

# DIAGNOSIS KEGAGALAN TRANSFORMATOR DAYA MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBORS DAN ARTIFICIAL NEURAL NETWORK BERBASIS DISSOLVED GAS ANALYSIS

Rafi Maulana Hidayat<sup>\*</sup>), Galih Media, Nike Sartika dan Lia Kamelia

Departemen Teknik Elektro, Universitas Islam Negeri Sunan Gunung Djati, Bandung, Indonesia

<sup>\*</sup>Penulis korespondensi, E-mail: rafimaulanahidayat@gmail.com

## Abstrak

*Dissolved Gas Analysis (DGA)* merupakan metode untuk mengidentifikasi jenis kegagalan pada transformator dengan menilai jumlah gas yang terkandung pada minyak isolasi transformator. DGA memiliki beberapa metode dalam menganalisis dan mengidentifikasi jenis kegagalan berdasarkan jenis gas yang terlarut. Tetapi, dalam jumlah data yang besar metode ini menjadi sulit dan memerlukan keahlian dalam mendeteksi kegagalan secara grafis. Penelitian ini bertujuan untuk meningkatkan akurasi diagnostik kegagalan transformator dengan mengimplementasikan serta membandingkan algoritma *K-Nearest Neighbors* (KNN) dengan algoritma *Artificial Neural Network* (ANN) pada setiap metode konvensional DGA yaitu *Roger Ratio*, *Duval Triangle*, *Four Gases* dan *Duval pentagon* dalam mengklasifikasikan jenis kegagalan. Sebanyak 822 sampel dataset digunakan untuk melatih dan memvalidasi model yang digunakan. Berdasarkan hasil penelitian diagnosis kegagalan transformator menunjukkan bahwa metode grafis memberikan hasil yang paling efektif dalam mendiagnosis jenis kegagalan transformator dibandingkan dengan metode rasio. Selain itu penggunaan algoritma KNN memberikan hasil yang lebih baik dibandingkan dengan algoritma ANN dalam meningkatkan akurasi diagnostik dengan memperoleh akurasi tertinggi sebesar 98% pada *duval triangle*.

*Kata kunci: Artificial Neural Network, Diagnosis Kegagalan, Dissolved Gas Analysis, K-Nearest Neighbors, Transformator.*

## Abstract

*Dissolved Gas Analysis (DGA)* is a method to identify the type of failure in a transformer by assessing the amount of gas contained in the transformer insulating oil. DGA has several methods of analyzing and identifying the type of failure based on the type of dissolved gas. However, in large amounts of data this method becomes difficult and requires expertise in graphically detecting failures. This study aims to improve the diagnostic accuracy of transformer failures by implementing and comparing the *K-Nearest Neighbors* (KNN) algorithm with the *Artificial Neural Network* (ANN) algorithm on each of the conventional DGA methods namely *Roger Ratio*, *Duval Triangle*, *Four Gases* and *Duval pentagon* in classifying failure types. A total of 822 dataset samples were used to train and validate the model used. Based on the results of the transformer failure diagnosis research, it shows that the graphical method provides the most effective results in diagnosing the type of transformer failure compared to the ratio method. In addition, the use of the KNN algorithm provides better results than the ANN algorithm in improving diagnostic accuracy by obtaining the highest accuracy of 98% on the *duval triangle*.

*Keywords: Artificial Neural Network, Dissolved Gas Analysis, Failure Diagnosis, K Nearest Neighbors, Transformer.*

## 1. Pendahuluan

Transformator merupakan komponen statis dalam infrastruktur tenaga listrik yang beroperasi berdasarkan prinsip induksi elektromagnetik untuk mentransfer daya antar sirkuit dengan frekuensi serupa [1]. Dalam sistem kelistrikan, transformator berperan dalam mengubah tegangan dan arus sesuai kebutuhan, namun selama operasinya, perangkat ini dapat mengalami tekanan listrik dan panas yang berpotensi menurunkan kinerja. Salah satu aspek kritis dalam pemeliharaan transformator adalah

minyak isolasi, yang tidak hanya berfungsi sebagai medium isolasi tetapi juga sebagai pendingin [2], [3]. Seiring waktu, kualitas minyak transformator dapat mengalami degradasi akibat adanya *thermal stress*, *over voltage*, dan *mechanical stress*, yang berkontribusi pada pembentukan gas terlarut seperti hidrogen, metana, etana, etilena, asetilena, karbon monoksida, dan karbon dioksida [4]. Konsentrasi gas ini dapat menjadi indikator kondisi operasional transformator, memungkinkan identifikasi dini terhadap potensi kegagalan.

Dalam konteks industri energi listrik, sistem transmisi dan distribusi harus mampu menyalurkan tenaga listrik secara efisien dan andal ke berbagai sektor. Transformator daya berperan sentral dalam jaringan listrik, di mana kondisi operasionalnya secara langsung memengaruhi kinerja sistem. Degradasi minyak isolasi, terutama dalam transformator berbasis minyak, dapat mempercepat pembentukan gas terlarut, yang apabila mencapai tingkat kontaminasi tinggi, dapat meningkatkan suhu serta menyebabkan gangguan kelistrikan [4], [5]. Oleh karena itu, tindakan preventif sangat diperlukan untuk menjaga keandalan sistem, salah satunya dengan metode *Dissolved Gas Analysis (DGA)*. Pengujian DGA memungkinkan analisis kondisi transformator melalui identifikasi dan kuantifikasi gas terlarut dalam minyak transformator, sehingga dapat menentukan jenis serta tingkat kegagalan yang terjadi [6]. Pendekatan ini memainkan peran penting dalam strategi pemeliharaan preventif guna memastikan kontinuitas operasional dan efisiensi sistem tenaga listrik.

DGA sendiri memiliki beberapa metode untuk menganalisis dan mengidentifikasi berdasarkan jenis gas terlarut diantaranya ada metode *Total Dissolved Combustible Gas*, *Key Gas*, *Roger's Ratio*, *The Four Gases*, *Duval Triangle* dan *Duval Pentagon* [6], [7], [8], [9], [10]. Metode konvensional ini yang bergantung pada nilai konsentrasi gas terlarut dalam minyak transformator untuk mendiagnosis gejala awal pada transformator [11]. Tetapi, dengan jumlah data yang besar, metode ini menjadi sulit digunakan karena memerlukan keahlian dalam mengidentifikasi kegagalan secara grafis.

