

METODE RANDOM FOREST UNTUK KLASIFIKASI KERUSAKAN TRANSFORMATOR DAYA BERDASARKAN GAS TERLARUT PADA DUVAL TRIANGLE DAN DUVAL PENTAGON

Hunayn Risatayn, Ekojono, Dhebys Suryani Hormansyah
Program Studi Teknik Informatika, Politeknik Negeri Malang
Jl. Soekarno Hatta No. 9 Malang, Indonesia
hunaynr@gmail.com

ABSTRAK

Penyaluran energi listrik merupakan hal yang esensial dalam operasional komponen elektronik. Kinerja transformator untuk memenuhi kebutuhan transmisi bergantung pada sistem isolasi. Sistem isolasi berupa minyak transformator akan mengalami degradasi kualitas seiring dengan periode pemanfaatan alat. Peninjauan mutu dari minyak transformator secara periodik diperlukan untuk memastikan performa sistem. Analisis konsentrasi gas terlarut menggunakan teknik interpretasi *duval triangle* dan *duval pentagon* diterapkan untuk mendapatkan informasi terkait kelas kerusakan transformator daya. *Duval Triangle* mampu memberikan lokasi titik koordinat *centroid* berdasarkan konsentrasi gas terlarut yang diperoleh secara riil di lapangan oleh operator. Fungsi serupa terdapat dalam metode *Duval Pentagon* dengan identifikasi gas meliputi hidrogen, etana, metana, etilen dan asetilen yang terkandung dalam sistem isolasi transformator. Kurangnya tingkat pemahaman operator dalam mendefinisikan kelas kerusakan memicu pengaplikasian metode klasifikasi. Metode *Random Forest* menawarkan proses klasifikasi dengan waktu yang relatif singkat tanpa diperlukan pemahaman yang mendalam bagi operator. Himpunan pohon keputusan yang terbentuk dalam metode memberikan keluaran independen dan secara kolektif dihasilkan konklusi dari mayoritas label pohon keputusan menggunakan teknik interpretasi *Duval Triangle* dan metode klasifikasi *Random Forest* sebesar 96.011% pada kombinasi *testing size*=25% dan *n_estimator*=125, sedangkan teknik interpretasi *Duval Pentagon* menghasilkan tingkat akurasi tertinggi dengan nilai 98.693% pada kombinasi *testing size*=15% dan *n_estimator*=100.

Kata kunci: DGA, Duval Triangle, Duval Pentagon, Random Forest, Decision Tree.

1. PENDAHULUAN

Transformator adalah alat yang digunakan untuk memindahkan energi listrik arus bolak balik dari satu rangkaian ke rangkaian yang lain. Salah satu parameter tingkat kualitas kerja transformator adalah sistem isolasi. Seiring dengan bertambahnya periode pengoperasian transformator, sistem isolasi berupa minyak transformator mengalami degradasi kualitas oleh kontaminan berupa partikel gas akibat adanya tegangan lebih, pemanasan, kerusakan mekanis dan gangguan lingkungan. Apabila minyak dengan kondisi tersebut tetap digunakan dalam pengoperasian transformator daya, maka akan terjadi kegagalan isolasi yang berdampak pada tingkat optimalitas transmisi aliran listrik. Untuk mencegah terjadinya hal tersebut, perlu dilakukan pengujian pada minyak transformator sehingga kualitas minyak dapat terpantau secara periodik.

Dissolved Gas Analysis (DGA) adalah metode yang digunakan secara luas untuk memonitoring transformator daya. Metode ini menggunakan konsentrasi dari gas-gas terlarut dalam minyak transformator. *Duval Triangle* salah satu metode analisis hasil DGA menggunakan tiga konsentrasi gas terlarut yaitu metana, etilen dan asetilen. Metode analisis hasil DGA yang diterapkan berikutnya adalah *Duval Pentagon* dengan parameter input berupa konsentrasi gas hidrogen, etana, metana, etilen dan asetilen. Teknik interpretasi tersebut akan

menghasilkan titik koordinat *centroid* yang mengindikasikan kelas kerusakan transformator. Penerapan metode analisis duval sangat bergantung pada penilaian dan kajian tenaga ahli dampak kurangnya pemahaman dalam penggunaan grafik sehingga dibutuhkan waktu yang lebih lama untuk mendapatkan suatu kesimpulan. Metode *Random Forest* diajukan terhadap permasalahan tersebut menggunakan mayoritas keputusan dari seluruh output *decision tree* yang dihasilkan. Berdasarkan paparan keunggulan metode, penulis terdorong untuk melakukan penelitian dengan membangun sistem yang dapat melakukan diagnosis secara cepat kerusakan transformator daya berdasarkan metode DGA menggunakan algoritma *Random Forest*. Sehingga dengan adanya sistem tersebut, diharapkan kerusakan transformator dapat dideteksi lebih awal.

