

BAB II

LANDASAN TEORI

Pada bab ini berisi tentang teori-teori yang digunakan dalam penyelesaian permasalahan pada penelitian khususnya pada perancangan dan implementasi sistemnya.

2.1 Pengertian Baja

Baja adalah logam paduan, logam besi sebagai unsur dasar dengan karbon sebagai unsur paduan utamanya. Kandungan unsur karbon dalam baja berkisar antara 0.2% hingga 2.1% berat sesuai gradenya. Fungsi karbon dalam baja adalah sebagai unsur penguat dengan mencegah dislokasi bergeser pada kisi kristal (*crystal lattice*) atom besi. Baja karbon ini dikenal sebagai baja hitam karena berwarna hitam, banyak digunakan untuk peralatan pertanian misalnya sabit dan cangkul. Unsur paduan lain yang biasa ditambahkan selain karbon adalah *titanium*, krom (*chromium*), *nikel*, *vanadium*, *cobalt* dan tungsten (*wolfram*). 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*) (Amstead, 1993).

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, misalnya 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.

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 merupakan bidang dari beberapa bidang keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, database, dan visualisasi untuk penanganan permasalahan pengambilan informasi dari database yang besar (Larose, Daniel T. 2005).

2.5.1 Pengelompokan Data Mining

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

Metode yang akan digunakan pada penelitian ini termasuk kedalam kelompok prediksi, karena menggunakan teknik klasifikasi yang hasilnya akan ada dimasa mendatang.

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.

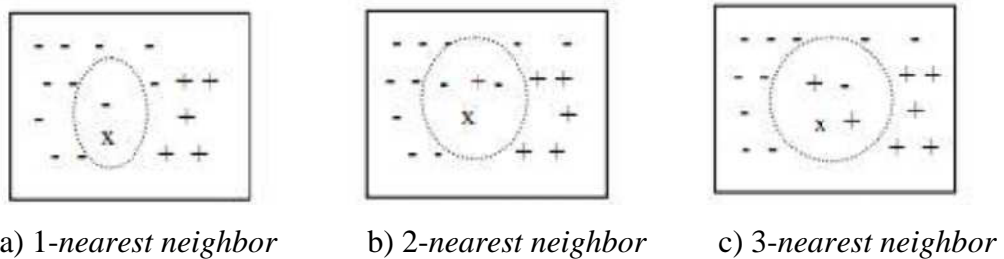
KNN memiliki beberapa kelebihan yaitu ketagguhan terhadap training data yang memiliki banyak noise dan efektif apabila training datanya besar. Sedangkan kelemahan KNN adalah KNN perlunya menentukan nilai parameter K (jumlah dari tetangga terdekat), training berdasarkan jarak tidak jelas mengenai jenis 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 .



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*:

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

Dengan:

x_1 = sampel data

x_2 = data uji

i = variabel data

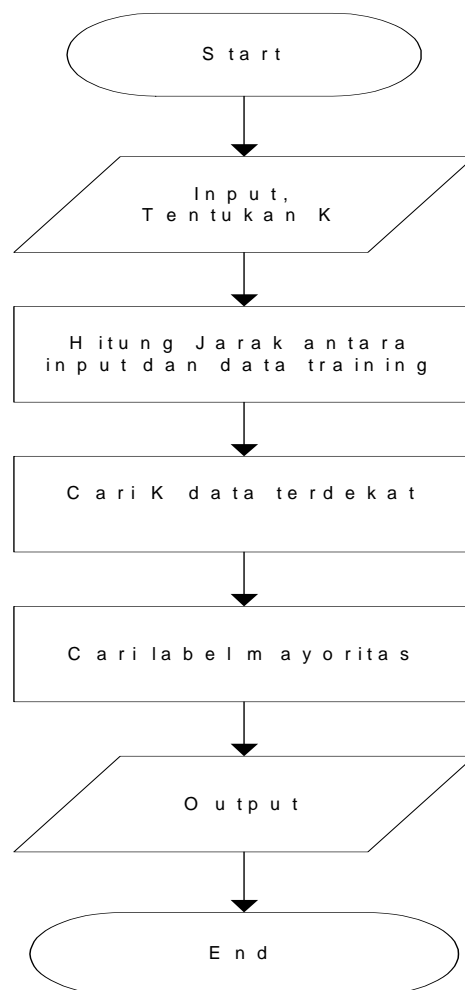
d = jarak

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.

2.6.1 Flowchart Perhitungan *K- Nearest Neighbor*



Gambar 2.2 Flowchart *K-Nearest Neighbor*

dengan d adalah jarak antara titik pada data training x dan titik data testing y yang akan diklasifikasi, dimana $x=x_1,x_2,\dots,x_i$ dan $y=y_1,y_2,\dots,y_i$ dan I merepresentasikan nilai atribut serta n merupakan dimensi atribut (Han & Kamber, 2001). Sebagai ilustrasi, pada Tabel 1 berikut ini disajikan contoh penerapan rumus *Euclidean*, pada empat data klasifikasi kualitas baik dan tidak baik sebuah kertas tisu yang dinilai berdasarkan daya tahan kertas tersebut dan fungsinya. Sebanyak tiga data yang sudah terklasifikasi yaitu data no 1,2, dan 3. masing-masing data dihitung jaraknya ke data no 4 untuk mendapatkan kelas yang sesuai bagi data no 4 maka $k=1$.