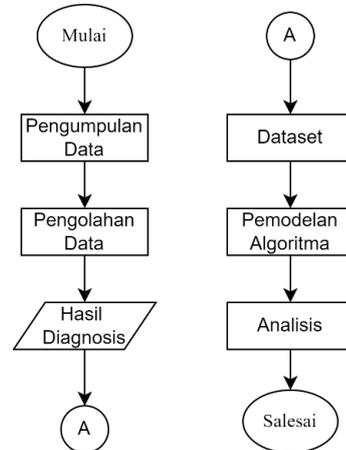
Dalam diagnosis DGA, *machine learning* digunakan untuk mengklasifikasikan jenis kegagalan transformator berdasarkan data konsentrasi gas terlarut atau untuk memprediksi tingkat keparahan kegagalan transformator berdasarkan data historis DGA [12]. Untuk mengatasi keterbatasan metode tersebut, sudah banyak penelitian yang menggunakan algoritma kecerdasan buatan seperti *Artificial Neural Network* dan *K-Nearest Neighbors* untuk meningkatkan akurasi diagnostik hingga 93% tetapi hanya beberapa metode konvensional DGA dengan hanya memperhatikan nilai k untuk menentukan jumlah tetangga tanpa memperhatikan metrik jarak yang digunakan [13].

Oleh karena itu, fokus penelitian ini menitikberatkan pada diagnosis kegagalan transformator daya pada beberapa metode konvensional DGA dengan mengimplementasikan algoritma *Artificial Neural Network* dan algoritma *K Nearest Neighbors* dalam meningkatkan akurasi diagnosis kegagalan transformator daya. Selain itu, penelitian ini juga berfokus pada penggunaan matrik jarak yang digunakan dalam algoritma *K Nearest Neighbors*.

## 2. Metode

Metode implementasi algoritma *K-Nearest Neighbors* dan *Artificial Neural Network* dalam mengklasifikasikan jenis

kegagalan pada transformator daya ditunjukkan oleh Gambar 1 Metode ini terdiri dari tahapan, mengumpulkan data historis DGA, pengolahan data menggunakan metode konvensional DGA, pembuatan model algoritma KNN dan ANN, dan kemudian menganalisis hasil pemodelan.



Gambar 1. Flowchart Penelitian

### 2.1. Pengumpulan Data

Penelitian ini bertujuan untuk mengidentifikasi jenis kegagalan isolasi pada minyak transformator daya berdasarkan data historis hasil analisis gas terlarut atau *Dissolved Gas Analysis (DGA)*. Data yang digunakan dalam penelitian ini diperoleh dari Unit Induk Transmisi Jawa Tengah PT PLN (Persero) serta *Database Dataport IEEE* dan mencerminkan konsentrasi gas-gas terlarut dalam satuan ppm yang telah diuji oleh pihak terkait sebelumnya. Tahapan pengumpulan data merupakan aspek krusial dalam proses diagnosis dan klasifikasi kegagalan isolasi, di mana data yang diperoleh digunakan sebagai data *training* dan *testing* dalam penerapan algoritma *K-Nearest Neighbor (KNN)* serta *Artificial Neural Network (ANN)*. Analisis ini berfokus pada nilai gas hidrokarbon yang telah diukur, sehingga dapat mendukung implementasi metode kecerdasan buatan dalam meningkatkan akurasi identifikasi kegagalan isolasi pada transformator daya.

### 2.2. Pengolahan Data

Setelah memperoleh data *Dissolved Gas Analysis (DGA)* pada transformator daya, langkah selanjutnya adalah pengolahan data guna menentukan jenis kegagalan yang terjadi. Data yang dianalisis berupa konsentrasi gas terlarut dalam satuan ppm, yang kemudian diklasifikasikan menggunakan empat metode konvensional, yaitu *Roger Ratio*, *Duval Triangle*, *Four Gases*, dan *Duval Pentagon*. Tahapan pengolahan data dilakukan untuk memastikan validitas dan kualitas data sebelum analisis lebih lanjut. Proses ini mencakup pembersihan data yang tidak diperlukan, termasuk data

*null*, data *outlier*, serta data yang tidak konsisten. Setelah melalui tahap validasi, data yang telah disaring kemudian diolah menggunakan keempat metode konvensional DGA guna mendukung identifikasi dan klasifikasi kegagalan transformator daya secara sistematis. Tahapan pengolahan data diperlukan untuk memastikan bahwa data valid dan berkualitas, yaitu dengan melakukan pembersihan data yang tidak diperlukan, seperti data *null* atau kosong, serta data *outlier*, dan data yang tidak konsisten. Setelah dilakukan validasi, selanjutnya data diolah dengan menggunakan keempat metode konvensional DGA yaitu:

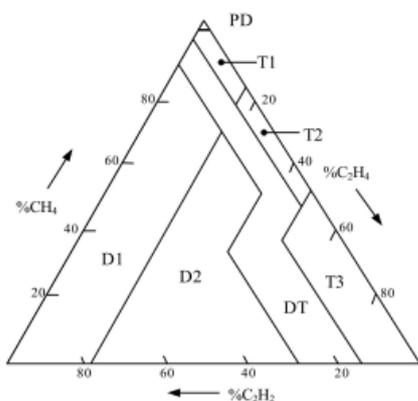
### 2.2.1. Roger Ratio

Roger ratio menggunakan rasio konsentrasi gas-gas utama untuk mengidentifikasi kegagalan pada transformator. Gas yang digunakan adalah: Hidrogen ( $H_2$ ), Metana ( $CH_4$ ), Asetilena ( $C_2H_2$ ), Etana ( $C_2H_6$ ), dan Etilen ( $C_2H_4$ ) [13]. Tabel 1 menunjukkan daftar kondisi transformator berdasarkan tiga rasio.

