

BAB III

METODE PENELITIAN

3.1 Lokasi dan Waktu Penelitian

3.1.1 Lokasi Penelitian

Penelitian ini dilakukan di PT. PLN (Indonesia Power) ULPL-TA Musi Unit Pelaksana Pembangkitan (UPDK) Bengkulu di Jl. Raya Bengkulu-Curup KM 72, Kel. Ujan Mas Atas, Kab. Kepahiang, Provinsi Bengkulu.

3.1.2 Waktu Penelitian

Penelitian ini diperkirakan akan dilaksanakan pada bulan Juni - Agustus 2025. Penelitian ini menggunakan data hasil DGA yang dimiliki oleh PT. PLN Indonesia Power ULPL-TA Musi Unit Pelaksana Pembangkitan (UPDK).

3.2 Objek Penelitian

Objek penelitian yang digunakan ialah minyak transformator step up 11kV/150kV dari PT. PLN Indonesia Power ULPL-TA Musi. Transformator yang akan dianalisis yaitu 3 unit transformator.

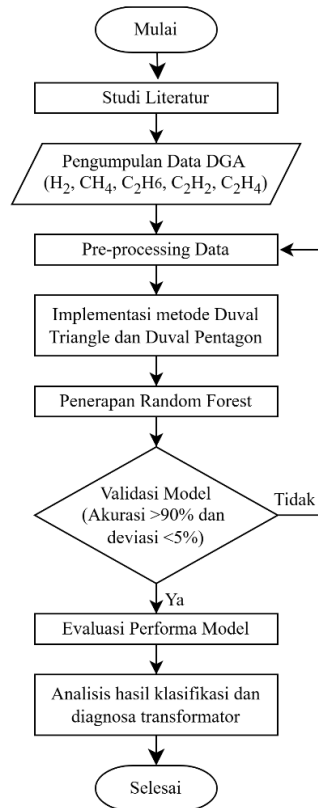
3.3 Alat dan Bahan

Penelitian ini membutuhkan sejumlah peralatan yang terbagi menjadi dua kategori utama, yaitu perangkat keras (*hardware*) dan perangkat lunak (*software*). Perangkat lunak yang digunakan mencakup sistem operasi serta aplikasi atau program pendukung yang dibutuhkan selama proses penelitian yaitu *Visual Studio Code* untuk memproses algoritma *Random Forest*. Sementara itu, perangkat keras yang diperlukan meliputi laptop dan berbagai perangkat tambahan lain yang mendukung kelancaran pelaksanaan penelitian.

3.4 Diagram Alir Penelitian

Proses penelitian DGA menggunakan metode *Random Forest* ini memiliki diagram alir penelitian agar penelitian dapat dijalankan terarah dan terstruktur.

Diagram alir penelitian DGA menggunakan *Random Forest* dapat dilihat pada Gambar 3.1



Gambar 3. 1 Diagram Alir Penelitian

3.5 Tahapan Penelitian

Tahapan penelitian merupakan rangkaian proses yang dilaksanakan secara bertahap dalam menjalankan suatu penelitian. Proses ini dilakukan secara terstruktur, sistematis, logis, dan relevan dengan kondisi aktual guna mencapai tujuan yang telah ditetapkan dalam penelitian. Adapun urutan tahapan dalam penelitian ini dijelaskan sebagai berikut.

3.4.1 Studi Literatur

Penelitian ini diawali dengan melakukan studi literatur guna memperoleh landasan teori yang kuat. Informasi dikumpulkan dari berbagai sumber, seperti jurnal internasional dan nasional, artikel ilmiah yang relevan, serta buku-buku yang mendukung topik penelitian. Selain itu, pengumpulan data juga mengacu pada studi

literatur yang dilakukan di PLTA Musi. Data yang dikumpulkan berupa nilai kandungan gas hasil dari *Dissolved Gas Analysis* (DGA) yang selanjutnya akan dianalisis menggunakan metode *Random Forest*. Studi literatur ini bertujuan untuk memastikan bahwa penelitian berjalan sesuai kaidah ilmiah dan berdasarkan teori yang telah terbukti secara empiris.