2. TINJAUAN PUSTAKA

2.1. Duval Triangle

Metode *Duval Triangle* untuk menganalisis gas terlarut pada transformator menggunakan konsentrasi gas metana, etilen dan asetilen. Konsentrasi tersebut ditampilkan dalam persentase total ($C_2H_2 + C_2H_4 + CH_4$) dan direpresentasikan sebagai titik ($\%C_2H_2, \%C_2H_4, \%CH_4$) pada diagram sistem koordinat triangular yang telah dipisahkan ke dalam daerah/zona kerusakan. Daerah kerusakan dimana titik terletak menggambarkan tipe kerusakan yang

dihasilkan oleh kombinasi konsentrasi gas-gas tersebut [1].

Untuk memetakan *Duval Triangle*, koordinat triangular harus dikonversi kedalam koordinat kartesius menggunakan trigonometri. Kalkulasi koordinat kartesius diperoleh dari tiga pecahan proporsi relatif gas penyusun. Titik pusat diletakkan pada koordinat asal (0,0) meliputi elemen panjang dari sisi segitiga [2].

2.2. Duval Pentagon

Metode analisis *Dissolved Gas Analysis* (DGA) *result Duval Pentagon* menggunakan tambahan dua gas dari tipe gas pada *Duval Triangle* yaitu hidrogen dan etana. Kelas baru hasil klasifikasi ditambahkan dengan nama *S*, *Stray gas zone*, yang berlaku ketika gas dalam keadaan normal saat operasi dari transformator [3]. Komponen dari gas-gas tersebut, H_2 , CH_4 , C_2H_6 , C_2H_4 dan C_2H_2 , dan total gas dihitung dalam persentase. Sumbu dari pentagon mencakup panjang dari 0% sampai 100% dimana pusat dari pentagon berperan sebagai titik mulai. Titik pusat dari konsentrasi gas dihitung dan dipetakan ke dalam *Duval Pentagon* [4].

Urutan dari kelima puncak gas dari pentagon menyatakan peningkatan energi yang dibutuhkan untuk memproduksi gas-gas tersebut dalam transformator daya, dari H_2 sampai C_2H_2 yaitu H_2 , C_2H_6 , CH_4 , C_2H_4 , C_2H_2 , berwalanan arah jarum jam. Urutan ini dapat memberikan solusi terbaik dalam topik identifikasi kerusakan transformator menggunakan representasi pentagon [5].

2.3. Random Forest

Random Forest adalah *supervised* dan *ensemble learning algorithm* yang digunakan untuk klasifikasi maupun prediksi. *Ensemble learning* pada *random forest* adalah gabungan *random decision tree classifiers* yang membuat prediksi dengan cara mengombinasikan prediksi dari seluruh pohon keputusan [6]. Sesuai dengan yang tersirat dalam namanya, *random forest* adalah kumpulan dari pohon keputusan yang dihasilkan secara acak [7], dimana merupakan salah satu metode yang terintegrasi berdasarkan algoritma *bagging* untuk klasifikasi [8]. *Bagging* (*Bootstrap Agorithm*) adalah desain inti yang membentuk fitur acak pada *random forest*. Secara umum, proses *bagging* memilih secara acak beberapa data sample dari dataset training untuk membentuk *decision tree*, dan proses tersebut biasanya berjalan beberapa putaran untuk membentuk *decision tree forest*. Data sampel terpilih yang telah digunakan dalam setiap putaran pada proses *bagging* akan dikembalikan lagi ke dalam dataset dan dapat juga digunakan untuk putaran berikutnya pada proses *bagging* untuk memastikan keacakan dari beberapa pohon keputusan [9].

2.4. Algoritma ID3

Algoritma ID3 fokus pada pengambilan nilai *information gain* sebagai kriteria pemilihan attribute pada *nodes* untuk seluruh level pohon keputusan (*decision tree*), sehingga memungkinkan bagi setiap *non-leaf node* yaitu *root node* dan *internal node* untuk menetapkan attribute dengan *information gain* terbesar sebagai attribute pada *node* sementara dalam pengujian [10].

