

Penggunaan Algoritma Random Forest untuk Klasifikasi berbasis Kinerja Efisiensi Energi pada Sistem Pembangkit Daya

Ghiffari Awliya Muhammad Ashfania^{a,*}, Toni Prahasto^a, Achmad Widodo^a, Tarwaji Warsokusumo^b

^aDepartemen Teknik Mesin, Fakultas Teknik, Universitas Diponegoro

Jl. Prof. Soedarto, Tembalang, Kec. Tembalang, Kota Semarang, Jawa Tengah 50275

^bP.T. Indonesia Power, Jakarta, Indonesia 12950

Jl. Gatot Subroto No.Kav.18, Kuningan Tim., Kecamatan Setiabudi, Kota Jakarta Selatan, Daerah Khusus Ibukota Jakarta 12950

*E-mail: ghiffari.awliya@gmail.com

Abstract

Energy is one of the main performance parameters in the generation system apart from the reliability, availability, maintenance and safety (RAMS) parameters. Energy Efficiency Indicators (EEI) have a very significant role in cost efficiency related to the Cost of Supply of electric power and the reduction of CO₂ emissions related to the environment combined with the application of New and Renewable Energy based generators. Other power plant operation data that are monitored according to available sensors include: main steam (temperature, flow, pressure, spray water flow), reheat steam (temperature, flow, pressure, spray water flow), feed water (temperature, pressure, flow), exhaust gas (O₂ content, temperature), condensate (vacuum pressure, temperature), coal (calorific value, flow), etc. So far, in equipment failure diagnosis, decision making is only based on features based on equipment performance. This paper aims to find new alternatives by using features based on Energy Efficiency Performance indicators as a basis for diagnosing equipment failure or system function degradation for maintenance decision making. Diagnosis process uses the Random Forest Classifier (RFC) model which provide information on which features have anomaly so that the equipment or system associated with these features can be determined. In addition, by using the RFC model, Feature Importance can be obtained which can be used to determine the scope of maintenance work to be carried out. The results of the case study in this study indicate that by using the Net Plant Heat Rate (NPHR) indicator label feature and the predictor features of the Top 10 Controlled Variables, it is known that the features that have degraded functions related to energy efficiency performance and have adjusted their priorities are: FFW Temperature, Main APH Gas out Temperature, and O₂ Content in flue Gas.

Keywords: Random forest, NPHR, Controlled variable operation indicator, Efficiency energy

Abstrak

Efisiensi energi adalah salah satu parameter kinerja utama dalam sistem pembangkit daya selain dari parameter keandalan, ketersediaan, pemeliharaan, dan keselamatan (RAMS). Indikator Efisiensi Energi (EEI) mempunyai peran yang sangat signifikan dalam *cost efficiency* terkait Biaya Pokok Penyediaan (BPP) tenaga listrik dan penurunan emisi CO₂ yang terkait dengan lingkungan hidup yang dikombinasikan dengan aplikasi pembangkit berbasis Energi Baru dan Terbarukan (EBT). Data operasi pembangkit lainnya yang banyak dimonitor sesuai dengan sensor yang tersedia, antara lain: *main steam* (temperature, flow, pressure, spray water flow), *reheat steam* (temperature, flow, pressure, spray water flow), *feed water* (temperature, pressure, flow), *exhaust gas* (O₂ content, temperature), *condensate* (vacuum pressure, temperature), *coal* (calory value, flow), dll. Selama ini dalam diagnosis kegagalan peralatan untuk pengambilan keputusan pemeliharaan hanya didasarkan pada fitur-fitur berbasis kinerja keandalan peralatan. Penelitian dalam paper ini bertujuan untuk mencari alternatif baru dengan penggunaan fitur-fitur berbasis *indicator* Kinerja Efisiensi energi sebagai basis dalam diagnosis kegagalan peralatan atau degradasi fungsi sistem untuk pengambilan keputusan pemeliharaan. Proses diagnosis menggunakan model *Random Forest Classifier* (RFC) yang memberikan informasi fitur-fitur mana yang mengalami anomali sehingga dapat ditentukan peralatan atau sistem yang berhubungan dengan fitur tersebut. Selain itu juga dengan menggunakan model RFC dapat diperoleh *Feature Importance* yang dapat dimanfaatkan untuk menentukan prioritas lingkup pekerjaan pemeliharaan yang akan dilakukan. Hasil kajian studi kasus dalam penelitian ini menunjukkan bahwa dengan menggunakan fitur label indikator *Net Plant Heat Rate* (NPHR) dan fitur-fitur prediktor dari Top 10 Variabel Terkendali, diketahui fitur-fitur yang mengalami degradasi fungsi operasi terkait kinerja efisiensi energi dan sudah disesuaikan prioritasnya adalah: *Final Feed Water Temperatur*, *Main Air Preheater Gas Out Temperature*, dan *O₂ Contents in Flue Gas*.

