

Klasifikasi Standar Produk Baja PT. Krakatau Steel (Persero) Tbk. Berdasarkan Komposisi Kimia dan Sifat Mekanis Baja Menggunakan *Fuzzy K-Nearest Neighbor (Fuzzy K-NN)*

Ardisa Tamara Putri¹, M. Tanzil Furqon², Randy Cahya Wihandika³

Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Brawijaya

Email: ¹ardisatamar@gmail.com, ²m.tanzil.furqon@ub.ac.id, ³rendicahya@ub.ac.id

Abstrak

Terjadinya penyimpangan komposisi kimia atau sifat mekanis dalam produksi baja menyebabkan pengelompokan baja berdasarkan standarnya tidak dapat terdefinisi. Penyimpangan yang terjadi adalah penyimpangan dari batas maksimum dari komposisi kimia dan batas minimum dari sifat mekanis baja. Hal tersebut melatar belakangi peneliti untuk membuat sistem dengan menggunakan metode *Fuzzy k-Nearest Neighbor*. Metode *Fuzzy K-Nearest Neighbor (Fuzzy K-NN)* digunakan untuk mengklasifikasikan standar baja berdasarkan komposisi kimia dari baja yang dihasilkan. Data yang diperlukan untuk penelitian ini adalah data yang berisi hasil produksi baja dengan spesifikasi komposisi baja dan sifat mekanis baja tersebut beserta klasifikasi standar baja yang diproduksi. Tahap-tahap yang dilakukan adalah normalisasi data, *Fuzzy k-Nearest Neighbor*, menghitung jarak *Euclidean*, mengambil jarak terpendek sebanyak *k*, menghitung nilai keanggotaan setiap kelas dan menentukan kelas target. Tingkat akurasi tertinggi yang dihasilkan dengan pengujian nilai ketetapan menggunakan *k-fold cross validation* adalah 74,44% dengan nilai ketetapan sama dengan 4 dan jumlah data latih sebanyak 267 data.

Kata Kunci: standar baja, komposisi kimia baja, sifat mekanis baja, baja, fuzzy, fuzzy k-nearest neighbor, k-fold cross validation

Abstract

The occurrence of chemical composition deviations or mechanical properties in steel production causes the classification of steels based on standards can't be defined. The deviation that occurs is the deviation from the maximum limit of the chemical composition and the minimum limit of the mechanical properties of steel. This is the background of researchers to create a system using the *Fuzzy k-Nearest Neighbor* method. The *Fuzzy k-Nearest Neighbor (Fuzzy K-NN)* method used for classifying steel standards based on the chemical composition of the steel produced. The data used for this study is data steel products with the specifications of the steel composition, the mechanical properties of the steel and the classification of standard of steel produced. The steps performed are data normalization, *Fuzzy k-Nearest Neighbor*, calculate *Euclidean* distance, take the shortest distance *k*, calculate the membership value of each class and determine the target class. The highest accuracy resulted by testing *k* values using *k-fold cross validation* is 74,44% with *k* value equal to 4 and total of training data is 267 data.

Keywords: steel standard, steel chemical composition, mechanical properties of steel, steel, fuzzy, fuzzy k-nearest neighbor, k-fold cross validation

1. PENDAHULUAN

PT. Krakatau Steel (Persero) Tbk memiliki enam buah fasilitas produksi yang membuat perusahaan ini menjadi satu-satunya industri baja terpadu di Indonesia. Pabrik-pabrik tersebut mengolah bahan mentah sehingga menghasilkan

berbagai jenis produk baja. Spektrum penggunaan dari baja sangat luas yang karakteristiknya berbeda sehingga diperlukannya standar-standar yang berbeda (Steel, 2015).

