Syntax Literate: Jurnal Ilmiah Indonesia p–ISSN: 2541-0849
e-ISSN: 2548-1398
Vol. 10, No. 2, Februari 2025

PEMODELAN PREDIKSI UMPAN PULVERIZED COAL PADA SISTEM ROTARY KILN PABRIK SEMEN BERDASARKAN PARAMETER OPERASI & KIMIA BERBASIS MACHINE LEARNING

Riduwan Maliki
Sekolah Interdisiplin Manajemen dan Teknologi, Indonesia
Email: riduwan.maliki@outlook.com

Abstrak

Penelitian ini bertujuan untuk mengembangkan model prediksi umpan pulverized coal (PC) pada sistem rotary kiln di pabrik semen menggunakan algoritma machine learning. Model ini diharapkan dapat membantu operator Control Center Room (CCR) dalam mengoptimalkan konsumsi PC dan meningkatkan efisiensi produksi semen. Industri semen di Indonesia mengalami persaingan ketat dan tuntutan efisiensi operasional akibat kenaikan biaya energi. Konsumsi PC merupakan salah satu komponen biaya terbesar dalam produksi semen, dengan potensi optimasi yang signifikan. Penelitian ini berfokus pada tiga pertanyaan utama, yaitu 1.) membangun model prediksi umpan PC berbasis machine learning pada rotary kiln dengan memanfaatkan data parameter operasi, dan sifat kimia bahan baku dan produk; 2.) menentukan algoritma machine learning yang paling akurat dalam memprediksi nilai umpan PC pada rotary kiln; 3.) menentukan faktor apa yang paling signifikan dalam mempengaruhi nilai umpan PC pada rotary kiln. Penelitian ini diharapkan menghasilkan model prediksi PC yang akurat dan handal, serta mengidentifikasi faktor-faktor kunci yang mempengaruhi konsumsi PC. Hasil penelitian menunjukkan bahwa algoritma Gradient Boosting memberikan akurasi terbaik dengan nilai R-squared sebesar 0.976. Model yang dikembangkan mampu memprediksi nilai umpan yang lebih rendah dibandingkan dengan pengaturan manual oleh operator. Implikasi dari hasil ini adalah potensi penghematan penggunaan batubara dan pengurangan emisi CO2 dalam proses produksi semen.

Kata kunci: rotary kiln, cement, coal, energy, machine learning, Support Vector Machine (SVM), Gradient Boosting, dan Neural Network

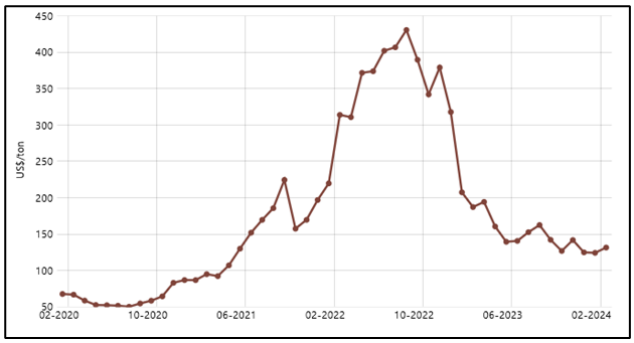
Abstract

This research aims to develop a prediction model for pulverized coal (PC) feed in the rotary kiln system at a cement plant using machine learning algorithms. This model is expected to assist Control Center Room (CCR) operators in optimizing PC consumption and improving cement production efficiency. The cement industry in Indonesia is experiencing intense competition and demands for operational efficiency due to rising energy costs. PC consumption is one of the largest cost components in cement production, with significant optimization potential. This research focuses on three main questions, namely 1.) building a machine learning-based PC feed prediction model in rotary kilns by utilizing data on operating parameters, and chemical properties of materials and products; 2.) determining the most accurate machine learning algorithm in predicting PC feed values in rotary kilns; 3.) determining what factors are most significant in influencing PC feed values in rotary kilns. This research is expected to produce an accurate and reliable PC prediction model, as well as identify key factors that affect PC consumption. The results indicate that the Gradient Boosting algorithm achieved the highest accuracy with an R-squared value of 0.976. The developed model can predict lower feed values compared to manual operator

adjustments. This implies potential savings in coal consumption and reduced CO2 emissions in the cement production process.
Keywords: *rotary kiln, cement, coal, energy, machine learning, Support Vector Machine (SVM), Gradient Boosting, dan Neural Network*