Kata kunci: random forest, NPHR, Indikator operasi variabel terkendali

1. Pendahuluan

Penggunaan bahan bakar fosil yang diterapkan pada industri pembangkit listrik dewasa ini telah mengalami penurunan. Pembangkit data yang menggunakan bahan bakar fosil sebagai energi primernya telah mendorong dalam peningkatan emisi gas rumah kaca, khususnya gas karbondioksida (CO₂) [1]. Meningkatnya jumlah CO₂ yang sangat tajam memberikan dampak yang cukup besar terhadap perubahan iklim. Untuk meningkatkan efektivitas dalam penggunaan energi pada era ini adalah dengan menggunakan teknik analisis data dalam menghasilkan *self-diagnosis fault*. Penggunaan *self-diagnosis fault* telah diterapkan dengan menggunakan berbagai manajemen index efisiensi energi [2]. Penelitian ini bertujuan untuk mengembangkan Sistem Diagnostik gangguan berbasis efisiensi energi menggunakan data sistem kendali terdistribusi (DCS) dengan Random Forest Classifier (RFC) [3,4] dan data real-time dari pembangkit listrik tenaga uap berbahan bakar batubara dengan kelas daya 400 MW. Output dari penelitian ini dapat digunakan untuk membuat keputusan mengenai ruang lingkup peralatan yang akan dipelihara.

Manajemen pemeliharaan konvensional yang dipakai selama ini masih menggunakan indikator *Reliability, Availability, Maintainability & Safety* (RAMS) sebagai basis untuk monitoring dan pengambilan keputusan. Indikator ini sifatnya berhubungan dengan kondisi kesehatan aset fisik mesin industri yang terkait dengan modus kegagalan fisiknya. Beberapa indikator kinerja RAMS secara praktis yang diaplikasikan untuk industri ketenagalistrikan adalah: *Equivalent Availability Factor (EAF)*, *Sudden Outage Factor (SdOF)*, *Equivalent Forced Outage Rate (EFOR)*, dan *Capacity Factor (CF)*. *Data-driven diagnostic and prognostic* berbasis data kinerja RAMS juga sudah banyak dilakukan dengan berbagai variasi algoritma yang sudah cukup maju dalam aplikasinya. Indikator kinerja *lagging* dari aset mesin industri yang juga selalu dimonitor adalah *Energy Efficiency Indicators (EEI)*, seperti: *Specific Fuel Consumption (SFC, kg/kWh)*, *Net Plant Heat Rate (NPHR, kCal/kWh)*, *Thermal Efficiency (η , %)* dan *Fuel Cost (Rp/kWh)*. Selanjutnya akan dibahas secara spesifik tentang pemanfaatan data indikator kinerja dalam pengambilan keputusan pemeliharaan dengan menggunakan algoritma *data-driven diagnostic & prognostic*.

1.1 Indikator Efisiensi Energi

Meningkatkan efisiensi energi berarti meningkatkan efektivitas penggunaan energi dengan menghasilkan output yang sama. Efisiensi energi tidak selalu berdampak pada emisi gas rumah kaca. Kondisi ini disesuaikan dari energi primer pada sistem tersebut. Jika energi primer pada sistem berupa bahan bakar fosil, dengan meningkatkan efektivitas penggunaan bahan bakar tersebut dapat mengurangi emisi gas rumah kaca, berbeda dengan energi primer seperti listrik yang mana tidak begitu mengurangi dampak emisi. Penilaian efisiensi energi (EE) dapat digunakan untuk pengambilan keputusan tentang strategi dan tindakan prioritas untuk mengurangi konsumsi energi, permintaan energi, dan masalah lingkungan. Untuk penilaian ini, EE diekspresikan dengan menggunakan lebih sedikit energi untuk menghasilkan jumlah layanan atau output yang berguna yang sama. Dengan demikian, persamaan EE dirumuskan dalam Persamaan 1.

$$EE = \frac{\text{work output}}{\text{energy input}} \quad (1)$$

Empat kategori EEI telah dikembangkan dalam berbagai penelitian seperti indikator termodinamika [5], indikator termodinamika-fisika [6], indikator termodinamika-ekonomika [7], dan indikator ekonomi atau keuangan [8]. Dalam perhitungan efisiensi energi, terdapat parameter lain yang digunakan oleh industri pembangkit yang disebut dengan *Net Plant Heat Rate (NPHR)*. NPHR adalah istilah yang digunakan oleh industri pembangkit yang juga untuk menentukan efisiensi termal pada suatu pembangkit. Nilai NPHR ditentukan dengan persamaan 2.