Table 1. Jenis Kegagalan Metode Roger Ratio [14].

Case	R1 $C_2H_2/C_2H_4$	R2 $CH_4/H_2$	R3 $C_2H_4/C_2H_6$	Type of Fault
0	<0.1	>0.1 to <1.0	<0.1	Unit Normal
1	<0.1	<0.1	<0.1	Low Energy Density Arching (PD)
2	0.1 to 3.0	0.1 to 1.0	>3.0	Arching – High Energy Discharge
3	<0.1	>0.1 to <1.0	1.0 to 3.0	Low Temperature Thermal (LT)
4	<0.1	>1.0	1.0 to 3.0	Thermal Fault <700°C (TF1)
5	<0.1	>1.0	>1.0	Thermal Fault >700°C (TF2)

### 2.2.2. Duval Triangle



Gambar 2. Duval Triangle [14].

Duval triangle adalah plot terner yang merupakan representasi segitiga dari persentase relatif tiga gas yang mudah terbakar  $CH_4$ ,  $C_2H_2$  dan  $C_2H_4$  [4], [14]. Ketiga gas mudah terbakar tersebut diplot dalam sumbu kartesius X-Y. Konsentrasi gas ini akan dihitung menggunakan (1)

sampai (3) dan direpresentasikan dari ketiga sisi diagram segitiga seperti pada Gambar 2.

$$CH_4 \% = \frac{100x}{x+y+z} \quad \text{for, } x = [CH_4] \text{ in ppm} \quad (1)$$

$$C_2H_4 \% = \frac{100y}{x+y+z} \quad \text{for, } y = [C_2H_4] \text{ in ppm} \quad (2)$$

$$C_2H_2 \% = \frac{100z}{x+y+z} \quad \text{for, } z = [C_2H_2] \text{ in ppm} \quad (3)$$

### 2.2.3. Four Gasses

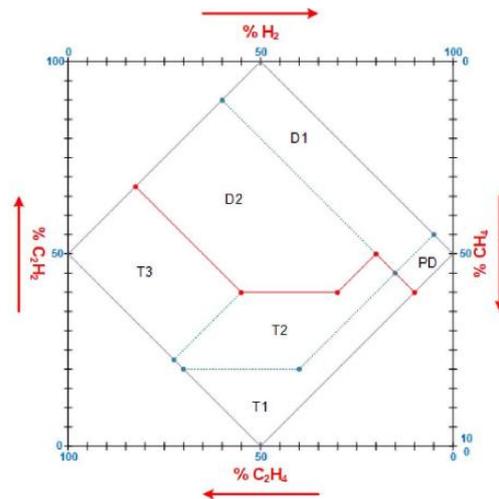
Metode ini menentukan jenis kegagalan dengan menggunakan persentase relatif dari empat gas yang mudah terbakar yaitu ( $C_2H_2$ ,  $C_2H_4$ ,  $CH_4$ ,  $H_2$ ) dan jenis kegagalan ditentukan dengan menghitung persentase setiap gas mudah terbakar [10]. Konsentrasi gas ini akan dihitung menggunakan (4) sampai (7) dan direpresentasikan dari keempat sisi diagram segiempat, kemudian dicari titik pusat dari diagonal yang akan saling berpotongan untuk menentukan jenis kegagalan seperti pada Gambar 3.

$$CH_4 \% = \frac{CH_4}{CH_4 + C_2H_2 + C_2H_4 + H_2} \quad (4)$$

$$C_2H_2 \% = \frac{C_2H_2}{CH_4 + C_2H_2 + C_2H_4 + H_2} \quad (5)$$

$$C_2H_4 \% = \frac{C_2H_4}{CH_4 + C_2H_2 + C_2H_4 + H_2} \quad (6)$$

$$H_2 \% = \frac{H_2}{CH_4 + C_2H_2 + C_2H_4 + H_2} \quad (7)$$



Gambar 3. Four Gasses [10].

### 2.2.4. Duval Pentagon

Duval Pentagon adalah plot terner yang merupakan representasi Pentagon dari persentase relatif lima gas yang mudah terbakar. Hal ini diplot dalam sumbu kartesius X-Y [4], [14]. Konsentrasi gas ini akan dihitung menggunakan (8) sampai (12) dan direpresentasikan dari

kelima sisi diagram *pentagon* menggunakan seperti pada Gambar 3.

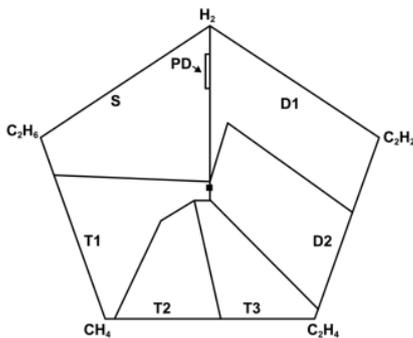
$$CH_4\% = \frac{CH_4}{CH_4 + C_2H_4 + C_2H_2 + C_2H_6 + H_2} \quad (8)$$

$$C_2H_4\% = \frac{C_2H_4}{CH_4 + C_2H_4 + C_2H_2 + C_2H_6 + H_2} \quad (9)$$

$$C_2H_2\% = \frac{C_2H_2}{CH_4 + C_2H_4 + C_2H_2 + C_2H_6 + H_2} \quad (10)$$

$$C_2H_6\% = \frac{C_2H_6}{CH_4 + C_2H_4 + C_2H_2 + C_2H_6 + H_2} \quad (11)$$

$$H_2\% = \frac{H_2}{CH_4 + C_2H_4 + C_2H_2 + C_2H_6 + H_2} \quad (12)$$



Gambar 4. *Pentagon* [14].

### 2.3. Hasil Diagnosis

Tahapan ini memberikan hasil diagnosis dari keempat metode konvensional DGA dan mengklasifikasikan jenis kegagalan pada transformator. Hasil analisis yang dilakukan sesuai dengan teori dan teknik pengolahan data DGA akan ditunjukkan dalam tahap ini.