Pendahuluan

Industri semen di Indonesia dalam kurun waktu 5 tahun terakhir (2019-2023) mengalami persaingan yang ketat dengan beberapa pemain utama antara lain Semen Indonesia Group, Indocement, dan beberapa pemain regional dan skala global seperti Bosowa, Anhui Conch Cement, Jui shin, dll. Persaingan yang ketat tersebut tidak diimbangi dengan besarnya jumlah permintaan di bagian hilir, sehingga menyebabkan kondisi oversupply. Disamping itu kondisi perekonomian di Indonesia dimana masih dalam proses pemulihan pasca Pandemi COVID-19, dan adanya kenaikan biaya energi (batubara) serta kondisi ketidakpastian global menuntut pemain industri semen untuk bisa bertahan salah satunya dengan melakukan banyak upaya efisiensi operasional (X. Hu, 2017; Kelleher et al., 2015; Natekin & Knoll, 2013). Biaya terbesar dalam produksi semen adalah terkait dengan pemakaian bahan bakar (energi), dimana biaya tersebut menyerap lebih dari 30% dari total manufacturing cost (Sumber Annual report SIG, 2022).

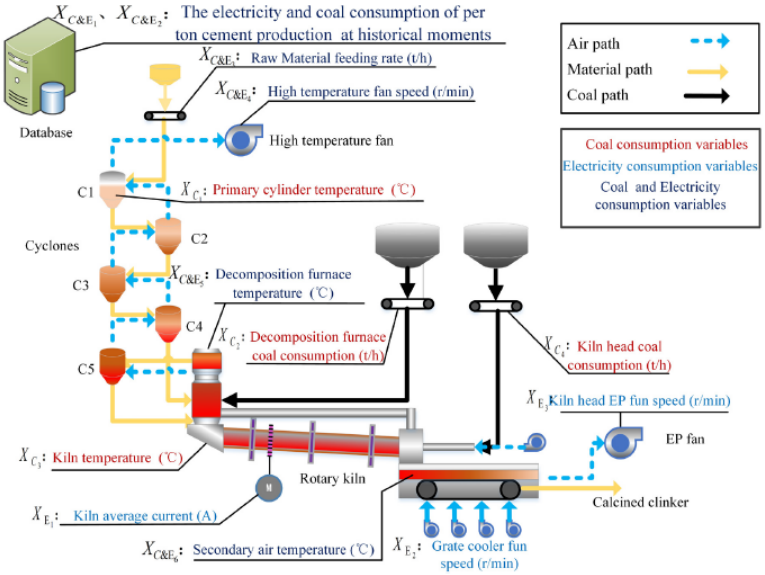


Gambar 1. Grafik Fluktuasi Harga Batubara Kurun 02/2020-02/2024

Pada gambar 1 diatas menunjukkan bahwa harga batubara pernah naik signifikan diatas 400% dalam kurun waktu kurang dari 2 tahun. Tentunya kondisi seperti ini akan menambah biaya produksi semen.

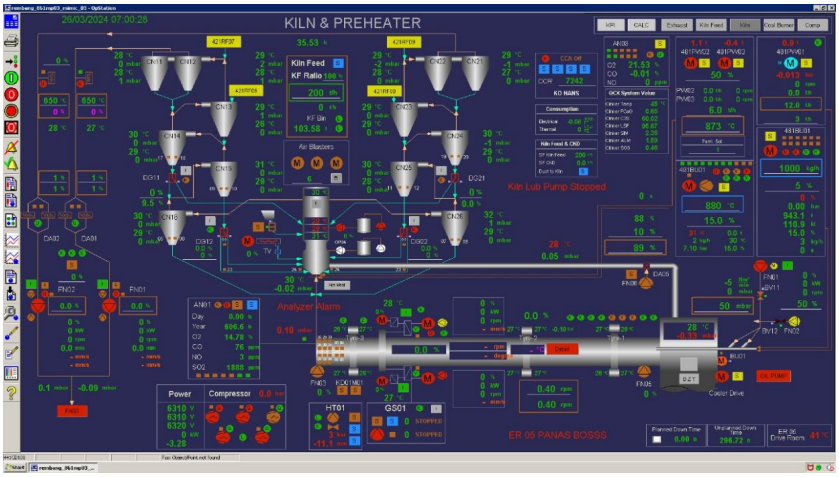
Terdapat 2 jenis energi yang digunakan dalam produksi semen, yaitu energi panas dari pembakaran batubara dan energi listrik. Konsumsi energi terbesar adalah pada energi termal dengan penggunaan sekitar 89% dari total konsumsi energi. Penggunaan energi termal dominan digunakan untuk pembakaran di rotary kiln untuk menghasilkan clinker atau terak. Energi termal didapatkan dari pembakaran batubara yang sebelumnya telah dilakukan penggilingan di coal mill hingga menjadi pulverized coal. Gambar 2 memperlihatkan layout untuk sistem rotary kiln.