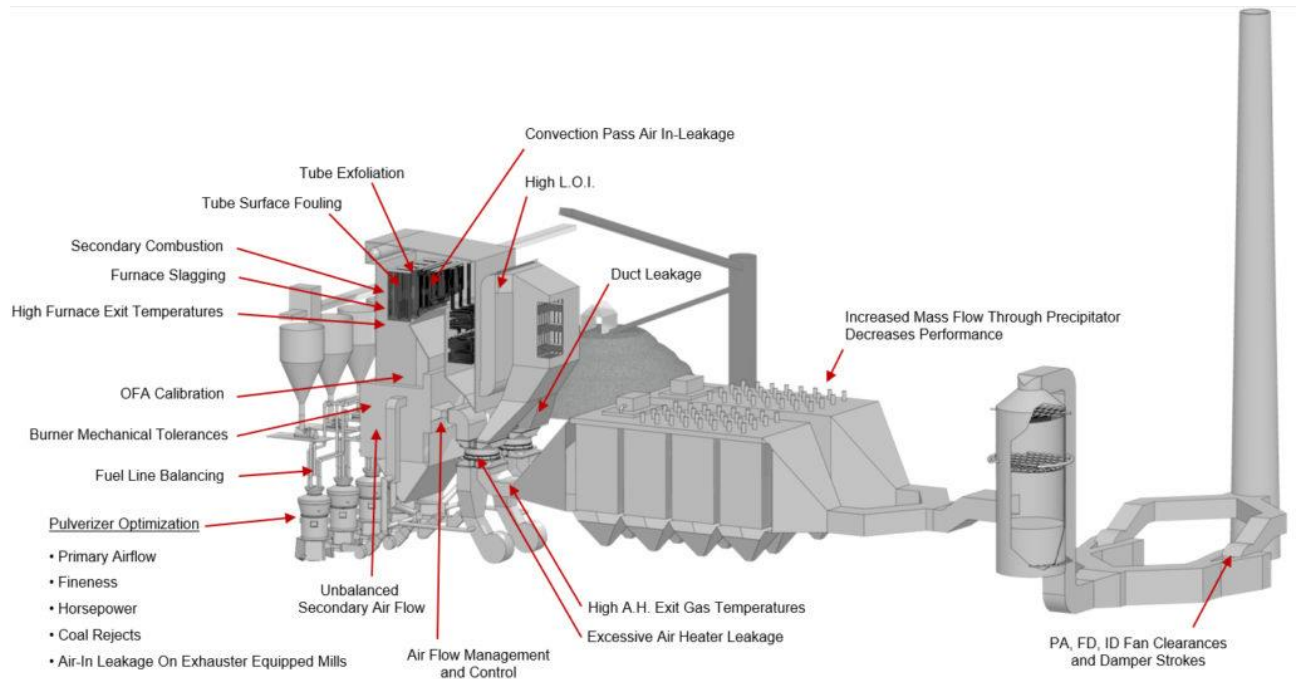
$$NPHR \left(\frac{\text{kcal}}{\text{kWh}} \right) = \frac{m_{\text{coal}} \times HHV}{NGO} \quad (2)$$

Yang mana, m_{coal} adalah Laju alir batu bara per jam (kg/h), HHV adalah Nilai kalori batu bara gross (kcal/kg), NGO adalah *Net Generator Output / Net Load*.

Selain dari indikator efisiensi energi, terdapat indikator operasi lain yang digunakan untuk meningkatkan efisiensi energi. EPRI telah menetapkan 22 faktor laju panas yang dapat dikontrol. Dari 22 faktor tersebut diperoleh 10 variabel terkendali untuk operasi dan pemeliharaan boiler yang diterapkan oleh Storm Tech Company dan digabungkan dengan referensi aktual dari pembangkit listrik terpilih [9], yaitu: (1) *main steam temperature*, (2) *hot reheat steam temperature*, (3) *main steam water spray flow*, (4) *hot reheat steam water spray flow*, (5) *main steam pressure*, (6) *main air preheater gas out temperature*, (7) *final feed water temperature*, (8) *LP turbine exhaust vacuum pressure*, (9) *condenser make up water flow*, dan (10) *O₂ content in flue gas out*.