3.4.2 Pengumpulan Data

Tahapan ini bertujuan untuk mengumpulkan informasi atau data yang berfungsi sebagai bahan pendukung dalam pelaksanaan penelitian. Jenis data yang digunakan adalah data sekunder, yaitu data yang telah tersedia sebelumnya dan diperoleh dari sumber yang terpercaya. Dalam penelitian ini, data yang dikumpulkan berupa informasi mengenai kandungan gas-gas terlarut dalam minyak isolasi, seperti CO, H₂, CH₄, C₂H₆, C₂H₂, dan C₂H₄. Data yang digunakan yaitu 90 data DGA dari PLN dan data tambahan dari jurnal pada penelitian sebelumnya.

3.4.3 Preprocessing Data

Preprocessing data merupakan tahap kritis untuk memastikan kualitas dataset sebelum digunakan dalam pemodelan. Selanjutnya, normalisasi fitur dilakukan untuk mengubah rentang nilai gas ke skala [0, 1] agar tidak ada fitur yang mendominasi akibat perbedaan satuan. Data kemudian dibagi menjadi dua subset, yaitu 80% untuk pelatihan (*training*) dan 20% untuk pengujian (*testing*), dengan *stratified sampling* untuk menjaga proporsi kelas gangguan. Label klasifikasi diubah menjadi nilai numerik agar kompatibel dengan algoritma *Random Forest* seperti pada Tabel 3.1 dan 3.2.

Tabel 3. 1 Inisialisasi *Output Duval Triangle* Pada *Random Forest*

Zone	Fault	Output
PD	<i>Partial Discharge</i>	1
D1	<i>Discharges of low energy</i>	2
D2	<i>Discharges of high energy</i>	3
T1	<i>Thermal fault < 300°C</i>	4

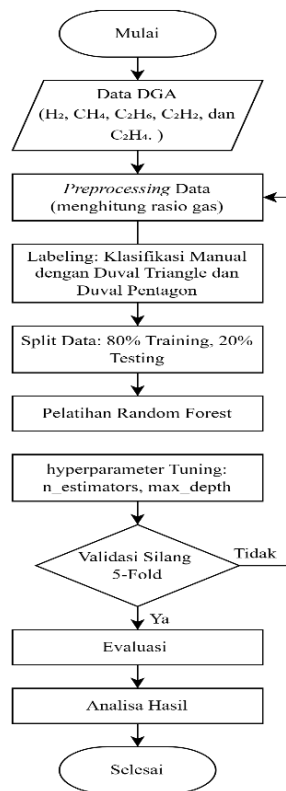
<i>Zone</i>	<i>Fault</i>	<i>Output</i>
T2	<i>Thermal fault 300°C - 700°C</i>	5
T3	<i>Thermal fault > 700°C</i>	6
DT	<i>Thermal fault and discharge</i>	7

Tabel 3. 2 Inisialisasi *Output Duval Pentagon* Pada *Random Forest*

<i>Zone</i>	<i>Fault</i>	<i>Output</i>
PD	<i>Partial Discharge (corona discharge)</i>	1
D1	<i>Discharges of low energi (spark type)</i>	2
D2	<i>Discharges of high energy</i>	3
T1	<i>Thermal fault < 300°C</i>	4
T2	<i>Thermal fault 300°C - 700°C</i>	5
T3	<i>Thermal fault > 700°C</i>	6
S	<i>Stray gassing</i>	7