Pembentukan material clinker membutuhkan panas dimana setiap kilogram clinker dipanaskan hingga mencapai suhu pembakaran. Energi termal untuk menjalankan proses reaksi kimia dan fisika material penyusun clinker (W. Hu & Mao, 2024; X. Liu et al., 2022; Zheng et al., 2022). Proses yang membutuhkan energi termal ini meliputi pengeringan, penguapan, peningkatan kemampuan meleleh, dan penggabungan material (kalsinasi) (X. Zhang & Zhao, 2011).

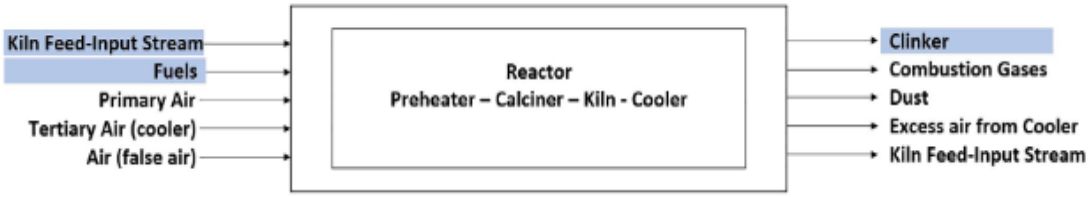


Gambar 2. Sistem Rotary Kiln

Pengoperasian rotary kiln semen untuk mencapai titik optimal merupakan hal yang cukup sulit dilakukan oleh operator. Gambar 3 memperlihatkan tampilan kendali operasi untuk operator rotary kiln yang mengendalikan secara remote dari Control Center Room (CCR). Ada tiga hal yang harus dicapai oleh operator CCR pada rotary kiln, yaitu: mengurangi konsumsi bahan bakar, mencapai volume produksi clinker, dan menjaga kualitas semen. Ketiga hal ini saling terkait dan dipengaruhi oleh faktor-faktor yang tidak bisa diprediksi, seperti kualitas input bahan baku yang berubah tergantung kandungan kimia bahan baku, pengaturan panas internal kiln, komponen kimia clinker, dan lain-lain. Untuk kompensasi perubahan variabel tersebut, operator CCR perlu menyesuaikan pembakaran di dalam rotary kiln dengan cara melakukan penyesuaian umpan pulverized coal. Setiap operator memiliki intuisi tersendiri dalam melakukan penyesuaian tersebut. Oleh karena itu perlu adanya standard atau assistance tools kepada setiap operator CCR agar didapatkan kombinasi pembakaran yang tepat dan tidak berlebihan ketika ada perubahan parameter operasi maupun kimia di dalam prosesnya (Qin & Dong, 2020; Ramasamy et al., 2023; L. Zhang et al., 2019). Secara umum blok diagram terkait parameter operasi dan kimia di rotary kiln semen dapat ditunjukkan pada gambar 4.