Terdapat contoh himpunan *S*, memuat data dengan *C* sebagai jumlah keluaran/label kelas data, nilai *entropy* dari *S* adalah

$$Entropy(S) = \sum_{i=1}^C (P_{(i)} \log_2 P_{(i)}) \quad (1)$$

Dimana $p_{(i)}$ adalah proporsi *S* yang termasuk dalam kelas *i*. Catatan bahwa *S* bukanlah fitur melainkan seluruh dataset. *Entropy* bernilai 0 jika seluruh member dari *S* termasuk dalam kelas yang sama. Range dari entropy adalah 0 (*purity*) sampai 1 (*impurity*). Pengukuran berikutnya adalah *information gain* yang digunakan untuk mengukur pengurangan nilai *entropy*. Untuk fitur tertentu *A*, $Gain(S, A)$ bermakna nilai *information gain* dari dataset *S* pada fitur *A* dan didefinisikan oleh persamaan berikut

$$Gain(S, A) = Entropy(S) - \sum_{n=1}^v \left[\left(\frac{|S_v|}{|S|} \right) Entropy(S_v) \right] \quad (2)$$

dimana:

Σ = Jumlah setiap kemungkinan nilai *v* dari fitur *A*,

S_v = Subset dari *S* dimana fitur *A* mempunyai nilai *v*,

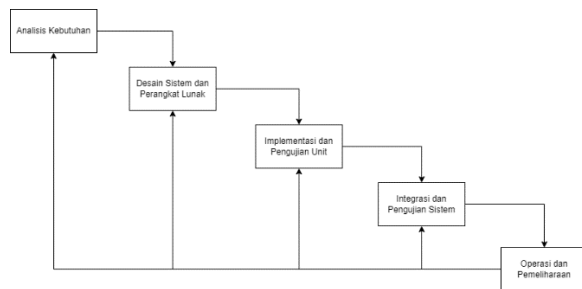
$|S_v|$ = jumlah elemen dalam S_v

$|S|$ = jumlah elemen dalam *S*

ID3 melakukan pencarian terhadap seluruh fitur dari dataset training dan menentukan fitur yang dapat memisahkan data terbaik. Apabila fitur dapat mengklasifikasikan dataset training dengan sempurna/baik, ID3 akan berhenti, selain itu algoritma akan tetap beroperasi [11].

3. METODE PENELITIAN

Pendekatan metode *waterfall* memberikan penekanan secara linier dalam proses pengembangan perangkat lunak melalui beberapa tahapan secara terstruktur. Lingkup dan tingkat kompleksitas penelitian yang tergolong spesifik serta durasi pelaksanaan yang terjadwal mendorong terbentuknya karakteristik yang sesuai untuk mengimplementasikan metode ini berdasarkan rangkaian aktivitas yang tercantum pada gambar 1 berikut



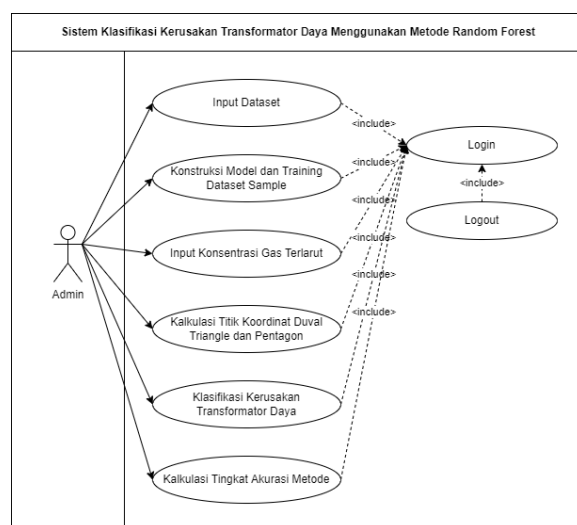
Gambar 1. Diagram waterfall

3.1. Teknik Pengumpulan Data

Data yang digunakan merupakan data sekunder berupa konsentrasi gas-gas terlarut pada minyak transformator. Data tersebut diperoleh dari hasil wawancara dengan Dosen Listrik Politeknik Negeri Malang yang telah melakukan pengujian di laboratorium terhadap minyak transformator untuk mengetahui substansi yang menjadi komponen dalam menentukan tingkat atau kelas kerusakan transformator daya melalui analisis berdasarkan *Duval Triangle* meliputi konsentrasi C_2H_2 , C_2H_4 , dan CH_4 sedangkan berdasarkan *Duval Pentagon* meliputi konsentrasi H_2 , C_2H_6 , CH_4 , C_2H_4 dan C_2H_2

3.2. Use Case Diagram

Use Case Diagram memberikan ilustrasi yang berlangsung antara actor terhadap sistem. Interaksi ini mencerminkan urutan aktivitas perilaku user dalam sistem sehingga dapat dihasilkan potret kegiatan secara sistematis. Kegiatan yang dimiliki oleh user Dosen Listrik Politeknik Negeri Malang yang memegang peran sebagai admin sistem memiliki akses diantaranya melakukan login, input dataset, input konsentrasi gas terlarut, klasifikasi kerusakan transformator daya, dan kalkulasi tingkat akurasi metode sesuai dengan visualisasi pada gambar 2