Standar-standar yang ada dibedakan berdasarkan jenis baja dan aplikasi dari baja tersebut. Hal yang terpenting dari standar baja

adalah *technical requirement* seperti komposisi kimia, sifat mekanis, standar material dan lain-lain. Pembuatan standar baja dari suatu negara berbeda dengan negara lainnya. Perbedaan komposisi kimia merupakan penyebab dari bedanya kriteria standar. Selain komposisi kimia, sifat mekanis juga berpengaruh dalam pemberian standar baja yang diproduksi. Sifat-sifat mekanis yang dimaksud adalah *elongation*, *tensile strength*, dan *yield strength*. Terjadinya penyimpangan komposisi kimia atau sifat mekanis dalam produksi baja menyebabkan pengelompokan baja berdasarkan standarnya tidak dapat terdefinisi. Penyimpangan yang terjadi adalah penyimpangan dari batas maksimal dari komposisi kimia dan batas minimal dari sifat mekanis baja. Hal tersebut melatar belakangi peneliti untuk membuat sistem dengan menggunakan metode *Fuzzy k-Nearest Neighbor*.

Metode *Fuzzy k-Nearest Neighbor* (*Fuzzy K-NN*) digunakan untuk mengklasifikasikan standar baja berdasarkan komposisi kimia dari baja yang dihasilkan. *Fuzzy k-Nearest Neighbor* merupakan penggabungan dari logika *fuzzy* dan *k-Nearest Neighbor* (*k-NN*). Logika *fuzzy* merupakan cara untuk menempatkan suatu input ke dalam suatu output untuk menentukan tingkat kebenaran dari nilai kebenaran dengan cara perhitungan aritmatika (Tjahjono, et al., 2015). *K-Nearest Neighbor* merupakan algoritma untuk mengklasifikasi data berdasarkan mayoritas dari kategori terhadap banyaknya tetangga terdekat (Liu, 2007).

2. DASAR TEORI

2.1 Standar Baja

Banyaknya karakteristik baja menyebabkan dilakukannya pengelompokan jenis-jenis baja berdasarkan keperluannya. Standar baja yang dikelompokkan bisa berdasarkan jenis baja atau fungsi. Hal yang terpenting dari standar baja adalah *technical requirement* seperti komposisi kimia, sifat mekanis, standar material dan lain-lain. Pembuatan standar baja dari suatu negara berbeda dengan negara lainnya. Perbedaan komposisi kimia merupakan penyebab dari bedanya kriteria standar. Sifat mekanis dapat mempengaruhi pengelompokan standar baja. Sifat mekanis berupa elongasi (*Elongation*), tegangan maksimal baja sebelum patah (*Tensile strength*), dan tegangan maksimal baja sebelum deformasi plastic (*Yield strength*). Kriteria

standar ini mengoptimasi keselamatan dengan biaya anantara pemilihan material desain struktur sehingga suatu standar tidak bisa dikatakan lebih baik daripada standar lainnya (Steel, 2015).

Standar-standar yang digunakan oleh PT. Krakatau Steel adalah ASTM (Amerika), DIN (Jerman), JIS (Jepang), Krakatau Steel Standard (Indonesia), API, SNV, NAW, SAE dan lain-lain.

2.2 K-Nearest Neighbor

K-Nearest Neighbor merupakan algoritma untuk mengklasifikasi data berdasarkan mayoritas dari kategori terhadap banyaknya tetangga terdekat (Liu, 2007). Berikut merupakan langkah-langkah proses algoritma k-Nearest Neighbor:

1. Tentukan nilai k. Nilai k merupakan nilai jumlah tetangga terdekat.
2. Hitung nilai jarak dengan menggunakan persamaan (1) yaitu persamaan kuadrat jarak euclidean data uji terhadap data training yang diberikan.

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^n (a_r(x_i) - a_r(x_j))^2} \quad (1)$$

Keterangan:

$d(x_i, x_j)$ = Jarak *Euclidean*.

(x_i) = *Record* ke-i.

(x_j) = *Record* ke-j.

(a_r) = Data ke-r.

i, j = 1, 2, 3, ..., n.

3. Urutkan hasil jarak yang telah dihitung secara *ascending*.
4. Kumpulkan kategori Y. Y merupakan klasifikasi tetangga terdekat berdasarkan nilai k.
5. Kategori mayoritas merupakan kategori untuk data uji yang di uji.