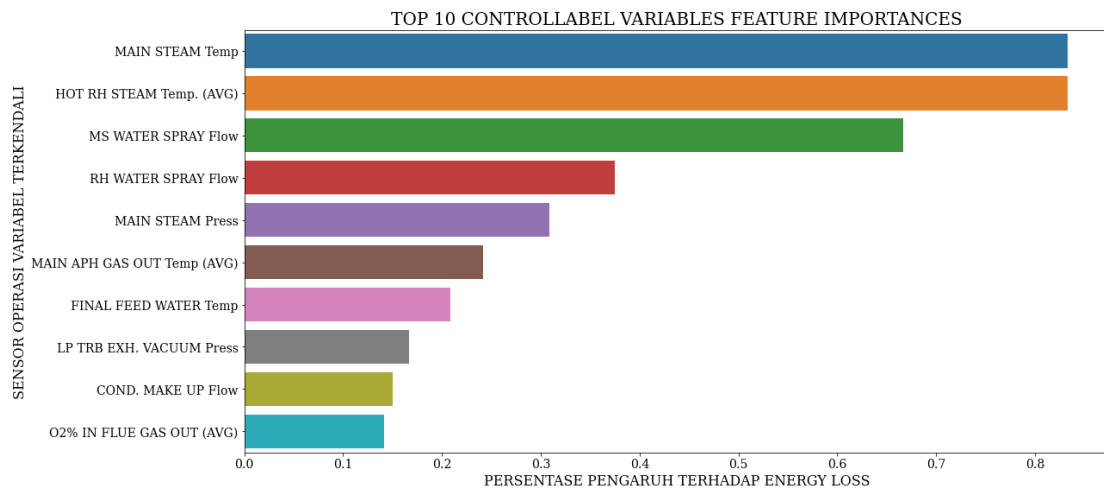
Sepuluh sensor teratas variabel terkendali ini dipilih berdasarkan data kejadian yang sensornya memiliki dampak tinggi pada efisiensi termal. parameter yang terjadi masalah akan dilihat dari seberapa besar rugi daya yang terjadi dan akan dipilih parameter mana yang paling besar pengaruhnya terhadap efisiensi termal. Berdasarkan pengalaman

penerapan 10 variabel referensi Storm yang telah diterapkan, rata-rata unit PLTU berpotensi meningkatkan heat rate sekitar 300 hingga 500 Btu per kWh [9]. Gambar 1 menunjukkan 22 sensor variable terkendali laju panas.



Gambar 1. Diagram 22 controllabel heat rate variables pada boiler [9]

Dengan menggunakan Analisis Pareto untuk memprioritaskan variabel-variabel yang paling berpengaruh atau berpengaruh signifikan terhadap indikator kinerja utama (efisiensi termal atau NPHR), maka dipilih 10 variabel terkendali yang paling berpengaruh secara signifikan terhadap pengendalian kerugian. Gambar 2 menunjukkan bagaimana analisis Pareto memprioritaskan parameter mana yang akan dievaluasi untuk diagnosis.



Gambar 2. Analisis pareto untuk memprioritaskan 10 variabel terkendali.

1.2 Machine Learning: Random Forest Classifier

Machine learning (ML) merupakan sebuah algoritma mesin pembelajaran yang dapat mencari korelasi dari karakteristik data yang diberikan. ML membutuhkan data latih (*training data*) yang cukup dalam membangun sebuah model algoritma cerdas dan perlu diverifikasi akurasi dengan menggunakan data pengujian (*testing data*). Selanjutnya model algoritma ML ini secara mandiri akan mampu memberikan prediksi berupa hasil regresi dan klasifikasi atau klastering yang dapat digunakan untuk proses diagnosis dan prognosis. *Decision Tree* atau pohon keputusan adalah algoritma yang memisahkan sebuah kelompok data dengan menggunakan struktur pohon. Konsep dari pohon keputusan ini adalah membuat pemisahan Data Berdasarkan kondisi yang mana kondisi tersebut dibuat dalam bentuk cabang.

Algoritma *decision tree* mempunyai beberapa algoritma, yaitu ID3 berdasarkan nilai *entropy* dan CART berdasarkan nilai *gini*. CART merupakan algoritma *decision tree* selain ID3 [10]. Pada tahun 1984, sekelompok ahli statistik menerbitkan buku *Classification and Regression Trees* atau CART [11], yang menjelaskan generasi dari *decision tree* biner. Rumus dari pencarian nilai *impurity* pada algoritma CART ditunjukkan pada Persamaan 3.

$$Gini(D) = 1 - \sum_{i=1}^m P_i^2 \tag{3}$$

Yang mana:

P_i : Nilai peluang dari sebuah nilai tuple D pada suatu kelas
 m : Jumlah label kelas

Indeks *Gini* mempertimbangkan pemisahan biner untuk setiap atribut. Untuk mempertimbangkan pemisahan biner, diperlukan perhitungan jumlah bobot dari *impurity* dari setiap partisi yang dihasilkan. Misalnya, jika biner dipecah pada partisi A, D menjadi D_1 dan D_2 , maka nilai indeks *gini* dari D yang diberikan partisi tersebut dapat ditentukan. Nilai indeks *gini* ditunjukkan pada persamaan 4.