3.4.4 Penerapan *Random Forest*

Algoritma *Random Forest* dipilih karena kemampuannya menangani data non-linear dan ketahanannya terhadap *overfitting*. Model dibangun menggunakan *RandomForestClassifier* dari pustaka *Scikit-learn* dengan parameter awal: $n_estimators=100$ (jumlah pohon keputusan), $max_depth=10$ (kedalaman maksimum pohon), dan $criterion='gini'$ (ukuran pemisahan node). *Hyperparameter tuning* kemudian dilakukan dengan *GridSearchCV* untuk mengoptimasi nilai $n_estimators$ (50–200) dan max_depth (5–15) berdasarkan akurasi validasi silang. Fitur penting (*feature importance*) dievaluasi untuk mengidentifikasi gas yang paling berpengaruh dalam klasifikasi. Model dilatih dengan data *training* dan hasil klasifikasi dari metode Duval (Triangle dan Pentagon) digunakan sebagai label referensi. Berikut merupakan diagram alir proses dari program *Random Forest* yang dirancang pada analisis DGA menggunakan *Duval Triangle* dan *Duval Pentagon* dapat dilihat pada Gambar 3.2



Gambar 3. 2 Diagram Alir Proses Dari Program *Random Forest* Dengan *Duval Triangle* Dan *Duval Pentagon*

3.4.5 Validasi Model

Validasi model bertujuan untuk memastikan keandalan prediksi sebelum diterapkan ke data baru. Pada penelitian ini, digunakan teknik *5-fold cross-validation* untuk membagi data *training* menjadi 5 subset, dimana setiap subset bergantian menjadi data validasi. Metode ini mengurangi risiko *overfitting* dan memberikan estimasi performa yang lebih stabil. Selain itu, *out-of-bag error* (OOB) dari *Random Forest* juga dihitung sebagai metrik validasi internal. Hasil validasi silang menunjukkan konsistensi model dengan akurasi rata-rata >90% dan deviasi standar <5%, mengindikasikan bahwa model cukup robust terhadap variasi data.

3.4.6 Evaluasi Model

Evaluasi model *Random Forest* dilakukan untuk mengetahui sejauh mana akurasi dan ketepatan klasifikasi gangguan transformator berdasarkan data DGA. Model diuji menggunakan data uji dan dinilai dengan metrik seperti *accuracy*,

precision, *recall*, dan *F1-score*. *Accuracy* mencerminkan jumlah prediksi benar secara keseluruhan, sementara *precision* dan *recall* membantu menilai performa tiap kelas gangguan. *F1-score* digunakan sebagai ukuran keseimbangan antara *precision* dan *recall*. Evaluasi dilakukan untuk masing-masing model *Duval Triangle* dan *Duval Pentagon* secara terpisah, guna membandingkan efektivitas kedua pendekatan.

3.4.7 Analisa Hasil Klasifikasi

Model Random Forest menghasilkan prediksi jenis gangguan transformator berdasarkan pola gas dari data DGA. Hasil klasifikasi dibandingkan dengan label aktual untuk menilai akurasi model. Analisis dilakukan secara terpisah untuk model *Duval Triangle* dan *Duval Pentagon* guna menilai kekuatan masing-masing pendekatan. Hasil ini mendukung diagnosa transformator secara otomatis dan membantu teknisi menentukan langkah perbaikan yang tepat.

Hasil analisa dari klasifikasi ini diinterpretasikan secara spesifik terhadap kerusakan pada minyak transformator. Pola gas yang terbentuk mencerminkan adanya degradasi atau gangguan tertentu yang terjadi di dalam minyak, seperti pemanasan berlebih, pelepasan parsial, atau arcing. Dengan demikian, model ini secara langsung membantu dalam mengidentifikasi kondisi kerusakan minyak dan mendeteksi potensi kegagalan dini sebelum kerusakan menyebar ke komponen lain.

3.6 Matriks Evaluasi

Matriks evaluasi digunakan untuk menilai performa model dalam mengklasifikasikan gangguan transformator. *Confusion matrix* menunjukkan distribusi hasil klasifikasi antar kelas dan membantu mengidentifikasi kesalahan prediksi. Matrik lain yang digunakan adalah *accuracy* untuk mengukur prediksi yang benar secara keseluruhan, serta *precision*, *recall*, dan *F1-score* untuk menilai performa tiap kelas. Hasil dari evaluasi ini menjadi dasar perbandingan antara model berdasarkan *Duval Triangle* dan *Duval Pentagon*.