2.3 Fuzzy K-Nearest Neighbor

Metode *Fuzzy k-Nearest Neighbor* (*Fuzzy K-NN*) untuk mengklasifikasikan Standar baja berdasarkan komposisi kimia dari baja yang dihasilkan. *Fuzzy k-Nearest Neighbor* merupakan penggabungan dari logika *fuzzy* dan *k-Nearest Neighbor* (*k-NN*). *Fuzzy k-Nearest Neighbor* merupakan metode klasifikasi untuk memprediksi data yang ingin diuji menggunakan nilai derajat keanggotaan data uji terhadap setiap kelas. Kelas yang memiliki nilai derajat

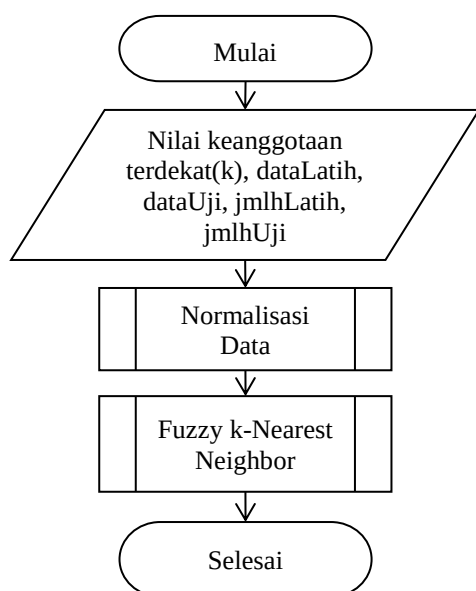
keanggotaan terbesar akan diambil dan dijadikan kelas untuk data yang diuji. (Keller, et al., 1985).

Setiap kelas memiliki nilai derajat keanggotaan yang berbeda terhadap data uji yang memiliki interval $[0, 1]$. Nilai keanggotaan suatu data terhadap masing-masing kelas ditentukan dengan teori himpunan *fuzzy* yang mengeneralisasi teori k-NN klasik (Keller, et al., 1985).

3. PERANCANGAN DAN IMPLEMENTASI

3.1 Perancangan Sistem

Pada tahap perancangan sistem, alur-alur dari metode *Fuzzy k-Nearest Neighbor* dijelaskan menggunakan diagram alur. Gambar 1 merupakan diagram alur *Fuzzy k-Nearest Neighbor*.



Gambar 1 Diagram Alur *Fuzzy K-Nearest Neighbor*

Proses dimulai dari pengguna memasukkan nilai k dan data uji lalu data dinormalisasikan. Setelah itu, tahap *Fuzzy k-Nearest Neighbor* di proses. Pada tahap *Fuzzy k-Nearest Neighbor*, terdapat tahap menghitung jarak *Euclidean*, mengambil jarak terpendek sebanyak k , menghitung nilai keanggotaan setiap kelas dan menentukan kelas target.

3.2 Data Penelitian

Data yang dibutuhkan untuk penelitian ini adalah data yang berisi hasil produksi baja dengan spesifikasi komposisi kimia dan sifat mekanis baja tersebut beserta klasifikasi standar

baja yang diproduksi. Data yang digunakan untuk penelitian ini adalah 300 data. Spesifikasi komposisi baja yang dimaksud adalah kandungan unsur karbon, mangan, sulfur, dan fosfor. Sifat mekanis berupa pertambahan panjang baja sebelum patah (*Elongation*), tegangan maksimal sebelum baja patah (*Tensile strength*), dan tegangan maksimal sebelum baja mengalami deformasi (*Yield strength*).

3.3 Normalisasi Data

Pada tahap ini digunakan normalisasi dengan metode *min-max normalization*. Hasil dari tahap ini adalah data numerik setiap atribut yang sudah dinormalisasi. Persamaan yang digunakan pada tahap ini adalah persamaan (2).

$$V' = \frac{V - \min_A}{\max_A - \min_A} \quad (2)$$

Keterangan:

V' = Hasil normalisasi yang nilainya berkisar antara 0 dan 1.

V = Nilai data numerik yang akan dinormalisasi.

$\min_A$ = Nilai minimum dari kelompok data numerik.

$\max_A$ = Nilai maksimum dari kelompok data numerik.

3.4 Perhitungan Jarak *Euclidean*

Pada tahap menghitung jarak *Euclidean* persamaan yang digunakan adalah persamaan (1). Tabel 1 merupakan tabel hasil perhitungan jarak *Euclidean*.