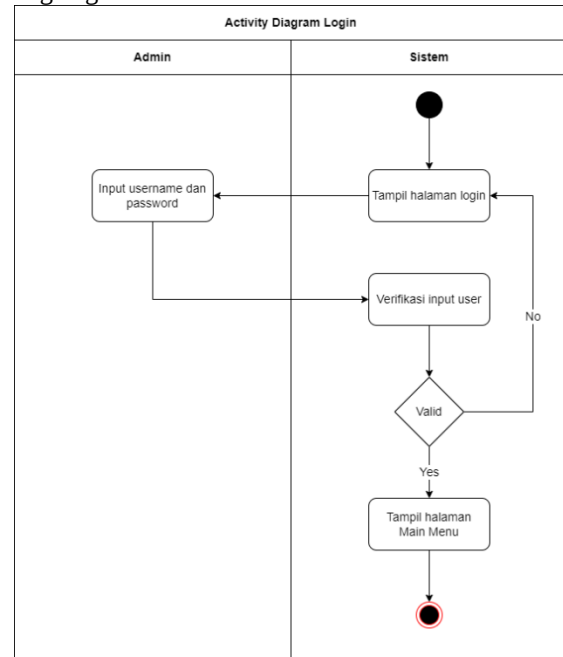


Gambar 2. Use case diagram

3.3. Activity Diagram Login

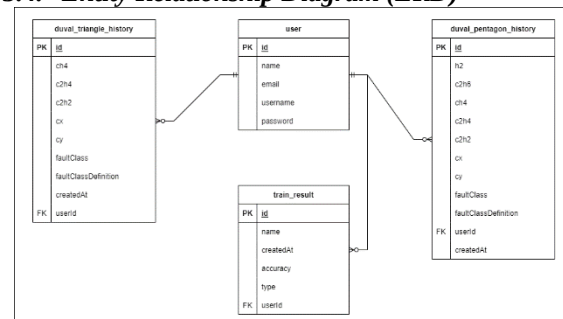
Kegiatan *login* diawali oleh tampilan sistem mewakili input *field username* dan *password* beserta

login button. Data yang diberikan oleh user akan dilakukan verifikasi terhadap sistem dalam pangkalan data, apabila data valid user akan diarahkan menuju ke halaman *main menu* dan jika tidak sistem memberikan *alert* bahwa percobaan *login* tidak berhasil kemudian menginstruksikan untuk melakukan *login* ulang sesuai dengan gambar 3



Gambar 3. Activity diagram login

3.4. Entity Relationship Diagram (ERD)



Gambar 4. Struktur entity relationship diagram

Komponen entitas bertindak sebagai objek dalam proses interaksi yang memiliki fungsi khusus dan berlainan sesuai dengan peran yang ditetapkan. Berikut fungsi dari entitas dalam peancangan relasi database sistem:

1. User

Entitas *user* berfungsi sebagai media untuk menyimpan data pengguna. Objek ini berelasi dengan seluruh tabel lain dengan identitas *id* yang menjadi atribut pembeda untuk setiap record data dalam tabel lain

2. Duval_Triangle_History

Entitas *duval_triangle_history* digunakan untuk menyimpan data hasil klasifikasi input user dalam sistem menggunakan teknik interpretasi *duval triangle*. Data tersebut meliputi persentase relatif

gas metana, etilen dan asetilen, titik koordinat hasil interpretasi dengan metode analisis gas terlarut, dan kelas/target klasifikasi beserta definisi kelas

3. *Duval_Pentagon_History*

Entitas *duval_pentagon_history* digunakan untuk menyimpan data hasil klasifikasi input user dalam sistem menggunakan teknik interpretasi *duval pentagon*. Data tersebut meliputi persentase relatif gas hidrogen, etana, metana, etilen dan asetilen, titik koordinat hasil interpretasi dengan metode analisis gas terlarut, dan kelas/target klasifikasi beserta definisi kelas

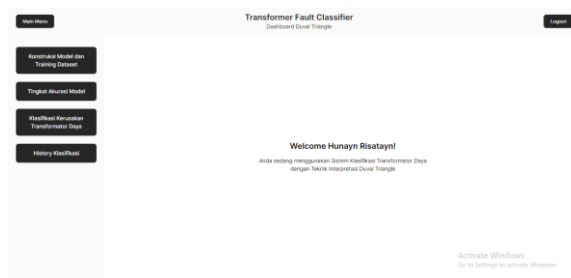
4. *Train_Result*

Entitas *train_result* berfungsi untuk menyimpan data hasil konstruksi dan training model beserta tingkat akurasi. Objek ini menyimpan nama model dalam bentuk string sehingga dapat dilakukan *reload* model ketika klasifikasi data baru dilakukan kembali