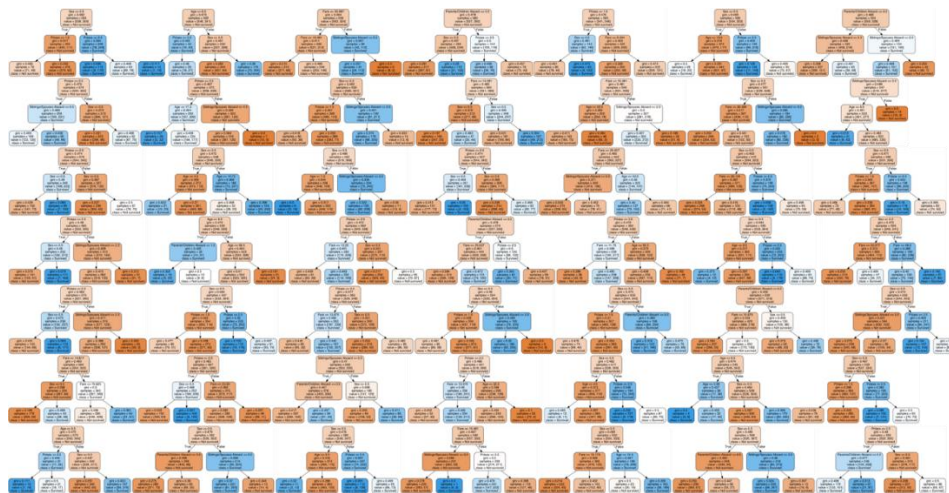
$$Gini_A(D) = \frac{|D_1|}{|D|} Gini(D_1) + \frac{|D_2|}{|D|} Gini(D_2) \tag{4}$$

Impurity reduction yang timbul dari pemisahan biner pada atribut bernilai diskrit atau bernilai kontinu A ditunjukkan pada persamaan 5.

$$\Delta Gini(A) = Gini(D) - Gini_A(D) \tag{5}$$

Decision tree adalah algoritma yang praktis. Model yang menggunakan algoritma ini mudah digunakan, mudah dikembangkan, mudah diinterpretasikan. Tetapi hampir dari setiap penelitian tidak diimplementasikan. *Decision tree* mempunyai aspek yang mana dapat menghalangi model untuk menjadi alat ideal dalam *predictive learning* yaitu *inaccuracy* [12]. Dengan kata lain, algoritma ini cocok untuk mengelompokkan data yang digunakan dalam pembuatan model tetapi tidak fleksibel dalam mengklasifikasikan sampel baru.

Random forest adalah metode *ensemble* dari *decision tree*. *Random forest* menggabungkan kesederhanaan *decision trees* dengan fleksibel menghasilkan peningkatan besar pada akurasi. *Random forest* menggunakan teknik *Bootstrap aggregating (Bagging)* dengan membuat dataset *bootstrapped* untuk mengurangi varians dalam kumpulan *noisy* data dan *aggregating* dengan melakukan voting pada hasil terbanyak. Dataset *bootstrapped* akan melatih setiap model *decision trees* secara paralel. Hasil yang bervariasi dari tiap model *decision trees* akan dilakukan voting dan jumlah terbanyak akan menjadi hasil akhir. Gambar 3. menunjukkan ketergantungan *random forest* terhadap tiap *decision trees*.



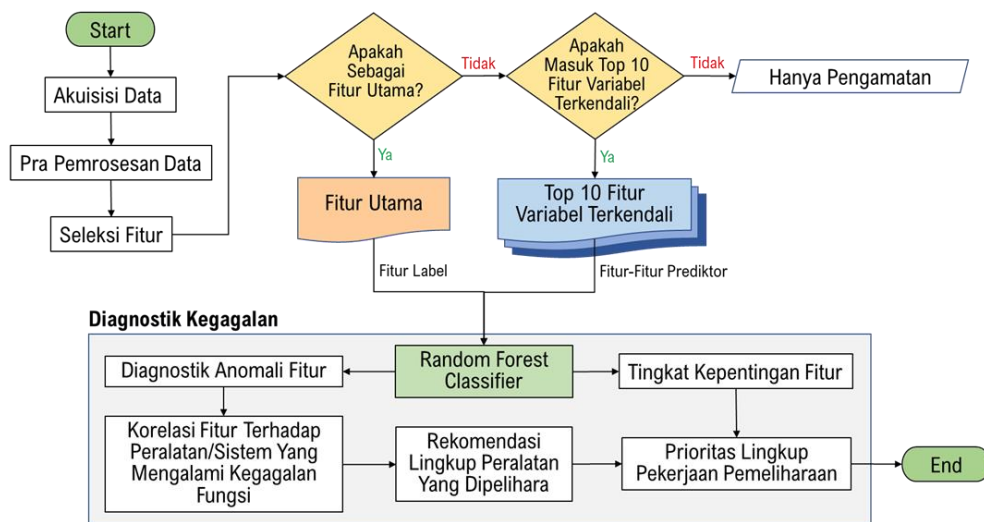
Gambar 3. *Random forest* dengan berbagai model *decision trees* [13].