Tabel 2.1 klasifikasi kualitas baik atau tidak baik sebuah kertas tisu

No	Fungsi	Daya Tahan	Klasifikasi
1	7	7	Tidak Baik
2	7	4	Tidak Baik
3	3	4	Baik
4	1	4	?

Berikut ini disajikan pula perhitungan yang dilakukan terhadap tiga data yang sudah terklasifikasi dengan data yang belum terklasifikasi pada Tabel di atas. Jarak data no satu ke data no empat:

$$d_{1,4} = \sqrt{(7 - 1)^2 + (7 - 4)^2} = \sqrt{6^2 + 3^2} = \sqrt{45} = 6.07$$

Jarak data no dua ke data no empat:

$$d_{2,4} = \sqrt{(7 - 1)^2 + (4 - 4)^2} = \sqrt{6^2 + 0^2} = \sqrt{36} = 6$$

Jarak data no tiga ke data no empat:

$$d_{3,4} = \sqrt{(3 - 1)^2 + (4 - 4)^2} = \sqrt{2^2 + 0^2} = \sqrt{4} = 2$$

Dari hasil perhitungan di atas diperoleh jarak antara data no tiga dan data no empat adalah jarak yang terdekat maka kelas data no empat adalah baik. Teknik ini akan diuji cobakan terhadap dataset yang belum terklasifikasi atau data yang belum dikenal, untuk menemukan kelas yang sesuai dengan berdasarkan pada data tetangga terdekatnya yang sudah terklasifikasi. Tingkat ketepatan klasifikasi

terhadap data dari kedua algoritma yang digunakan menjadi titik fokus analisa dalam penelitian.

2.7 Fuzzy K-Nearest Neighbor

Selain K-NN yang melakukan prediksi secara tegas pada uji berdasarkan perbandingan K tetangga terdekat, ada pendekatan lain yang dalam melakukan prediksi juga berdasarkan K tetangga terdekat, tetapi tidak secara tegas memprediksi kelas yang harus diikuti oleh data uji, pemberian label kelas data uji pada setiap kelas dengan memberikan nilai keanggotaan seperti halnya teori himpunan *fuzzy*. Algoritma *Fuzzy K-Nearest Neighbor* (FK-NN) diperkenalkan oleh Keller *et al* (1985) dengan mengembangkan K-NN yang digabungkan dengan teori *fuzzy* dalam menyampaikan definisi pemberian label kelas pada data uji yang diprediksi.

Seperti halnya pada teori *fuzzy*, sebuah data mempunyai nilai keanggotaan pada setiap kelas, yang artinya sebuah data bisa dimiliki oleh 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 (Liao, 2007):

$$u(x, c_i) = \frac{\sum_{k=1}^K u(x_k, c_i) * d(x, x_k)^{\frac{-2}{(m-1)}}}{\sum_{k=1}^K d(x, x_k)^{\frac{-2}{(m-1)}}} \dots\dots\dots(2.2)$$

Dengan:

$u(x, c_i)$ = Nilai keanggotaan data x ke kelas c_i .

K = Jumlah tetangga yang digunakan.

$u(x_k, c_i)$ = Nilai keanggotaan data tetangga dalam K tetangga pada kelas c_i , nilainya 1 jika data latih x_k milik kelas c_i atau 0 jika bukan milik kelas c_i .

$d(x, x_k)$ = Jarak dari data x ke data x_k dalam K tetangga terdekat.

m = Bobot pangkat yang besarnya >1 .

Nilai keanggotaan suatu data pada kelas sangat dipengaruhi oleh jarak data itu ke tetangga terdekatnya. Semakin dekat ke tetangganya, semakin besar nilai

keanggotaan data tersebut. Jarak tersebut diukur dengan N dimensi (fitur) data. Pengukuran jarak (ketidak miripan) dua data yang digunakan dalam FK-NN digeneralisasi dengan formula (Liao, 2007) :

$$d(x_i, x_j) = \left(\sum_{l=1}^N |x_{il} - x_{jl}|^p \right)^{\frac{1}{p}} \dots \dots \dots (2.3)$$

Dimana:

N = Dimensi (jumlah fitur) data.

p = Penentu jarak yang digunakan. Jika p=1, jarak yang digunakan adalah Manhattan. Jika p=2, jarak yang digunakan adalah Euclidian. Namun jika p= ∞, jarak yang digunakan adalah Chebyshev.