4. HASIL DAN PEMBAHASAN

4.1. Implementasi Tampilan Dashboard

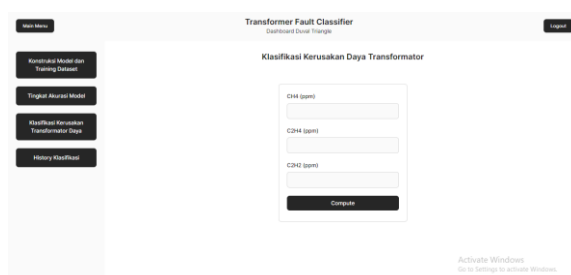
Halaman Dashboard memberikan informasi terkait teknik interpretasi hasil analisis gas terlarut yang sedang digunakan dan identitas user berupa nama yang sedang login dalam sistem sesuai dengan Gambar 5 berikut



Gambar 5. Halaman dashboard

4.2. Implementasi Tampilan Klasifikasi Kerusakan Tranformator Daya

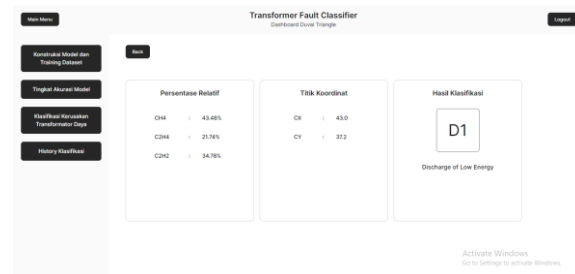
Halaman Klasifikasi Kerusakan Transformator Daya berfungsi untuk menerima input user berupa data konsentrasi gas terlarut meliputi gas metana, etilen dan asetilen yang akan dipetakan kedalam titik koordinat kemudian ditentukan kelas kerusakan transformator daya menggunakan metode sesuai dengan Gambar 6



Gambar 6. Halaman klasifikasi kerusakan transformator daya

4.3. Implementasi Tampilan Hasil Klasifikasi Model

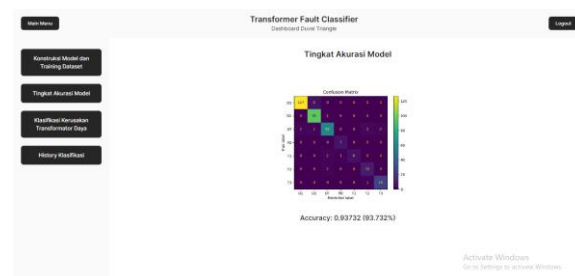
Halaman Hasil Klasifikasi Model berfungsi untuk menampilkan persentase relatif tiap gas dan letak titik koordinat serta hasil klasifikasi kelas kerusakan transformator daya menggunakan model sesuai dengan Gambar 7 berikut



Gambar 7. Halaman hasil klasifikasi model

4.4. Implementasi Tampilan Tingkat Akurasi Model

Halaman Tingkat Akurasi Model berfungsi untuk menampilkan tingkat akurasi dari model yang terbentuk menggunakan confusion matrix dengan membandingkan nilai prediksi dengan nilai actual dataset sesuai dengan Gambar 8 berikut



Gambar 8. Halaman tingkat akurasi model

4.5. Implementasi Tampilan Histori Klasifikasi

Halaman Histori Klasifikasi berfungsi untuk menampilkan seluruh *record* hasil klasifikasi berdasarkan *id* user meliputi persentase relatif gas metana, etilen dan asetilen, letak titik koordinat serta kelas kerusakan transformator daya sesuai dengan Gambar 9 berikut

ID	%CH4	%C2H4	%C2H2	CX	CY	KELAS KERUSAKAN	DETAIL KERUSAKAN	ACTION
1	34.00	43.01	23.00	57.0	210	D2	Discharge of High Energy	Detail
2	30.00	34.84	47.0	42.0	10.0	D2	Discharge of High Energy	Detail
3	43.46	21.76	34.76	43.0	312	D1	Discharge of Low Energy	Detail

Gambar 9. Halaman Histori Klasifikasi

4.6. Pengujian Fitur Login

Fitur login disematkan dalam sistem untuk melakukan filter pengguna sehingga fungsi dan hak yang ditugaskan dapat dijalankan sesuai dengan ketentuan aplikasi. Pengujian fitur tercantum pada tabel 1 berikut

Tabel 1. Tabel Pengujian Fitur Login

Pengujian Fitur Login			
Kasus dan Hasil Uji Data Normal			
Data Input	Harapan	Pengamatan	Hasil
username: hunaynr password: 123456789	Data terinput dalam Text Box Input Field	Data terinput dalam Text Box Input Field	Sesuai
Klik tombol login	Tampil halaman Main Menu	Tampil halaman Main Menu	Sesuai
Kasus dan Hasil Uji Data Tidak Normal			
Data Input	Harapan	Pengamatan	Hasil
username: hunaynr password: hunaynr12345	Data terinput dalam Text Box Input Field	Data terinput dalam Text Box Input Field	Sesuai
Klik tombol login	Muncul notifikasi warning "Please check your login details and try again"	Muncul notifikasi warning "Please check your login details and try again"	Sesuai