Dari setiap pohon keputusan yang menjadi hutan bebas akan menjadi acuan dalam penentuan label. Hasil yang bervariasi dari tiap model *decision trees* akan dilakukan voting dan jumlah terbanyak akan menjadi hasil akhir dari prediksi label.

2. Metodologi Penelitian

Langkah awal dalam pembuatan model diagnostik, terlebih dahulu ditentukan indikator yang menjadi dasar terjadinya suatu anomali fungsional dalam pengoperasian pembangkit. Indikator sensor operasi tersebut akan menjadi data prediktor untuk menentukan kondisi nilai NPHR. NPHR akan diatur *threshold* untuk menentukan kondisi NPHR. *Threshold* diatur dengan nilai sebesar 2650, yang mana jika nilai NPHR kurang dari 2650 dianggap baik dan jika nilai NPHR lebih dari 2650 dianggap buruk. Data prediktor yakni 10 sensor operasi terkendali teratas akan dipelajari oleh model RFC guna untuk memprediksi label baik buruknya nilai NPHR. Model RFC dengan akurasi terbaik akan memberikan keluaran sensor mana saja yang paling mempengaruhi label NPHR yang disebut juga sebagai *feature importances*. *Feature importances* akan memberikan persentase dari setiap sensor yang mempengaruhi nilai NPHR dan dari persentase tersebut diambil yang paling tinggi yang menunjukkan adanya degradasi performa pada sensor tersebut yang menyebabkan nilai NPHR naik.

Indikasi anomali tersebut dapat ditelusuri dari hasil trending data terhadap waktu dimana batas yang diijinkan telah ditentukan ketika data melebihi ambang batas. Indikasi anomali akan menjadi validasi dari sensor operasi variabel terkendali. Peralatan sistem pembangkit listrik yang mengalami penurunan fungsi operasi akan mengakibatkan peningkatan konsumsi energi yang kemudian secara bersama-sama berkontribusi pada peningkatan nilai indikator kinerja NPHR. Terjadinya penurunan fungsi operasi peralatan dapat dideteksi dari data indikator operasi masing-masing peralatan berdasarkan *baselinevalue*. Flowchart penelitian ditunjukkan pada Gambar 4.



Gambar 4. Flowchart penelitian *fault* efisiensi energi dengan RFC

3. Hasil dan Pembahasan

3.1 Diagnostik berbasis Data

Diagnosis pada Data A dengan label NPHR menggunakan algoritma klasifikasi *random forest classifier*. Sensor NPHR dikelompokkan menjadi label Good dan Bad dengan mengatur nilai *threshold* sebesar 2650. Nilai NPHR yang melebihi *threshold* akan dikategorikan menjadi label Bad dan nilai NPHR yang kurang dari *threshold* akan dikategorikan menjadi label Good. Dengan data 10 sensor operasi variable terkendali, data tersebut akan menjadi prediktor untuk melakukan klasifikasi label Bad dan Good yang sudah ditentukan. Klasifikasi menggunakan *hyperparameter tuning* pada RFC. Parameter yang digunakan untuk *hyperparameter tuning* pada klasifikasi ini ditunjukkan pada Tabel 1. Tujuan dari *hyperparameter* ini adalah untuk mengurangi kemungkinan terjadinya model yang terlalu *overfit* dalam memprediksi nilai NPHR