### 2.4. Dataset

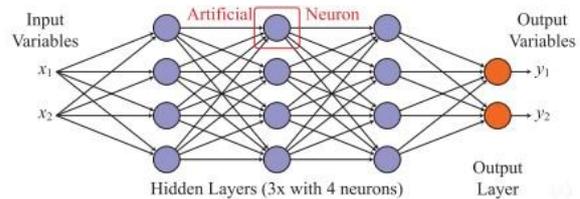
Tahapan ini mengkonversi hasil diagnosis yang diperoleh menjadi bentuk file *Comma Separated Values* (CSV). Dataset yang diperoleh kemudian dibagi menjadi empat dataset sesuai metode konvensional DGA yang digunakan.

### 2.5. Pemodelan Algoritma KNN dan ANN

Setelah dataset diperoleh, data akan dibagi menjadi data pelatihan sebesar 80% dan data pengujian sebesar 20%. Data pelatihan akan digunakan untuk melatih pemodelan algoritma KNN dengan cara menentukan nilai *k* dan menghitung jarak antara titik data. Sedangkan pada algoritma ANN data pelatihan dilatih dengan menggunakan beberapa epoch hingga model mencapai konvergensi atau kinerja yang diinginkan. Kemudian, data pengujian akan digunakan untuk menguji kinerja pemodelan KNN dan ANN dengan cara mengklasifikasikan data berdasarkan data pelatihan.

### 2.5.1. Algoritma ANN

*Artificial Neural Network* (ANN) atau bisa disebut Jaringan Saraf Tiruan (JST) merupakan metode kecerdasan buatan yang memiliki prinsip seperti otak biologis manusia [15]. Aplikasi utama ANN adalah untuk tugas pemrosesan data, seperti pengenalan dan klasifikasi pola yang kompleks tanpa memerlukan banyak pengetahuan tentang prosesnya [15], [16]. Suatu contoh penerapan konsep dari jaringan syaraf biologis adalah adanya komponen pemrosesan pada jaringan syaraf tiruan yang sering saling terhubung dan beroperasi secara paralel, meniru struktur neuron-neuron dalam jaringan syaraf biologis [16]. Gambar 5 Menunjukkan Struktur komputasi ANN



Gambar 5. Struktur Komputasi Algoritma ANN [16].

### 2.5.2. Algoritma KNN

Algoritma *K-Nearest Neighbors* (KNN) merupakan metode *supervised learning* yang digunakan untuk mengklasifikasikan suatu objek dengan *distance matrix* yang digunakan. *Distance measures* berperan dalam menentukan tingkat kemiripan dan ketidakmiripan antara dua vektor data, sehingga metode ini banyak diterapkan dalam pengenalan pola [17], [18], [19]. Dalam analisis ini, model KNN menghitung tingkat akurasi setiap metode klasifikasi menggunakan fungsi dari matriks evaluasi, kemudian membandingkannya untuk menentukan metode yang paling sesuai dalam penerapan algoritma KNN. Beberapa jenis *distance matrix* yang dapat digunakan dalam perhitungan meliputi *Canberra*, *Bray-Curtis*, dan *Euclidean*, masing-masing dengan karakteristik khusus dalam mengukur perbedaan antar data.

#### 2.5.2.1. Canberra Distance

*Canberra distance* merupakan varian terukur dari *Manhattan distance* yang telah dimodifikasi dengan pemberian bobot berdasarkan nilai absolut dari variabel yang digunakan. Metode ini mengkuantifikasi jumlah relatif perbedaan antara dua vektor, dengan mempertahankan nilai absolut untuk memastikan hasil perhitungan tetap dalam rentang 0 hingga 1. Jika salah satu koordinat bernilai nol, komponen jarak akan mencapai nilai maksimum, terlepas dari besaran nilai elemen lainnya, sehingga nilai jarak yang dihasilkan tetap konstan [19], [20]. Fenomena ini didefinisikan melalui persamaan:

$$d_{ij} = \sum_{k=1}^n \frac{|X_{ik} - X_{jk}|}{|X_{ik}| + |X_{jk}|} \quad (13)$$

Selain itu, dalam kasus di mana kedua elemen memiliki nilai nol, diperlukan pendekatan khusus guna mendefinisikan hasil perhitungan agar tetap valid [20]. Jarak Canberra memiliki tingkat sensitivitas yang tinggi terhadap perubahan kecil, terutama ketika nilai kedua elemen berada dalam kisaran yang sangat rendah, menjadikannya metode yang relevan dalam analisis data dengan variabilitas signifikan.

### 2.5.2.2. Euclidean Distance

*Euclidean distance* merupakan metode pengukuran jarak antara dua titik dalam ruang *Euclidean* dengan menghitung akar kuadrat dari jumlah kuadrat selisih setiap koordinat. Jarak ini digunakan untuk menentukan tingkat kesamaan atau perbedaan antara dua elemen dalam ruang multidimensi [19], [20]. Persamaan matematis yang digunakan untuk menghitung Jarak *Euclidean* diberikan sebagai berikut:

$$d_{ij} = \sqrt{\sum_{k=1}^n (X_{ik} - X_{jk})^2} \quad (16)$$

Dimana

- $d_{ij}$  adalah jarak antara titik  $i$  dan  $j$
- $n$  adalah jumlah dimensi atau atribut yang dibandingkan
- $X_{ik}$  adalah nilai atribut ke- $k$  dari pusat kluster  $i$
- $X_{jk}$  adalah nilai atribut ke- $k$  dari pusat kluster  $j$

*Euclidean Distance* sering digunakan dalam berbagai bidang, termasuk pemrosesan data, analisis kluster, dan pembelajaran mesin, karena sifatnya yang intuitif dan kemampuannya dalam mengukur kesamaan antara data dengan presisi tinggi.