4.7. Perhitungan Metode

Dua tahap dalam mekanisme konstruksi model *random forest* diantaranya *bagging* (*bootstrap aggregating*), himpunan *bootstrap* sample dibentuk dari dataset yang ditentukan secara acak dengan konsep *random sampling and replacement*. Data sample dilakukan training untuk membentuk *decision tree*. Seleksi fitur secara acak berperan dalam penentuan fitur pemisah pada *root node*. Subset fitur dipilih secara acak dari fitur dalam dataset dengan jumlah elemen fitur dalam subset berjumlah minimal dua. Dataset berjumlah 1016 records dengan parameter *testing size* = 25% melakukan partisi pada dataset menjadi dua bagian data training = 762 records dan data testing = 254 records. Parameter *n_estimator* menentukan jumlah *decision tree* yang dibangun menggunakan data sample hasil *bagging*. Sejumlah 100 *decision trees* dilakukan konstruksi dengan basis data training secara acak. Pembangunan *decision tree* mengikuti aturan berikut:

Fitur *root node* ditentukan menggunakan *information entropy* dengan persamaan (1). Jumlah record dengan fitur $C_y \leq 1.45$ yang termasuk dalam seluruh tipe kelas kerusakan dan nilai probabilitas

setiap kelas kerusakan dalam rentang nilai fitur tercantum dalam tabel 2 berikut

Tabel 2. Jumlah data dengan fitur $C_y \leq 1.45$ setiap kelas kerusakan dan probabilitas

No.	Kelas Kerusakan	n	Probabilitas
1.	C	71	0.093176
2.	D1	144	0.188976
3.	D2	144	0.188976
4.	O	119	0.156168
5.	PD	46	0.060367
6.	S	158	0.207349
7.	T3-H	80	0.104987
Total		762	

Information entropy

$$= -\left[\begin{aligned} &0.093176 * \frac{\log(0.093176)}{\log(2)} \\ &+ \left(0.188976 * \frac{\log(0.188976)}{\log(2)} \right) \\ &+ \left(0.188976 * \frac{\log(0.188976)}{\log(2)} \right) \\ &+ \left(0.156168 * \frac{\log(0.156168)}{\log(2)} \right) \\ &+ \left(0.060367 * \frac{\log(0.060367)}{\log(2)} \right) \\ &+ \left(0.207349 * \frac{\log(0.207349)}{\log(2)} \right) \\ &+ \left(0.104987 * \frac{\log(0.104987)}{\log(2)} \right) \end{aligned} \right]$$

$$\text{Information entropy} = 2.702401$$

Fitur *internal node* ditentukan menggunakan nilai *information gain* tertinggi dari kandidat fitur sesuai dengan persamaan (2)

Jumlah *record* pada *internal node* dengan fitur $C_x \leq -1.25$ yang termasuk dalam seluruh tipe kelas kerusakan dan nilai probabilitas setiap kelas kerusakan dalam rentang nilai fitur tercantum dalam tabel 3 berikut

Tabel 3. Jumlah data dengan fitur $C_x \leq -1.25$ setiap kelas kerusakan dan probabilitas

No.	Kelas Kerusakan	n	Probabilitas
1.	C	71	0.196676
2.	D1	2	0.00554
3.	D2	92	0.254848
4.	O	116	0.32133
5.	PD	0	0
6.	S	0	0
7.	T3-H	80	0.221607
Total		361	

Information entropy

$$= -\left[\left(0.196676 * \frac{\log(0.196676)}{\log(2)} \right) + \left(0.00554 * \frac{\log(0.00554)}{\log(2)} \right) + \left(0.254848 * \frac{\log(0.254848)}{\log(2)} \right) + \left(0.32133 * \frac{\log(0.32133)}{\log(2)} \right) + \left(0.221607 * \frac{\log(0.221607)}{\log(2)} \right) \right]$$

Information entropy = 2.01364

Nilai *information gain* merupakan selisih atau total pengurangan *entropy* sebelum dilakukan separasi/pemisahan dengan *entropy* setelah dilakukan pemisahan menggunakan kandidat fitur. Nilai tertinggi dari kandidat merupakan fitur terbaik dalam melakukan separasi/pemisahan.