Gambar 3. Tampilan Kendali Rotary Kiln Oleh Operator di CCR



Gambar 4. Blok Diagram Pada Rotary Kiln (Ali, 2022)

Kompleksitas dalam operasi rotary kiln memberikan peluang dalam pengembangan model yang mampu bekerja dalam situasi matematis kompleks seperti algoritma machine learning. Beberapa penelitian membuktikan bahwa model algoritma machine learning mampu memodelkan untuk sistem operasi kompleks seperti algoritma Support Vector Machine (SVM), Gradient Boosting, Adaboost, Random Forest, dan Neural Network. Algoritma machine learning juga akan menghasilkan model yang bisa digunakan untuk memprediksi kebutuhan umpan pulverized coal berdasarkan 3 kelompok parameter yaitu: operasi pembakaran, sifat kimia bahan baku, dan sifat kimia produk clinker. Hasil prediksi pemodelan diharapkan dapat memberikan masukan kepada operator CCR dalam melakukan setting umpan pulverized coal pada rotary kiln agar didapatkan konsumsi pulverized coal yang tidak berlebihan dan produk clinker yang dihasilkan sesuai dengan parameter kimia yang dipersyaratkan (Fatahi et al., 2023; Q. Liu et al., 2020; Ramasamy et al., 2019).

Dalam kajian mengenai industri semen di Indonesia, terdapat beberapa penelitian terdahulu yang relevan. Pertama, penelitian oleh Hu (2017) yang membahas tentang efisiensi operasional dalam industri semen, menunjukkan bahwa perusahaan perlu melakukan inovasi untuk bertahan dalam kondisi oversupply dan meningkatnya biaya energi. Kedua, penelitian oleh Kelleher et al. (2015) yang mengeksplorasi penggunaan teknologi dalam meningkatkan efisiensi energi di pabrik semen, menekankan pentingnya penerapan teknologi baru untuk mengurangi biaya produksi. Namun, terdapat gap dalam penelitian ini, yaitu kurangnya fokus pada penerapan algoritma machine learning dalam memprediksi kebutuhan umpan pulverized coal pada rotary kiln di pabrik semen.

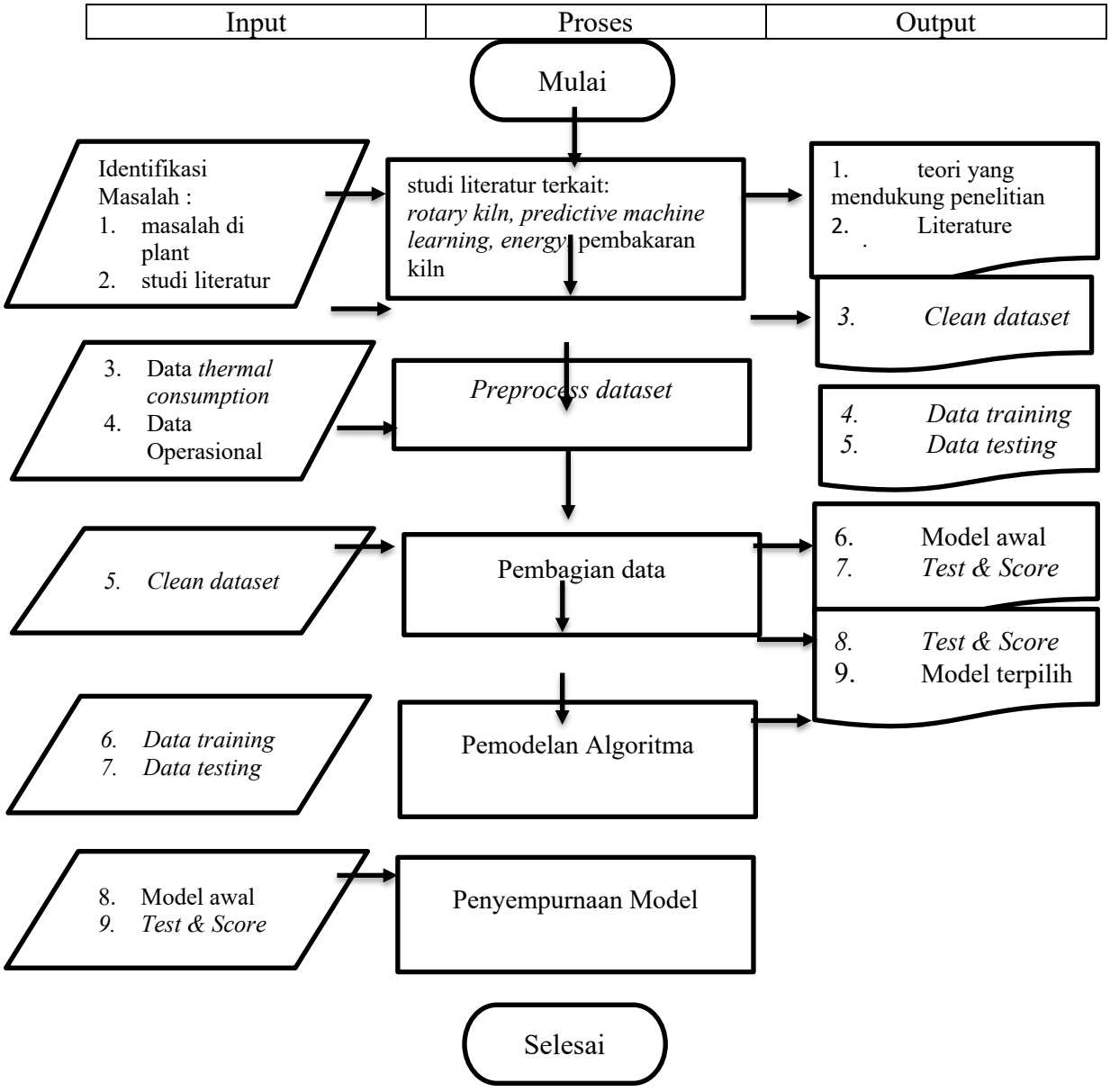
Tujuan penelitian yaitu Untuk menganalisis faktor yang paling signifikan dalam mempengaruhi nilai umpan pulverized coal pada rotary kiln di pabrik semen, Untuk menganalisis dan mengevaluasi model algoritma machine learning yang mampu memprediksi nilai umpan pulverized coal pada rotary kiln di pabrik semen, Untuk menganalisis dan mendapatkan model prediksi machine learning terbaik untuk prediksi umpan pulverized coal pada sistem rotary kiln pabrik semen.