Tabel 1. Hyperparameter pada RFC

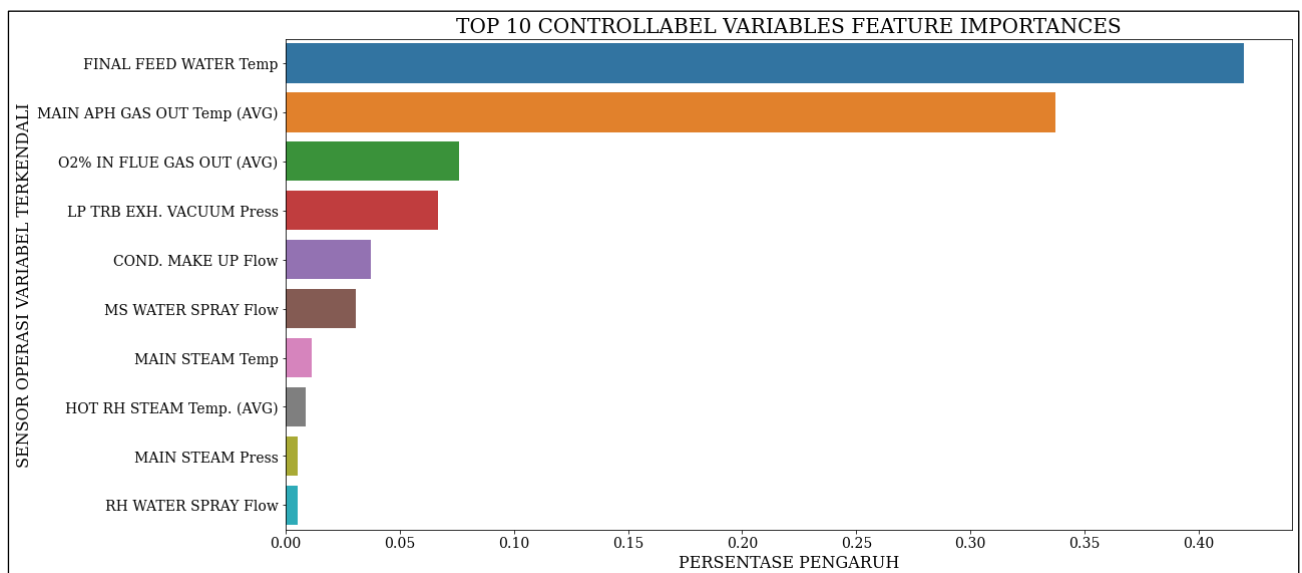
Machine Learning	Hyperparameter
<i>Random Forest Classifier</i>	Max Depth: 4,5,6 Criterion: 'gini' Estimator: 50,100

Dari berbagai parameter dengan *hyperparameter* yang ditentukan, didapatkan model dengan akurasi terbaik adalah model RFC dengan parameter *max_depth* = 6 dan *estimator* sebesar 50 pohon keputusan. Nilai akurasi yang didapatkan pada model terbaik ini adalah sebesar 92,6%. Skor ini cukup besar sehingga cukup baik sebagai model klasifikasi pengelompokan nilai NPHR. Skor dari model RFC ditunjukkan pada Gambar 5.

Random Forest Classifier			
Confusion Marix		Data Prediction	
		Kelas Normal	Kelas Gagal
Data Test	Kelas Normal	243	28
	Kelas Gagal	9	220
kelas	precision	recall	f1-score
Normal	96.43	89.67	92.93
Gagal	88.71	96.07	92.24
accuracy	92.60		

Gambar 5. Skor model RFC dengan parameter terbaik

Dari model RFC pada Gambar 5 menunjukkan selain nilai akurasi yang cukup tinggi, nilai presisi, recall dan skor-F1 dari model tersebut juga sangat tinggi, mendekati nilai 90%. Sehingga persiapan data yang digunakan untuk melatih model RFC cukup baik. Model RFC mempunyai keluaran feature importances. Feature importances menunjukkan seberapa pengaruh sensor variabel terkendali dalam pengelompokan nilai NPHR yang buruk maupun baik. Diagram batang dari feature importances ditunjukkan pada Gambar 6.



Gambar 6. Feature importances dari model RFC terhadap data indikator variabel terkendali

Pada diagram batang ini dapat diasumsikan bahwa sensor yang paling memberikan pengaruh terhadap pengelompokan nilai NPHR adalah Temperatur final feed-water, Temperatur Main APH Gas Out dan dilanjutkan dengan konten oksigen pada gas buang. Sensor yang diambil adalah 3 teratas sehingga peralatan yang perlu dilakukan pemeliharaan adalah 3 sensor tersebut.

3.2 Validasi Dan Penentuan Lingkup Pemeliharaan

Sensor variabel terkendali memiliki nilai baseline yang menjadi acuan dalam pembuatan sebuah threshold yang menentukan status dari kondisi sensor tersebut. Nilai baseline direkomendasikan oleh SME dan menjadi acuan utama dalam penentuan kondisi peralatan pada pembangkit. Di saat sensor tersebut melewati nilai threshold, maka status dari sensor tersebut dalam keadaan buruk. Sebaliknya, jika masih dalam threshold, maka status dari sensor tersebut dalam keadaan baik. Jumlah dari baik buruknya sensor tersebut akan menjadi acuan dalam penentuan status Normal atau Kritis. Cakupan data yang akan didiagnostik berdasarkan nilai base line value adalah nilai NPHR yang buruk atau melebihi 2650. Tabel 3 menunjukkan status dari 10 sensor variabel terkendali berdasarkan nilai baseline dari nilai NPHR yang melebihi 2650.