Tabel 1. Hasil Perhitungan Jarak *Euclidean*

No	Jarak <i>Euclidean</i>
1	0,87624
2	0,86728
3	0,88780
4	0,89359
5	0,85564
...	...
...	
50	0,98252

3.5 Pengambilan Jarak Terpendek Sebanyak K

Pada tahap ini, hasil dari perhitungan jarak *Euclidean* diurutkan dari jarak terpendek hingga terjauh. Setelah itu, jumlah data yang diambil untuk perhitungan selanjutnya ditentukan

berdasarkan dari nilai k . Tabel 2 merupakan tabel jarak terpendek dengan k bernilai 3.

Tabel 2. Jarak Terpendek dengan K Bernilai 3

No	Jarak Euclidean
6	0,14874
9	0,15095
31	0,21397

3.6 Menghitung Nilai Keanggotaan dan Penentuan Kelas Target

Tahap menghitung nilai keanggotaan digunakanlah persamaan (3).

$$u_{ij} = \begin{cases} 0.51 + \left(\frac{n_j}{k}\right) \times 0.49, & \text{jika } j = i \\ \left(\frac{n_j}{k}\right) \times 0.49, & \text{jika } j \neq i \end{cases} \quad (3)$$

Keterangan:

n_j = Jumlah anggota kelas j pada suatu dataset

k = Nilai tetangga terdekat

j = Kelas

Nilai k adalah 3, nilai jumlah kelas j pada dataset (n_j) adalah 5, u_1 adalah nilai keanggotaan kelas API5L, u_2 adalah nilai keanggotaan kelas ASTM A283, u_3 adalah nilai keanggotaan kelas SNI 07 0601, u_4 adalah nilai keanggotaan kelas Krakatau Steel, u_5 adalah nilai keanggotaan kelas JIS G3131, u_6 adalah nilai keanggotaan kelas SAE1006, u_7 adalah nilai keanggotaan kelas JIS G3113, u_8 adalah nilai keanggotaan kelas JIS G3116, u_9 adalah nilai keanggotaan kelas JIS A5525 dan u_{10} adalah nilai keanggotaan kelas JIS G3101. Setelah didapat nilai keanggotaan, maka nilai keanggotaan data dihitung menggunakan persamaan (4).

$$u_i(x) = \frac{\sum_{j=1}^k u_{ij}(1/\|x-x_j\|^{2/(m-1)})}{\sum_{j=1}^k (1/\|x-x_j\|^{2/(m-1)})} \quad (4)$$

Keterangan:

u_i = Fungsi Fuzzy k -Nearest Neighbor

u_{ij} = Nilai keanggotaan fuzzy pada contoh pengujian

k = Banyaknya nilai ketetanggaan terdekat

j = Variabel data untuk keanggotaan data latih

i = Variabel data untuk keanggotaan data uji

m = Berat pangkat

Tabel 3 merupakan tabel hasil perhitungan nilai keanggotaan suatu data tiap kelas. Nilai keanggotaan maksimum terdapat pada u_2 sehingga data uji yang diuji memiliki standar baja ASTM A283.

Tabel 3. Nilai Keanggotaan Setiap Kelas

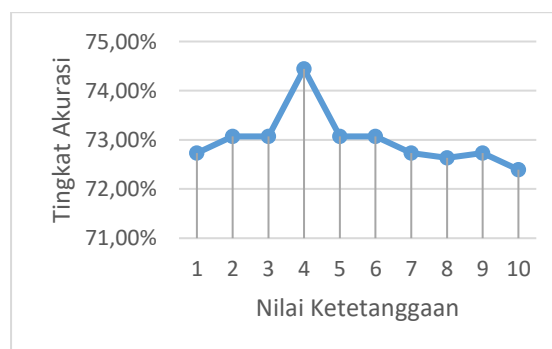
u_i	Nilai Keanggotaan
u_1	0,29451
u_2	0,70408
u_3	0,29451
u_4	0,29451
u_5	0,29451
u_6	0,29451
u_7	0,39493
u_8	0,29451
u_9	0,29451
u_{10}	0,29451

4. PENGUJIAN DAN ANALISIS

Pengujian dilakukan bertujuan untuk menguji tingkat akurasi dari sistem klasifikasi standar baja. Pengujian yang dilakukan adalah pengujian tingkat akurasi terhadap sepuluh data uji, pengujian tingkat akurasi berdasarkan nilai k dan pengujian akurasi berdasarkan jumlah data latih dengan metode k -fold cross validation.