Manfaat dalam penelitian ini adalah sebagai sarana melatih dan berfikir ilmiah pada bidang manajemen energi dan bahan, bagi perusahaan terkait diharapkan hasil penelitian ini dapat dijadikan bahan masukan dalam strategi menerapkan prediksi kebutuhan pulverized coal dengan memanfaatkan algoritma machine learning, unutk mendapatkan insight faktor-faktor yang mempengaruhi nilai umpan pulverized coal pada rotary kiln di pabrik semen, dan bagi peneliti lain sebagai acuan dan perbandingan dalam penelitian mengenai objek masalah yang sama di masa mendatang.

Metode Penelitian

Pendekatan Penelitian

Berikut diagram metodologi dalam penulisan penelitian ini.



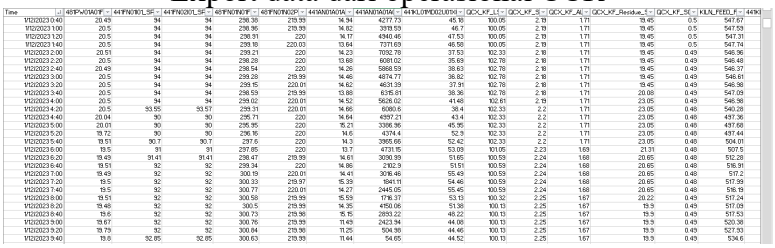
Gambar 5. *Flowchart* Penelitian

Pengumpulan Data

Data penelitian didapatkan dari data export pada sistem kontrol operasi Kiln pabrik semen di Pulau Jawa, Indonesia. Data export dalam bentuk raw data berformat .xlsx yang memuat nama parameter dan nilai parameter. Data yang digunakan adalah data dalam kurun waktu operasi bulan September 2023 hingga bulan Mei 2024. Gambar 6 merupakan tampilan CCR untuk sistem control kiln dan proses melakukan export data ke dalam format .xlsx.



Export data dari operasional CCR



Gambar 6. Ilustrasi Export Data Operasional CCR

Pemilahan Parameter Feature dan Target

Data yang telah terkumpul dalam format .xlsx selanjutnya dilakukan pemilahan dengan memisahkan parameter yang tidak digunakan dalam pemodelan. Data yang digunakan memuat 24 parameter yang dapat dilihat pada tabel 1.