Tabel 2. Status 10 teratas variabel terkendali berdasarkan nilai *baseline*

No.	Sensor	Upper Limit	Lower Limit	Good	Bad	Status
1	MAIN STEAM Temp	541.695	536.305	1321	33	Normal
2	HOT RH STEAM Temp. (AVG)	541.695	536.305	1267	87	Normal
3	MS WATER SPRAY Flow	21.494	17.586	67	1287	Kritis
4	RH WATER SPRAY Flow	1.76	1.44	895	459	Normal
5	MAIN STEAM Press	169.302	164.298	1157	197	Normal
6	MAIN APH GAS OUT Temp (AVG)	162.18	155.82	126	1228	Kritis
7	FINAL FEED WATER Temp	246.7686	237.0914	232	1122	Kritis
8	LP TRB EXH. VACUUM Press	711.747	697.653	1351	3	Normal
9	COND. MAKE UP Flow	1.045	0.855	1288	66	Normal
10	O2% IN FLUE GAS OUT (AVG)	2.92116	2.52282	88	1266	Kritis

Dari Tabel 2 dapat disimpulkan bahwa terdapat 4 sensor yang mengalami anomali dengan status kritis, sensor tersebut adalah Main Steam Water Spray Flow, Main APH Gas Out Temp (Avg), Final Feed Water Temperature dan O2% in Flue Gas Out. Tabel 3 memvalidasi bahwa 3 dari 4 sensor yang memiliki status kritis ini adalah sensor yang mempengaruhi kenaikan nilai NPHR sehingga peralatan yang terkait dengan indikator operasi ini menjadi bagian dari lingkup pemeliharaan. Sensor Main Steam Water Spray Flow termasuk dalam kondisi kritis tetapi tidak termasuk dari bagian yang mempengaruhi kenaikan nilai NPHR dikarenakan sensor Main Steam Water Spray Flow sudah kritis dari awal pada saat nilai NPHR masih rendah.

4. Kesimpulan

Model diagnosis pada Data Top 10 Variabel Terkendali dan NPHR yang dibangun dengan algoritma RFC mempunyai akurasi 92,8% dengan model terbaik menggunakan parameter *max_depth* sebesar 5 dengan estimator pohon keputusan sebanyak 50. Keluaran model RFC didapatkan data sensor yang mempengaruhi kenaikan nilai NPHR adalah data penurunan temperatur FFW dan kenaikan temperatur Main APH gas out dan konten oksigen pada gas buang. Pengaruh dari sensor yang menyebabkan nilai NPHR divalidasi dengan status kondisi kritis dari setiap sensor dengan menggunakan baseline value. Dari 4 sensor kritis, didapatkan 3 sensor yang sama yang mempengaruhi kenaikan nilai NPHR.

Dari penelitian ini diperlukan adanya peningkatan terhadap teknologi akuisisi data, yang terdapat di PI system untuk mendapatkan nilai data sensor yang akurat. Model yang digunakan pada penelitian ini masih dalam tahap pengembangan sehingga belum dapat digunakan dengan semestinya dan perlu dikembangkan lebih lanjut pada penelitian selanjutnya

Referensi

- [1] Moriarty, P., & Honnery, D. (2019). Energy efficiency or conservation for mitigating climate change?. *Energies*, 12(18), 3543.
- [2] Wang, J. M., Ge, X., Zhang, L., & Zhang, H. (2016). Management index systems and energy efficiency diagnosis model for power plant: cases in China. *Mathematical Problems in Engineering*, 2016.
- [3] Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
- [4] Loh, W. Y. (2011). Classification and regression trees. *Wiley interdisciplinary reviews: data mining and knowledge discovery*, 1(1), 14-23.
- [5] Economidou, M., Todeschi, V., Bertoldi, P., D'Agostino, D., Zangheri, P., & Castellazzi, L. (2020). Review of 50 years of EU energy efficiency policies for buildings. *Energy and Buildings*, 225, 110322.
- [6] Jørgensen, S. E. (2010). Ecosystem services, sustainability and thermodynamic indicators. *Ecological Complexity*, 7(3), 311-313.
- [7] Giaccone, E., & Mancò, S. (2012). Energy efficiency measurement in industrial processes. *Energy*, 38(1), 331-345.
- [8] Rosenquist, G., McNeil, M., Iyer, M., Meyers, S., & McMahon, J. (2006). Energy efficiency standards for equipment: Additional opportunities in the residential and commercial sectors. *Energy Policy*, 34(17), 3257-3267.
- [9] Storm, D. (2009). Applying the Fundamentals for Best Heat Rate Performance of Pulverized Coal Fueled Boilers. EPRI Heat Rate Conference, 1-14.
- [10] Han, J., Pei, J., & Tong, H. (2022). *Data mining: concepts and techniques*. Morgan kaufmann.

- [11] Loh, W. Y. (2011). Classification and regression trees. *Wiley interdisciplinary reviews: data mining and knowledge discovery*, 1(1), 14-23.
- [12] Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (Vol. 2, pp. 1-758). New York: springer.
- [13] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, 12, 2825-2830.