkan bahan bakunya, yaitu baja. Dalam penelitian yang dilakukan proses pengujiannya menggunakan 100 data yang terdiri dari 70 data latih dan 30 data uji. Hasil dari pengujian sistem tersebut memakai 2 kali percobaan dengan menggunakan data latih yang berbeda didapatkan tingkat akurasi data rata-rata sebesar 95 % dan laju error rata-rata 5 %. memiliki nilai akurasi kebenaran sistem sebesar terbaik 96.7 % yang didapatkan dari 1-NN 4-NN dan 7-NN.

Eko Prasetyo(2012) dalam jurnal di Seminar Nasional Teknik Informatika (Santika) menggunakan metode Fuzzy K-Nearest Neighbor In Every Class (FK-NNC) untuk klasifikasi data. Metode FK-NNC menggunakan K tetangga terdekat pada setiap kelas dari sebuah data uji, bukan dari K tetangga terdekat seperti pada K-NN atau FK-NN. Setiap data uji x_i harus dicarikan K tetangga terdekat pada setiap kelas. Penelitian ini mengusulkan Fuzzy K-NN in every Class (FK-NNC) untuk memperbaiki akurasi kinerja pada saat prediksi. Skenario pengujian dilakukan dengan tiga

metode yakni K-NN, F-KNN dan FK-NNC pada dataset Vertebral Coloumn dan dataset Iris. Hasil prediksi untuk dataset Iris pada ketiga metode bisa dikatakan hampir sama untuk K bernilai 3, 5, 7, dan 9 karena perbedaan hasil akurasi yang tipis. Sedangkan untuk pengujian dataset Vertebral Column, untuk FK-NNC memberikan hasil akurasi yang lebih tinggi daripada K-NN dan FK-NN. Akurasi yang diberikan oleh FK-NNC terbukti lebih tinggi dari pada K-NN atau FK-NN sehingga hasil prediksi yang didapatkan lebih dapat dipertanggungjawabkan. FK-NNC dapat menjadi alternatif metode K-NN, FK-NN dan varian-varian yang lain untuk melakukan pekerjaan klasifikasi data.

Najmatun Nabila. 2013, Universitas Muhammadiyah Gresik. Penelitiannya yang berjudul "*Pengklasifikasian Jenis Hadits Dengan Menggunakan Metode Fuzzy K-Nearest Neighbor In Every Class*". Masyarakat umum seringkali ikut-ikutan dalam mengambil hadits sebagai dasar hukum tanpa diketahui kebenaran dari hadits tersebut. Masyarakat lebih muda mengetahui hadits shahih dibandingkan hadits yang *dha'if* (lemah), dikarenakan banyak kitab-kitab yang membukukan hadits-hadits shohih. Sedangkan untuk hadits-hadits *dha'if* sendiri masih sulit dibedakan, terutama bagi masyarakat awam karena masih belum banyak buku yang menjelaskan mengenai contoh-contoh hadits *dha'if*. dalam penelitiannya, Nabila membandingkan hadits-hadits dari berbagai sumber periwayatan untuk kemudian diteliti kredibilitasnya dengan mempelajari atau mengkaji ilmu-ilmu hadits (*'ulum al-hadits*) yang membahan analisis atau pengklasifikasian hadits berdasarkan kriteria-kriteria tertentu. Dengan menggunakan metode Fuzzy K-Nearest Neighbor in Every Class yang dapat digunakan untuk membedakan jenis hadits tidak *dha'if* dan hadits *dha'if* dengan nilai rata-rata Accuracy 88,54% dan nilai eror 11,46%.