Tabel 1. Parameter Untuk Pemodelan

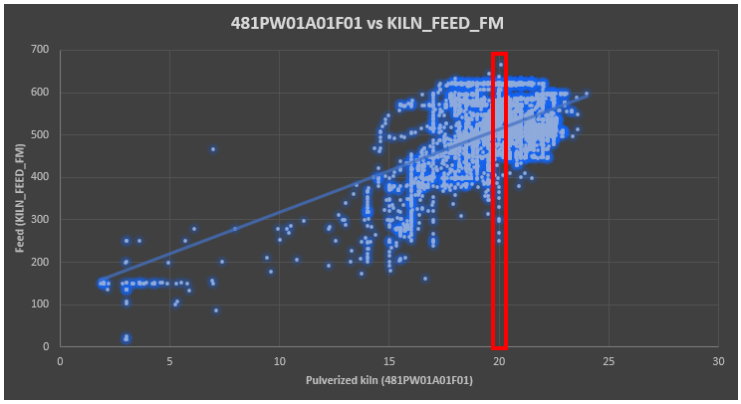
No.	Kategori	Jumlah Parameter	Nama Parameter
1	Parameter Target	1	• *Pulverized kiln* (481PW01A01F01)
2	Parameter Operasi	13	• *Flow Primary* (481FN01N01F01) • *Pressure Primary* (481FN01N02P01) • *O2 inlet* (441AN01A01A01) • *CO inlet* (441AN01A01A02) • *Torque kiln* (441KL01MD02U01X01) • *Feed* (KILN_FEED_FM) • *Temp. Kiln* (441KH01N01T01) • *Temp. TAD* (491DU01N01T01) • *Speed kiln* (441KL01MD02I01_SPMR) • *Pulverized calciner* (481PW02I01_SPM) • *Temp. clinker* (QCX_Clinker_Temp) • *Press. Inlet kiln* (441KL01MD02N01P01) • *Temp. inlet kiln* (441CI01N03T01)
3	Parameter Kimia Material input	4	• *LSF kiln feed* (QCX_KF_LSF) • *Silica modulus* KF (QCX_KF_SIM) • *Alumina modulus* KF (QCX_KF_ALM) • *SO3 kiln feed* (QCX_KF_SO3)
4	Parameter Kimia	6	• *Freelime clinker* (QCX_Clinker_FCaO) • *C3S clinker* (QCX_Clinker_C3S)

No.	Kategori	Jumlah Parameter	Nama Parameter
	Produk (Clinker)		• *LSF clinker* (QCX_Clinker_LSF) • *Silica Clinker* (QCX_Clinker_SIM) • *Alumina clinker* (QCX_Clinker_ALM) • SO3 *clinker* (QCX_Clinker_SO3)