Information Gain

$$= 2.702401 - \left[\left(\frac{71}{361} \right) * 2.01364 \right] + \left(\frac{2}{361} \right) * 2.01364 + \left(\frac{92}{361} \right) * 2.01364 + \left(\frac{116}{361} \right) * 2.01364 + \left(\frac{80}{361} \right) * 2.01364$$

Information Gain = 0.688762

Jumlah *record* pada internal *node* dengan fitur $C_x \leq 0.25$ yang termasuk dalam seluruh tipe kelas kerusakan dan nilai probabilitas setiap kelas kerusakan dalam rentang nilai fitur tercantum dalam tabel 4 berikut

Tabel 4. Jumlah data dengan fitur $C_x \leq 0.25$ setiap kelas kerusakan dan probabilitas

No.	Kelas Kerusakan	n	Probabilitas
1.	C	0	0
2.	D1	142	0.354115
3.	D2	52	0.129676
4.	O	3	0.007481
5.	PD	46	0.114713
6.	S	158	0.394015
7.	T3-H	0	0
Total		401	

Information entropy

$$= -\left[\left(0.354115 * \frac{\log(0.354115)}{\log(2)} \right) + \left(0.129676 * \frac{\log(0.129676)}{\log(2)} \right) + \left(0.007481 * \frac{\log(0.007481)}{\log(2)} \right) + \left(0.114713 * \frac{\log(0.114713)}{\log(2)} \right) + \left(0.394015 * \frac{\log(0.394015)}{\log(2)} \right) \right]$$

Information entropy = 1.8531366

Information Gain

$$= 2.702401 - \left[\left(\frac{142}{401} \right) * 1.8531366 \right] + \left(\frac{52}{401} \right) * 1.8531366 + \left(\frac{3}{401} \right) * 1.8531366 + \left(\frac{46}{401} \right) * 1.8531366 + \left(\frac{158}{401} \right) * 1.8531366$$

Information Gain = 0.8531366

Tingkat akurasi diperoleh dari kombinasi nilai input *testing size* dan *n_estimator* pada sistem klasifikasi perangkat lunak. Kedua nilai tersebut berperan dalam proses pembentukan model yang dapat diukur performa terhadap tugas klasifikasi untuk menghasilkan *confusion matrix*. Komparasi atau perbandingan antara output model bernilai benar sesuai dengan data sampel dengan output tidak tepat klasifikasi membentuk persentase tingkat akurasi model. Untuk mendapatkan performa maksimal dari sistem diperlukan pengujian dari berbagai ukuran parameter. Ragam parameter dan tingkat akurasi yang dihasilkan tercantum dalam tabel 5 berikut

Tabel 5. Ragam parameter dan tingkat akurasi *duval pentagon*

No	Testing Size (%)	n_estimator	Tingkat Akurasi (%)
1.	15	50	96.732
2.	15	75	96.078
3.	15	100	98.693
4.	15	125	94.118
5.	15	150	95.425
6.	20	50	94.608
7.	20	75	96.078
8.	20	100	93.137
9.	20	125	95.098
10.	20	150	96.569
11.	25	50	94.882
12.	25	75	97.244
13.	25	100	94.094

No	Testing Size (%)	$n_{estimator}$	Tingkat Akurasi (%)
14.	25	125	96.063
15.	25	150	95.276
16.	30	50	94.754
17.	30	75	93.115
18.	30	100	96.393
19.	30	125	96.066
20.	30	150	94.098
21.	35	50	96.348
22.	35	75	95.786
23.	35	100	94.944
24.	35	125	94.382
25.	35	150	94.663

Berdasarkan pengujian menggunakan sejumlah nilai parameter diperoleh tingkat akurasi tertinggi dengan gabungan persentase *testing size* dan $n_{estimator}$ secara berurutan sebesar 15 dan 100 menghasilkan tingkat akurasi model sebesar 98.693%. Kompilasi parameter *testing size* dan $n_{estimator}$ untuk mencapai tingkat akurasi maksimal menggunakan teknik interpretasi *duval triangle* tercantum dalam tabel 6 berikut

Tabel 6. Ragam parameter dan tingkat akurasi *duval triangle*

No	Testing Size (%)	$n_{estimator}$	Tingkat Akurasi (%)
1.	15	50	93.839
2.	15	75	92.891
3.	15	100	93.365
4.	15	125	95.261
5.	15	150	91.943
6.	20	50	93.594
7.	20	75	91.815
8.	20	100	92.171
9.	20	125	93.95
10.	20	150	94.662
11.	25	50	92.593
12.	25	75	94.302
13.	25	100	91.738
14.	25	125	96.011
15.	25	150	90.883
16.	30	50	93.824
17.	30	75	94.774
18.	30	100	91.687
19.	30	125	93.587
20.	30	150	94.537
21.	35	50	92.057
22.	35	75	92.668
23.	35	100	92.261
24.	35	125	93.075
25.	35	150	94.094

Berdasarkan pengujian menggunakan teknik interpretasi *duval triangle* dari sejumlah nilai parameter diperoleh tingkat akurasi tertinggi dengan gabungan persentase *testing size* dan $n_{estimator}$ secara berurutan sebesar 25 dan 125 menghasilkan tingkat akurasi model sebesar 96.011%. Merujuk pada pengujian yang telah dilakukan dapat ditarik kesimpulan bahwa besaran tingkat akurasi dipengaruhi

oleh kombinasi attribute yang berperan dalam pembentukan model yaitu nilai persentase *testing size* dan $n_{estimator}$ yang menyatakan jumlah *decision tree* yang terlibat dalam pemetaan kelas kerusakan transformator daya