Meskipun FK-NN menggunakan nilai keanggotaan untuk menyatakan keanggotaan data pada setiap kelas, tetapi untuk memberikan keluaran akhir, FK-NN tetap harus memberikan kelas akhir hasil prediksi, untuk keperluan ini, FKNN memilih kelas dengan nilai keanggotaan terbesar pada data tersebut. Algoritma Prediksi dengan FK-NN (Prasetyo, E. 2012)

1. Normalisasikan data menggunakan nilai terbesar dan terkecil data pada setiap fitur.
2. Cari K tetangga terdekat untuk data uji x menggunakan persamaan (2.2).
3. Hitung nilai keanggotaan $u(x, y_i)$ menggunakan persamaan (2.3) untuk setiap i, dimana $1 \leq i \leq C$.
4. Ambil nilai terbesar $c = u(x, y_i)$ untuk semua $1 \leq i \leq C$, C adalah jumlah kelas.
5. Berikan label kelas c ke data uji x.

2.8 Penelitian Sebelumnya

Sebagai bahan perbandingan, pada penelitian sebelumnya yang dilakukan oleh:

Puspasari, M., Dewi, C., dan Rahman, M.A. 2013, Universitas Brawijaya. Penelitiannya yang berjudul : “*Prediksi Tingkat Resiko Penyakit Jantung Koroner (PJK) Menggunakan Metode Fuzzy K-Nearest Neighbor (FK-NN)*”. Pada

penelitian ini digunakan metode gabungan Data Mining dengan *Logika Fuzzy*, yaitu *Fuzzy K-Nearest Neighbor* (FK-NN). Dimana metode *Fuzzy K-Nearest Neighbor* dapat melakukan prediksi Penyakit Jantung Koroner sesuai dengan faktor-faktor yang ada. Dimana pada penelitian ini faktor PJK yang digunakan yaitu umur, HDL, trigliserida, LDL, kolesterol dan sistolik. FK-NN melakukan prediksi dengan mencari nilai tetangga terdekat kemudian menggunakan basis nilai keanggotaan data uji dari setiap kelas dengan mengambil nilai keanggotaan terbesar sebagai hasil akhir prediksi. Pengujian metode ini digunakan pada data latih PJK dengan jumlah berbeda. Dimana pada data latih dengan jumlah 40, 50, 60 dan 70 didapat akurasi tertinggi sebesar 66,67% pada data latih 60 dan 70. Semakin banyak data latih yang digunakan, maka kemungkinan semakin banyaknya jarak record yang mendekati kelas data prediksi.

Wisdarianto, A., Ridok, A., dan Rahman, M.A. 2013, Universitas Brawijaya. Penelitiannya yang berjudul "*Penerapan Metode Fuzzy K - Nearest Neighbor (FK-NN) untuk Pengklasifikasian Spam Email*". Pada penelitian ini digunakan metode *Fuzzy K-NN* yang merupakan gabungan dari logika *Fuzzy* dan *K-NN*. Pada metode *Fuzzy K-NN* untuk klasifikasi email spam ini, email uji yang akan diklasifikasi akan diberikan nilai keanggotaan untuk kelas spam dan kelas ham berdasarkan sejumlah k dokumen yang memiliki similarity tertinggi. Proses klasifikasi akan dilakukan dengan memilih nilai keanggotaan kelas tertinggi pada email dokumen uji tersebut. Pada penelitian ini dilakukan pengujian 50 email uji terhadap penggunaan 50, 75,100, 125, 150, 175, hingga 200 data latih dan didapatkan hasil nilai f -measure paling baik yaitu 0,8979 pada penggunaan data latih 175 data latih untuk penggunaan nilai $k=10$ dan $k=15$.

Shofa, R.A., Muflikhah, L., dan Ridok, A. 2013, Universitas Brawijaya. Penelitiannya yang berjudul "*Penerapan Metode Fuzzy K-Nearest Neighbor (FK-NN) untuk Menentukan Kualitas Hasil Rendemen Tanaman Tebu*". Bahan pokok pembuatan gula hampir sebagian besar berasal dari tanaman tebu. Kadar kandungan gula yang dihasilkan dari proses pengolahan nira tebu disuatu pabrik gula biasa disebut dengan rendemen. Algoritma fuzzy k-nearest neighbor dapat digunakan untuk mengklasifikasikan kualitas hasil rendemen tanaman tebu

berdasarkan faktor-faktor yang berpengaruh (panjang tebu (m), diameter tebu (cm), banyak ruas per tebu, berat per meter (kg), persentase (%) brix, harkat kemurnian (HK), dan nira). Pengujian ini dilakukan terhadap 4 macam data latih yaitu 100, 125, 150, dan 175. Pengujian akurasi juga dilakukan pada empat macam jumlah data uji yaitu sebanyak 50, 60, 70, 80 data Uji. Pengujian dilakukan pada tiga tahapan yaitu pengujian k untuk mengetahui pengaruhnya terhadap tingkat akurasi, pengujian pengaruh jumlah data latih terhadap tingkat akurasi, dan pengujian pengaruh jumlah data uji terhadap akurasi. Akurasi tertinggi yang didapatkan pada pengujian ini adalah sebesar 98%. Penelitian ini menunjukkan bahwa algoritma fuzzy k-nearest neighbor memiliki kinerja yang baik dalam klasifikasi kualitas hasil rendemen tanaman tebu.