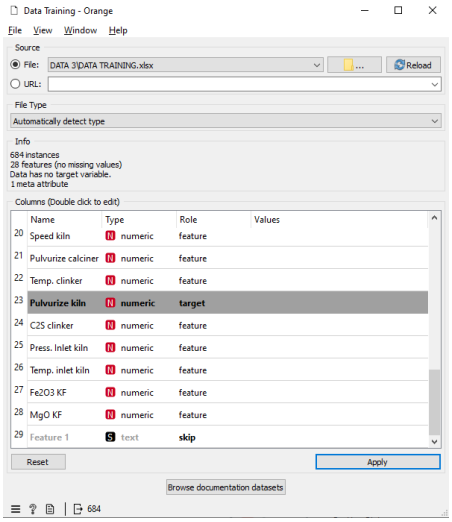
Hasil dan Pembahasan

Penentuan Dataset Untuk Pelatihan Pemodelan Algoritma

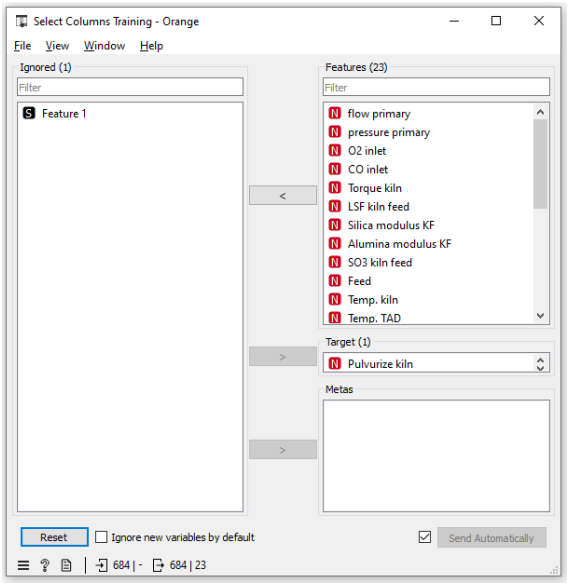
Data yang didapatkan dari operasional CCR berupa data kontinyu dengan time series. Melalui data operasional tersebut maka dilakukan perbandingan antara parameter Pulverized kiln (481PW01A01F01) terhadap Feed (KILN_FEED_FM) seperti yang ditunjukkan pada Gambar 7 yang menggambarkan bagaimana pola operasi yang dilakukan oleh operator CCR untuk melakukan perubahan Pulverized kiln (481PW01A01F01) terhadap Feed (KILN_FEED_FM). Dataset yang akan digunakan untuk model algoritma adalah untuk kelompok data yang memiliki range penggunaan Pulverized kiln (481PW01A01F01) yang optimal terhadap Feed (KILN_FEED_FM). Penggunaan Pulverized kiln (481PW01A01F01) yang optimal terhadap Feed (KILN_FEED_FM) bisa didapatkan pada area diatas garis regresi linier antara parameter Pulverized kiln (481PW01A01F01) terhadap Feed (KILN_FEED_FM). Pada grafik terlihat, bahwa untuk jumlah umpan yang sama (misal 20 tph) pada parameter Pulverized kiln (481PW01A01F01) pada sumbu x menunjukkan bahwa operator CCR menggunakannya untuk membakar Feed (KILN_FEED_FM) yang berada di atas garis regresi linier maupun dibawah garis regresi linier. Area dimana Penggunaan Pulverized kiln (481PW01A01F01) dapat dikatakan efisien atau optimal adalah pada area luasan diatas garis regresi linier. Kelompok data yang berada diatas garis regresi akan digunakan sebagai Dataset model. Sehingga output / hasil prediksi juga diharapkan mengarah pada nilai umpan pulverized coal yang paling efisien atau optimal terhadap feed kiln.



Gambar 7. Grafik Pulverized Kiln (481PW01A01F01) Terhadap Feed (KILN_FEED_FM)



Gambar 8. Window Widget File



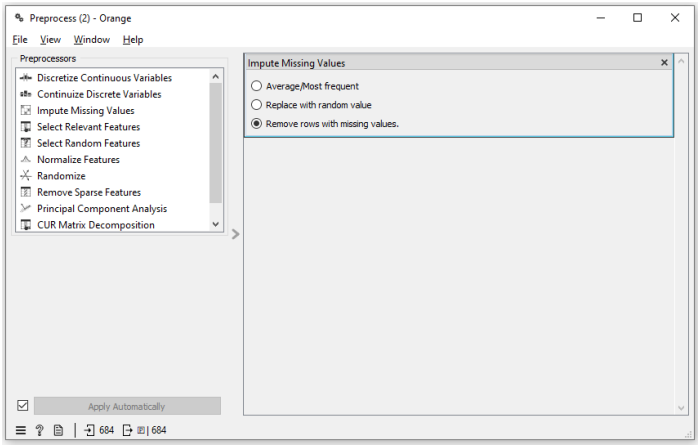
Gambar 9. Window Widget Select Columns

Dataset pelatihan yang telah didapatkan dilakukan penginputan ke dalam Aplikasi Orange Datamining melalui widget: File. Semua parameter dirubah ke dalam type: numeric. Parameter Pulverized kiln menggunakan Role: Target dan 23 parameter lainnya menggunakan Role: feature. Parameter yang tidak digunakan dalam pemodelan bisa menggunakan Role: Skip (Gambar 8).

Kita juga dapat menggunakan widget: Select Columns untuk melakukan pemilahan parameter ke dalam kategori features, target, atau ignored (Gambar 9). Melalui widget: Select Columns kita dapat melihat secara keseluruhan jumlah parameter yang masuk dalam kelompok features. Features adalah kelompok parameter yang akan menjadi input algoritma pemodelan, sedangkan kolom target hanya dapat diisi oleh satu buah parameter saja, dimana kita sudah tentukan dalam penelitian ini untuk parameter target adalah pulverized kiln.