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang telah dilakukan, maka penulis dapat menyimpulkan bahwa klasifikasi kerusakan transformator daya berdasarkan gas terlarut menggunakan teknik interpretasi *duval triangle* dan metode klasifikasi *random forest* menghasilkan tingkat akurasi terbaik sebesar 96.011% pada kombinasi *testing size* = 25% dan $n_{estimator}$ = 125, sedangkan teknik interpretasi *duval pentagon* menghasilkan tingkat akurasi tertinggi dengan nilai 98.693% pada kombinasi *testing size* = 15% dan $n_{estimator}$ = 100, melalui sudut pandang implementasi sistem dapat beroperasi dari segi fungsionalitas baik mekanisme klasifikasi model maupun utilitas komponen sesuai dengan spesifikasi perencanaan. Mengacu pada analisis yang telah dilakukan, maka penulis dapat memberikan saran untuk penelitian lebih lanjut yakni penerapan tipe metode analisis hasil DGA (*Dissolved Gas Analysis*) lain dalam memberikan interpretasi seperti *Roger Ratio* atau *Doernenburg Ratio Method*, selain itu aplikasi kriteria dalam pengukuran kualitas *purity/kemurnian* lain saat dilakukan pemisahan terhadap suatu fitur dalam rangka optimalisasi tingkat akurasi model seperti *gini index* atau *log loss*

DAFTAR PUSTAKA

- [1] Zope, N., Ali, S. I., Padmanaban, S., Bhaskar, M. S., and Mihet-Popa, L., 2018, February. Analysis of 132kV/33kV 15MVA power transformer dissolved gas using transport-X Kelman Kit through Duval's triangle and Roger's Ratio prediction. In *Industrial Technology (ICIT), IEEE International Conference on* (pp. 1160-1164). IEEE.
- [2] Akbari, A., Setayeshmehr, A., Borsi, H., and Gockenbach, E., 2008, June. A Software Implementation of the Duval Triangle Method. In *Electrical Insulation, Conference Record of the 2008 IEEE International Symposium on* (pp. 124-127). IEEE.
- [3] Martin, D., Lelekakis, N., Wijaya, J., Duval, M., and Saha, T. Investigations Into the Stray Gassing of Oils in the Fault Diagnosis of Transformers. In *Power Delivery, IEEE Transactions on* Oct 2014, vol. 29, no. 5 (pp. 2369-2374). IEEE.
- [4] Pattanadech, N. and Wattakapaiboon, W., 2019, July. Application of Duval Pentagon Compared with Other DGA Interpretation Techniques: Case Studies for Actual Transformer Inspections Including Experience from Power Plants in Thailand. In *Engineering, Applied Sciences and*

- Technology (ICEAST), 2019 5th International Conference on* (pp. 1-4). IEEE.
- [5] Duval, M. and Lamarre, L. The Duval Pentagon- A New Complementary Tool for the Interpretation of Dissolved Gas Analysis in Transformers. *IEEE Electrical Insulation Magazine*, on Dec 2014, vol. 30, no. 6 (pp. 9-12). IEEE.
- [6] Vens, C. and Costa, F., 2011. Random Forest Based Feature Induction. In *Data Mining, 2011 IEEE 11th International Conference on* (pp. 744-753). IEEE.
- [7] Safavian, S. R. and Landgrebe D. A survey of decision tree classifier methodology. In *Systems, Man, and Cybernetics, IEEE Transactions on* 1991 Jun, vol. 21, no. 3 (pp. 660-674). IEEE.
- [8] Guo, Y., Zhou, Y., Hu, X., and Cheng, W., 2019 Nov. Research on Recommendation of Insurance Products Based on Random Forest. In *Machine Learning, Big Data and Business Intelligence (MLBDBI), 2019 International Conference on* (pp. 308-311). IEEE.
- [9] Ho, Y. T., Wu, C.-F., Yang, M.-C., Chen, T.-Y., and Chang, Y.-H., 2019, Aug. Replanting Your Forest: NVM-friendly Bagging Strategy for Random Forest. In *Non-Volatile Memory Systems and Applications Symposium (NVMSA), 2019 IEEE on* (pp. 1-6). IEEE.
- [10] Zhongtao, L. and Hong, W., 2011, May. ID3 Optimization Algorithm Based On Interestingness Gain. In *Communication Software and Networks, 2011 IEEE 3rd International Conference on* (pp. 73-77). IEEE.
- [11] Jearanaitanakij, K., 2005, Dec. Classifying Continuous Data Set by ID3 Algorithm. In *Information Communications & Signal Processing, 2005 5th International Conference on* (pp. 1048-1051). IEEE.