### 2.5.2.3. Bray Curtis Distance

*Bray Curtis Distance* merupakan bentuk modifikasi dari jarak Manhattan yang diperoleh melalui proses normalisasi terhadap selisih nilai variabel. Normalisasi ini menjadikan ruang analisis menyerupai model kisi-kisi yang karakteristiknya mirip dengan *City Block Distance*. Apabila seluruh elemen dalam perhitungan jarak *Bray Curtis* bernilai positif, maka nilai jarak ini akan berada dalam rentang antara nol hingga satu. Kondisi ketika nilai *Bray Curtis Distance* bernilai nol menunjukkan bahwa kedua vektor yang dibandingkan memiliki kesamaan sempurna [19], [20]. *Bray Curtis Distance* dapat dihitung menggunakan persamaan berikut:

$$d_{ij} = \sum_{k=1}^n \frac{|X_{ik} - X_{jk}|}{(X_{ik} + X_{jk})} \quad (18)$$

Pada umumnya, algoritma *K-Nearest Neighbor* (KNN) menggunakan jarak *Euclidean* sebagai metode utama dalam proses klasifikasinya. Namun, variasi ukuran jarak kemiripan, seperti *Bray Curtis*, menawarkan alternatif yang dapat meningkatkan analisis diferensiasi kelas. Perhitungan jarak kemiripan ini memungkinkan ekstraksi informasi kesamaan objek data secara lebih presisi, terutama jika dikombinasikan dengan teknik klasifikasi berbasis algoritma yang efisien.

## 2.6. Analisis

Tahap berikutnya yaitu membandingkan tingkat hasil akurasi pengolahan data menggunakan algoritma KNN dan ANN pada setiap metode konvensional. Hasil analisis yang dilakukan sesuai dengan teori dan teknik pengolahan data DGA akan ditunjukkan dalam tahap ini.

## 3. Hasil dan Pembahasan

### 3.1. Hasil Diagnosis Kegagalan Metode Konvensional DGA

Dalam studi kasus ini, terdapat 822 sampel data pengujian DGA yang digunakan dalam implementasi algoritma KNN untuk mendiagnosis jenis kegagalan berdasarkan metode konvensional DGA. Terdapat lima jenis gas hidrokarbon yakni hidrogen ( $H_2$ ), metana ( $CH_4$ ), etana ( $C_2H_6$ ), etilen ( $C_2H_4$ ), dan asetilen ( $C_2H_2$ ). Hasil jumlah diagnosis dari keempat metode konvensional DGA ditabulasikan pada Tabel 2.

Dari hasil pengolahan data yang diperoleh dengan keempat metode konvensional DGA, Setiap metode memperoleh konvensional memberikan hasil yang berbeda beda dalam menentukan jenis kegagalan pada transformator seperti yang ditunjukkan oleh Gambar 5, hal ini disebabkan dikarenakan setiap metode konvensional DGA yang digunakan memiliki jenis gas yang berbeda beda dalam mendiagnosis jenis kegagalan pada transformator daya. Metode *roger's ratio* dapat memberikan hasil diagnosis normal pada transformator, tetapi metode ini memiliki kekurangan yaitu banyaknya data pengujian DGA yang tidak dapat didiagnosis. Kemudian untuk ketiga metode *graphical* DGA yaitu *duval triangle*, *four gases* dan *duval pentagon* memberikan hasil yang signifikan dalam mendiagnosis jenis kegagalan pada transformator, tetapi ketiga metode ini tidak dapat mendiagnosis transformator dalam keadaan normal. Secara keseluruhan, metode interpretasi *graphical* dapat memberikan hasil yang lebih baik daripada metode rasio dalam mendiagnosis jenis kegagalan pada transformator. Tetapi, keempat metode konvensional DGA ini cukup baik dalam diagnosis jenis kegagalan dengan masing-masing kelebihan mengingat terdapat perbedaan jenis gas hidrokarbon yang digunakan dalam setiap metode tersebut.

Tabel 2. Jumlah Diagnosis Pada Setiap Metode Konvensional

Jenis Kegagalan	Label Kegagalan	Jumlah Diagnosis			
		Roger Ratio	Duval Triangle	Four Gasses	Duval Pentagon
Thermal Fault < 300° C / Low Thermal Temperature	T1	19	117	312	171
Thermal Fault, 300° C < T < 700° C	T2	31	50	123	45
Thermal Fault > 700° C	T3	63	148	104	121
Discharge Low Energy	D1	-	120	55	85
Discharge High Energy	D2	109	125	164	153
Thermal and Electrical fault	DT	-	41	-	-
Partial Discharge	PD	9	221	64	0
Stray Gassing pada temperature < 200° C	S	-	-	-	247
Normal	N	21	-	-	-
No Diagnosis	ND	570	-	-	-
<b>Total Data</b>		<b>822</b>	<b>822</b>	<b>822</b>	<b>822</b>

### 3.2. Hasil Implementasi Algoritma KNN

Dalam sistem pemantauan transformator daya berbasis *Dissolved Gas Analysis (DGA)*, pemilihan algoritma klasifikasi yang optimal menjadi aspek kritis dalam peningkatan ketepatan diagnosis. *K-Nearest Neighbors (KNN)* merupakan salah satu algoritma yang umum digunakan dalam klasifikasi kegagalan transformator berdasarkan pola konsentrasi gas terlarut dalam minyak isolasi. Keberhasilan implementasi KNN sangat dipengaruhi oleh parameter utama, yaitu jumlah tetangga (K) dan matriks jarak yang digunakan untuk mengukur kedekatan antar data.

Matriks jarak berfungsi sebagai metrik untuk mengukur kesamaan atau perbedaan antar vektor data dalam ruang multidimensi. Pemilihan matriks jarak yang tepat dapat meningkatkan akurasi sistem klasifikasi, mengoptimalkan normalisasi data, serta mengurangi pengaruh outlier. Berdasarkan hasil pengujian terhadap matriks jarak *Canberra*, *Bray-Curtis*, dan *Euclidean*, diperoleh variasi akurasi klasifikasi berdasarkan vektor input yang digunakan, sebagaimana yang tersaji dalam Tabel 3, Tabel 4, dan Tabel 5.