4.1 Pengujian Tingkat Akurasi Berdasarkan Nilai K dengan Metode K-fold Cross Validation

Pengujian tingkat akurasi berdasarkan nilai k dengan metode k -fold cross validation dilakukan dengan data yang dikelompokkan menjadi data latih dan data uji. Pada pengujian ini, nilai k diinisialisasi dengan nilai 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 dan data yang digunakan sebanyak 267 data. Nilai k -fold yang digunakan adalah 9. Gambar 2 merupakan grafik pengaruh nilai k terhadap tingkat akurasi.



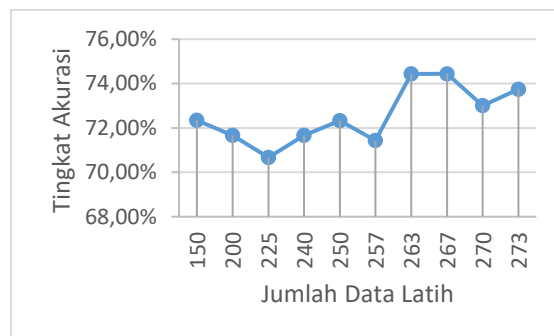
Gambar 2 Grafik Pengaruh Nilai K Terhadap Tingkat Akurasi

Gambar 2 menunjukkan bahwa nilai ketetanggaan mempengaruhi tingkat akurasi. Jika nilai ketetanggaan adalah 1 maka tingkat akurasi sebesar 72,73%. Pada saat nilai

ketetanggaan adalah 2 maka tingkat akurasi sebesar 73,07% dan terjadi peningkatan tingkat akurasi. Saat nilai ketetanggaan adalah 3 maka tingkat akurasi yang dihasilkan sebesar 73,07%. Pada saat nilai ketetanggaan adalah 4 maka tingkat akurasi sebesar 74,44% dan terjadi peningkatan tingkat akurasi. Jika nilai ketetanggaan adalah 5 maka tingkat akurasi sebesar 73,07% dan terjadi penurunan tingkat akurasi. Pada saat nilai ketetanggaan adalah 6 maka tingkat akurasi sebesar 73,07%. Saat nilai ketetanggaan adalah 7 maka tingkat akurasi yang dihasilkan sebesar 72,73%. Pada saat nilai ketetanggaan adalah 8 maka tingkat akurasi sebesar 72,63% dan terjadi penurunan tingkat akurasi. Jika nilai ketetanggaan adalah 9 maka tingkat akurasi sebesar 72,73%. Pada saat nilai ketetanggaan adalah 10 maka tingkat akurasi sebesar 72,39% dan terjadi peningkatan tingkat akurasi. Tingkat akurasi tertinggi didapatkan dengan nilai ketetanggaan sebesar 4. Rata-rata yang dihasilkan dalam pengujian ini adalah 72,99%. Hal ini menunjukkan bahwa jika nilai ketetanggaan bernilai kecil maka nilai keanggotaan yang dihasilkan belum dapat mewakili kelas target. Jika nilai ketetanggaan semakin besar maka data yang memiliki jarak yang jauh ikut dihitung pada proses perhitungan nilai keanggotaan setiap kelas yang mengakibatkan kesalahan dalam klasifikasi.

4.2 Pengujian Tingkat Akurasi Berdasarkan Jumlah Data Latih dengan Metode *K-fold Cross Validation*

Pengujian tingkat akurasi berdasarkan jumlah data latih dilakukan dengan metode *k-fold cross validation*. Pada Penelitian sebelumnya didapatkan tingkat akurasi yang tinggi ketika nilai *k* adalah 4 sehingga pada pengujian ini digunakan nilai *k* tersebut. Nilai *k-fold* yang digunakan adalah 2, 3, 4, 5, 6, 7, 8, 9, 10 dan 11. Jumlah data latih yang digunakan sebanyak 150, 200, 225, 240, 250, 257, 263, 267, 270 dan 273 data. Gambar 3 merupakan grafik mengenai pengaruh nilai *k* terhadap tingkat akurasi dengan metode *k-fold cross validation*.