Table 3. Nilai K Terbaik dengan Matrik Jarak Canberra

Input	Canberra Distance			
	Roger Ratio	Duval Triangle	Duval Pentagon	Four Gasses
K	1	3	1	3
Fold	5	4	1	2
Akurasi	94,51%	96,95%	89,90%	90,30%

Table 4. Nilai K Terbaik dengan Matrik Jarak Euclidean

Input	Euclidean Distance			
	Roger Ratio	Duval Triangle	Duval Pentagon	Four Gasses
K	1	5	1	3
Fold	5	4	6	5
Akurasi	85,37%	97,56%	91,46%	93,29%

Table 5. Nilai K Terbaik dengan Matrik Jarak Bray Curtis

Input	Bray Curtis Distance			
	Roger Ratio	Duval Triangle	Duval Pentagon	Four Gasses
K	1	5	1	4
Fold	1	4	5	5
Akurasi	90,30%	98,17%	90,24%	93,28%

Hasil eksperimen menunjukkan bahwa setiap matriks jarak memiliki karakteristik yang berbeda dalam menangani jenis vektor input tertentu. Pada vektor *Roger's Ratio*, akurasi tertinggi diperoleh dengan matriks jarak *Canberra Distance* dengan nilai K = 1, menghasilkan akurasi 94.51%. Faktor utama yang menyebabkan keunggulan *Canberra Distance* adalah kemampuannya dalam menangani data dengan nilai nol. *Roger's Ratio* memiliki banyak atribut dengan nilai nol akibat absennya jenis gas tertentu dalam sampel, sehingga matriks jarak yang mampu mempertimbangkan faktor ini lebih unggul. Selain itu, *Canberra Distance* menerapkan skala absolut dalam perhitungan jarak, di mana setiap elemen dibagi dengan jumlah nilai absolutnya. Hal ini berkontribusi dalam meningkatkan *robustnes* terhadap *outlier*, sehingga membantu dalam normalisasi data dengan skala yang berbeda.

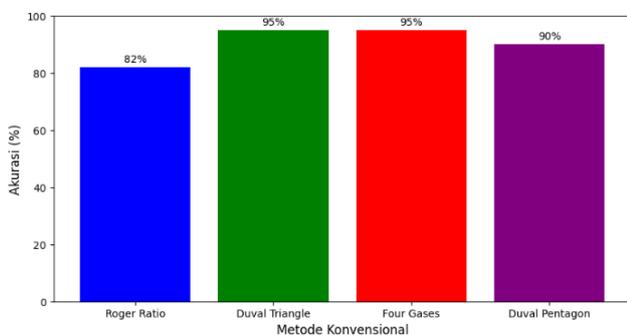
Pada vektor *Duval Triangle*, matriks *Bray-Curtis Distance* memberikan akurasi tertinggi sebesar 98.17% dengan nilai K = 4. *Duval Triangle* memiliki karakteristik unik, yakni proporsi data dengan nilai nol yang cukup tinggi serta jumlah label yang lebih kompleks dibandingkan vektor lainnya. Keunggulan *Bray-Curtis* dalam menangani data bernilai nol serta ketahanan terhadap distribusi label yang bervariasi menjadikannya pilihan yang optimal. *Bray-Curtis* juga mampu mengukur perbedaan komposisi antar dua titik data berdasarkan hubungan antara jumlah total elemen. Sifat ini sangat relevan dalam analisis gas terlarut, di mana konsentrasi relatif antar gas menjadi lebih signifikan daripada sekadar nilai absolutnya.

Pada vektor *Four Gases* dan *Duval Pentagon*, matriks *Euclidean Distance* menunjukkan kinerja optimal dengan akurasi masing-masing 93.29% dan 91.46%, dengan nilai  $K = 3$  dan  $K = 1$ . *Euclidean Distance* merupakan metrik jarak klasik yang mengukur perbedaan absolut antara titik-titik dalam ruang multidimensi. Matriks ini sangat efektif dalam menangani data numerik kontinu yang memiliki distribusi kompleks.

Namun, salah satu kelemahan *Euclidean Distance* adalah sensitivitas terhadap perbedaan skala data. Oleh karena itu, normalisasi sangat diperlukan agar perbedaan nilai antar atribut tidak menyebabkan bias dalam klasifikasi. Vektor *Four Gases* dan *Duval Pentagon* memiliki jumlah atribut yang lebih banyak dibandingkan vektor lainnya, sehingga metode perhitungan jarak berbasis akar kuadrat sangat efektif dalam membandingkan variasi kompleksitas antar variabel.

### 3.3. Hasil Implementasi Algoritma ANN

Setelah melakukan klasifikasi dari keempat metode konvensional untuk mengklasifikasikan jenis kegagalan yang terjadi menunjukkan bahwa hasil klasifikasi memiliki nilai akurasi yang baik untuk masing-masing metode DGA konvensional yang digunakan. Gambar 6 menunjukkan grafik akurasi untuk klasifikasi jenis kegagalan dari semua metode konvensional yang digunakan.



Gambar 6. Grafik perbandingan akurasi algoritma ANN pada setiap metode konvensional

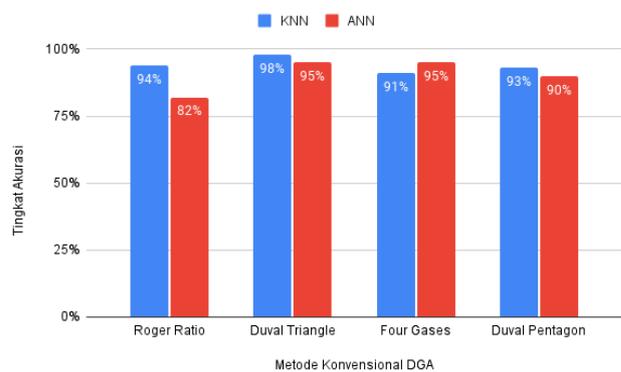
Dari grafik pada Gambar 6 dapat dilihat bahwa hasil klasifikasi terbaik ada pada metode *duval triangle* dan *four gases* dengan mencapai nilai akurasi 95%, hasil ini menunjukkan bahwa secara keseluruhan model bekerja dengan sangat baik dalam menjalankan tugas klasifikasi. Hasil klasifikasi dari *duval pentagon* mendapat nilai akurasi sebesar 90%, angka tersebut menandakan bahwa model masih bekerja dengan cukup baik dalam melakukan klasifikasi jenis kegagalan. Model klasifikasi pada metode *duval pentagon* ini memiliki kekurangan pada pembagian dataset yang mana terjadi ketidakseimbangan data pada jenis kegagalan stray atau S yang menyebabkan jenis atau kelas kegagalan lain tidak

seimbang. Selanjutnya adalah hasil klasifikasi metode *roger ratio* yang mendapatkan nilai akurasi dalam klasifikasi sebesar 82%, secara keseluruhan model klasifikasi ini masih bekerja dengan cukup baik, namun terdapat beberapa kelas yang tidak seimbang sehingga menyebabkan menurunnya akurasi keseluruhan pada model dan dataset yang digunakan.

### 3.4. Evaluasi Model

Dalam sistem pemantauan transformator daya berbasis *Dissolved Gas Analysis (DGA)*, pemilihan algoritma pembelajaran mesin berperan krusial dalam menentukan ketepatan klasifikasi jenis kegagalan. Pada penelitian ini, dua algoritma populer *K-Nearest Neighbors (KNN)* dan *Artificial Neural Network (ANN)* dievaluasi berdasarkan performanya dalam menganalisis metode konvensional DGA, seperti *Duval Triangle*, *Four Gases*, *Duval Pentagon*, dan *Roger's Ratio*.

Hasil analisis menunjukkan bahwa algoritma KNN mencapai akurasi tertinggi 98% pada metode *Duval Triangle*, sementara algoritma ANN memiliki performa terbaik 95% pada metode *Four Gases*. Namun, pada metode *Roger's Ratio*, KNN menunjukkan akurasi lebih tinggi dibandingkan ANN. Perbedaan ini menegaskan bahwa efektivitas algoritma sangat bergantung pada jenis data serta metode konvensional DGA yang digunakan. Secara keseluruhan, algoritma KNN menunjukkan performa yang lebih baik dibandingkan dengan algoritma ANN pada sebagian besar metode konvensional DGA. Algoritma KNN mencapai akurasi tertinggi pada metode *Duval Triangle* dengan nilai 98%, sedangkan algoritma ANN mencapai akurasi tertinggi pada metode *Four Gases* dengan nilai 95%. Namun, pada metode *Roger Ratio*, algoritma KNN justru memiliki akurasi yang lebih tinggi dibandingkan algoritma ANN.



Gambar 7. Grafik perbandingan akurasi algoritma KNN dan ANN

Algoritma KNN menunjukkan performa superior dalam beberapa aspek klasifikasi DGA, terutama dalam metode berbasis pola grafis seperti *Duval Triangle*, *Four Gases*,

dan *Duval Pentagon*. KNN bekerja dengan prinsip pencocokan pola melalui tetangga terdekat (*K-nearest neighbors*), menjadikannya sangat efektif dalam mendeteksi kesamaan distribusi gas dalam metode grafis. Pemilihan matriks jarak yang tepat memungkinkan KNN menangani dataset dengan proporsi nol yang signifikan, seperti yang ditemukan dalam metode *Roger's Ratio* ketika menggunakan *Canberra Distance*. Meskipun memiliki keunggulan dalam metode tertentu. Nilai K yang terlalu kecil dapat menyebabkan model terlalu responsif terhadap outlier, sedangkan nilai K yang terlalu besar bisa menyebabkan kesalahan klasifikasi akibat kelebihan generalisasi. Berbagai matriks seperti *Euclidean*, *Canberra*, dan *Bray-Curtis* memiliki dampak signifikan terhadap kinerja model, sehingga pemilihan metrik yang tidak sesuai dapat menurunkan akurasi klasifikasi.

ANN memberikan akurasi tinggi pada metode *Four Gases*, mencapai 95%, menunjukkan keunggulannya dalam menangkap pola hubungan kompleks antar variabel. ANN dapat mengoptimalkan klasifikasi dengan mempelajari hubungan kompleks antar atribut gas, yang sulit dideteksi menggunakan metode berbasis similarity seperti KNN. ANN menunjukkan stabilitas dalam menangani metode *Four Gases*, yang memiliki kompleksitas lebih tinggi dalam distribusi rasio gas dibandingkan *Duval Triangle*. ANN memiliki fleksibilitas tinggi dalam penyesuaian parameter seperti jumlah hidden layers, learning rate, dan fungsi aktivasi yang dapat meningkatkan akurasi klasifikasi. Meskipun ANN memiliki keunggulan dalam menangkap pola kompleks, ANN membutuhkan jumlah data yang besar untuk pelatihan optimal. Jika dataset terbatas atau distribusi label tidak seimbang, akurasi bisa menurun akibat *overfitting* atau *underfitting*. ANN memerlukan tuning hyperparameter yang lebih kompleks dibandingkan KNN, termasuk pemilihan jumlah neuron dalam setiap lapisan, teknik regularisasi, dan optimasi fungsi aktivasi.

#### 4. Kesimpulan

Berdasarkan hasil pengujian dan analisis yang diperoleh metode plotting grafis lebih baik dibandingkan dengan metode ratio dalam mendiagnosis jenis kegagalan transformator berdasarkan kandungan gas hidrogen di dalamnya. Model algoritma ANN menunjukkan performa terbaik saat menggunakan metode *duval triangle* dan *four gases* dengan akurasi 95%, kemudian untuk akurasi hasil klasifikasi *duval pentagon* sebesar 90% dan *roger ratio* sebesar 82%. Sedangkan, Model algoritma KNN menunjukkan performa terbaik saat menggunakan metode *duval triangle* dengan akurasi 98%, kemudian untuk akurasi hasil klasifikasi *duval pentagon* sebesar 93% dan *roger ratio* sebesar 94%. Hal ini menunjukkan penggunaan algoritma KNN dapat membantu meningkatkan akurasi dalam mendiagnosis jenis kegagalan dengan lebih baik dibandingkan dengan algoritma ANN. Penelitian lanjutan dapat dilakukan dengan menambahkan jumlah dataset dan

metode konvensional DGA digunakan serta penggunaan *distance matrik* lainnya pada algoritma KNN.