Gambar 3 Grafik Pengaruh Jumlah Data Latih Terhadap Tingkat Akurasi

Berdasarkan Grafik 3, jika jumlah data latih sebanyak 150 data maka tingkat akurasi sebesar 72,34%. Pada saat jumlah data latih sebanyak 200 data maka tingkat akurasi sebesar 71,67% dan terjadi penurunan tingkat akurasi. Saat jumlah data latih sebanyak 225 data maka tingkat akurasi yang dihasilkan sebesar 70,67%. Pada jumlah data latih 225 data, tingkat akurasi mengalami penurunan. Pada saat jumlah data latih sebanyak 240 data maka tingkat akurasi sebesar 71,67% dan terjadi peningkatan tingkat akurasi. Jika jumlah data latih sebanyak 250 data maka tingkat akurasi sebesar 72,33%. Pada saat jumlah data latih sebanyak 257 data maka tingkat akurasi sebesar 71,43% dan terjadi penurunan tingkat akurasi. Saat jumlah data latih sebanyak 263 data maka tingkat akurasi yang dihasilkan sebesar 74,43%. Pada jumlah data latih 263 data, tingkat akurasi mengalami peningkatan. Pada saat jumlah data latih sebanyak 267 data maka tingkat akurasi sebesar 74,44%. Jika jumlah data latih sebanyak 270 data maka tingkat akurasi sebesar 73,00%. Pada saat jumlah data latih sebanyak 273 data maka tingkat akurasi sebesar 73,74%. Tingkat akurasi tertinggi didapatkan dengan jumlah data latih sebesar 267 data dengan nilai ketetanggaan sebesar 4. Rata-rata yang dihasilkan dalam pengujian ini adalah 72,57%. Hal ini menunjukkan bahwa jumlah data latih berpengaruh terhadap tingkat akurasi yang dihasilkan. Hal ini disebabkan oleh pengambilan data uji secara acak sehingga komposisi data setiap kelas tidak seimbang.

5. KESIMPULAN

Berdasarkan hasil penelitian mengenai klasifikasi standar baja berdasarkan sifat mekanis dan komposisi kimia baja menggunakan metode *Fuzzy k-Nearest Neighbor*, dapat disimpulkan bahwa metode *Fuzzy k-*

Nearest Neighbor dapat digunakan untuk kasus klasifikasi standar baja berdasarkan komposisi kimia dan sifat mekanis baja. Tingkat akurasi tertinggi yang dihasilkan dengan pengujian nilai k menggunakan *k-fold cross validation* adalah 74,44% dengan nilai k sama dengan 4 dan jumlah data latih sebanyak 267 data. Terjadi penurunan tingkat akurasi ketika nilai k semakin besar. Jumlah data latih berpengaruh terhadap tingkat akurasi. Tingkat akurasi tertinggi yang didapatkan sebesar 74,44% dengan nilai k sama dengan 5. Jumlah data yang digunakan tiap kelas berbeda-beda juga menyebabkan tingkat akurasi yang berbeda dan tidak stabil.

DAFTAR PUSTAKA

- Han, J., Kamber, M. & Pei, J., 2006. *Data Mining: Concepts and Techniques*. Morgan Kaufmann.
- Keller, M. J., Gray, R. M. & Givens, A. J., 1985. *A Fuzzy K-Nearest Neighbor*. *IEEE Transactions on System, Man and Cyberetics*, SMC-15(4).
- Liu, B., 2007. *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*. Berlin: Springer.
- Steel, P. K., 2015. *PT. Krakatau Steel (Persero) Tbk. Product Facilities*. Cilegon: PT. Krakatau Steel (Persero) Tbk.
- Tjahjono, A., Purwanto, E. & Rizky, Y., 2015. *Penerapan Logika Fuzzy Pada Proses Pencucian Filter Air Dengan Menggunakan Plc Di Pdam Karangpilang I Surabaya*.