#### Referensi

- [1]. J. J. Winders, *Power transformers: principles and applications*. Marcel Dekker, 2002.
- [2]. A. Fauji, R. Kurnianto, dan M. Rajagukguk, "Trafo Oil Failure Based on Dissolved Gas Analysis and Voltage Breakdown Results," *Telecommunications, Computers, and Electricals Engineering Journal*, vol. 1, no. 1, 2023, doi: 10.26418/telectrical.v1i1.69800.
- [3]. V. A. Thivyanathan, P. J. Ker, Y. S. Leong, F. Abdullah, A. Ismail, dan M. Zaini Jamaludin, "Power Transformer Insulation System: A Review on the Reactions, Fault Detection, Challenges and Future Prospects," 1 Oktober 2022, *Elsevier B.V.* doi: 10.1016/j.aej.2022.01.026.
- [4]. R. Furqaranda dan S. Suwarno, "Analisa Minyak Isolasi Transformator Daya dengan Metode Dissolved Gas Analysis (DGA) dan Purifikasi," *Jurnal Ilmiah Ecosystem*, vol. 23, no. 2, hlm. 441–449, Agu 2023, doi: 10.35965/eco.v23i2.2871.
- [5]. J. F. Vidal dan A. R. G. Castro, "Diagnosing Faults in Power Transformers with Variational Autoencoder, Genetic Programming, and Neural Network," *IEEE Access*, vol. 11, 2023, doi: 10.1109/ACCESS.2023.3258544.
- [6]. D. C. Febrina, "Feasibility Test of Transformer Insulating Oil Lifetime Based on Breakdown Voltage (bdv) and Furan Analysis," *JEECS (Journal of Electrical Engineering and Computer Sciences)*, vol. 7, no. 1, hlm. 1177–1184, Jan 2023, doi: 10.54732/jeeecs.v7i1.212.
- [7]. A. P. Junfithrana, L. Hakim, R. Nopiyanti, dan M. Muslih, "Improving TDGC on Oil Dissolved Tranformator Gas Using Fuzzy Logic," dalam *6th International Conference on Computing, Engineering, and Design, ICCED 2020*, 2020. doi: 10.1109/ICCED51276.2020.9415819.
- [8]. O. E. Gouda, S. H. El-Hoshy, dan S. S. M. Ghoneim, "Enhancing the Diagnostic Accuracy of DGA Techniques Based on IEC-TC10 and Related Databases," *IEEE Access*, vol. 9, 2021, doi: 10.1109/ACCESS.2021.3107332.
- [9]. H. Sutikno, R. A. Prasajo, dan Suwarno, "Integration of Duval Pentagon to the Multi-Method Interpretation to Improve the Accuracy of Dissolved Gas Analysis Technique," dalam *Proceedings of the IEEE International Conference on Properties and Applications of Dielectric Materials*, 2021. doi: 10.1109/ICPADM49635.2021.9493929.
- [10]. S. J. Lee, Y. M. Kim, H. D. Seo, J. R. Jung, H. J. Yang, dan M. Duval, "New methods of DGA diagnosis using IEC TC 10 and related databases Part 2: Application of relative content of fault gases," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 20, no. 2, 2013, doi: 10.1109/TDEI.2013.6508774.
- [11]. A. Siswanto, A. Rohman, S. Suprijadi, M. Baehaqi, dan A. Arifudin, "Analisis Karakteristik Minyak Transformator Menggunakan Pengujian Dissolved Gas Analysis (DGA) pada IBT 1 Gardu Induk," *Foristek*, vol. 12, no. 1, 2022, doi: 10.54757/fs.v12i1.142.

- [12]. M. K. Ngwenyama dan M. N. Gitau, "Discernment of transformer oil stray gassing anomalies using machine learning classification techniques," *Sci Rep*, vol. 14, no. 1, 2024, doi: 10.1038/s41598-023-50833-7.
- [13]. O. Kherif, Y. Benmahamed, M. Tegar, A. Boubakeur, dan S. S. M. Ghoneim, "Accuracy Improvement of Power Transformer Faults Diagnostic Using KNN Classifier With Decision Tree Principle," *IEEE Access*, vol. 9, hlm. 81693–81701, 2021, doi: 10.1109/ACCESS.2021.3086135.
- [14]. C57.104-2019 - *IEEE Guide for the Interpretation of Gases Generated in Mineral Oil-Immersed Transformers*. IEEE, 2019.
- [15]. T. Guillod, P. Papamanolis, dan J. W. Kolar, "Artificial neural network (ann) based fast and accurate inductor modeling and design," *IEEE Open Journal of Power Electronics*, vol. 1, 2020, doi: 10.1109/OJPEL.2020.3012777.
- [16]. A. R. E. Soto, S. L. Lima, dan O. R. Saavedra, "Incipient fault diagnosis in power transformers by DGA using a machine learning ANN - Mean shift approach," dalam *2019 IEEE International Autumn Meeting on Power, Electronics and Computing, ROPEC 2019*, 2019. doi: 10.1109/ROPEC48299.2019.9057143.
- [17]. S. Zhang, "Challenges in KNN Classification," *IEEE Trans Knowl Data Eng*, 2021, doi: 10.1109/TKDE.2021.3049250.
- [18]. Z. Zuriati dan N. Qomariyah, "Klasifikasi Penyakit Stroke Menggunakan Algoritma K-Nearest Neighbor (KNN)," *ROUTERS: Jurnal Sistem dan Teknologi Informasi*, 2022, doi: 10.25181/rt.v1i1.2665.
- [19]. A. Pulungan, M. Zarlis, dan S. Suwilo, "Performance Analysis of Distance Measures in K-Nearest Neighbor," 2020. doi: 10.4108/eai.3-8-2019.2290748.
- [20]. M. Faisal, E. M. Zamzami, dan Sutarman, "Comparative Analysis of Inter-Centroid K-Means Performance using Euclidean Distance, Canberra Distance and Manhattan Distance," dalam *Journal of Physics: Conference Series*, 2020. doi: 10.1088/1742-6596/1566/1/012112.