Cleaning Dataset

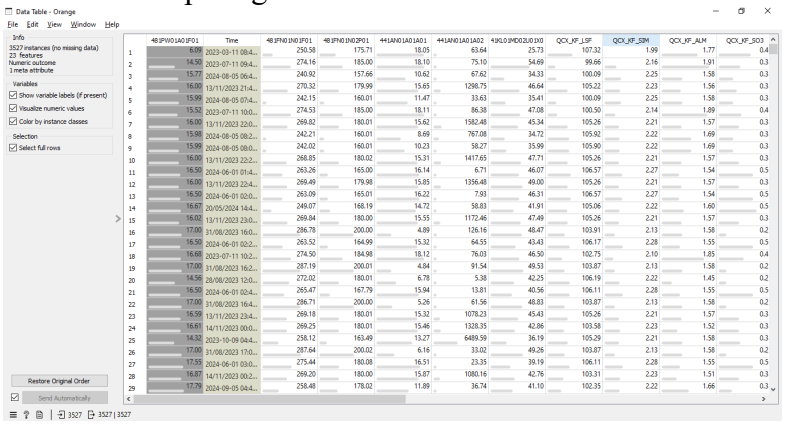
Selanjutnya input data dilakukan preprocess dengan menghilangkan baris data yang tidak memiliki nilai (value). Hal tersebut perlu dilakukan agar algoritma pemodelan tidak mengkonsumsi data yang tidak valid, sehingga mencegah terbentuknya hasil algoritma pemodelan yang tidak akurat atau tidak mewakili proses / sistem pembakaran kiln. Widget yang digunakan untuk tahapan ini adalah widget: Preprocess yang dapat dilihat pada gambar 10 berikut.



Gambar 10. Window Widget Preprocess

Pengecekan Dataset

Pengecekan akhir pada dataset perlu dilakukan untuk memastikan tidak ada data dengan missing value dan parameter yang digunakan untuk training pemodelan sudah masuk semua. Pengecekan bisa kita lakukan dengan menggunakan widget: Data Table dengan tampilan window seperti gambar 11 berikut.

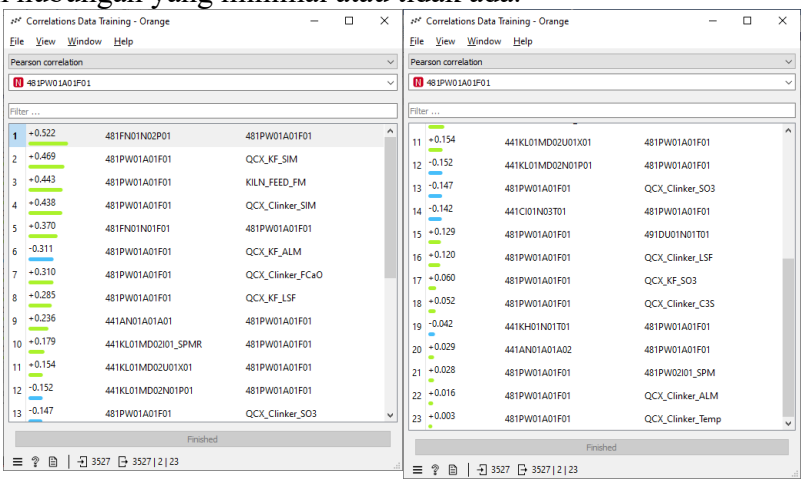


Gambar 11. Window Widget Data Table

Analisis Korelasi

Analisis korelasi digunakan untuk melihat parameter feature yang paling signifikan dalam mempengaruhi parameter target (Pulverized kiln -481PW01A01F01). Penelitian ini menggunakan Pearson Correlation untuk melihat keeratan hubungan antar 2 variabel. Pada Orange Datamining dapat menggunakan widget: Correlations dimana hasil untuk Pearson Correlation dapat dilihat pada Gambar 4.7. Pada gambar tersebut dapat dilihat bahwa terdapat beberapa variabel yang memiliki korelasi positif yang cukup kuat dengan variabel Pulverized kiln (481PW01A01F01). Variabel-variabel feature ini antara lain Pressure Primary (481FN01N02P01), Silica modulus KF (QCX_KF_SIM),

Feed (KILN_FEED_FM), dan Silica Clinker (QCX_Clinker_SIM). Hal ini menunjukkan bahwa ketika nilai variabel-variabel feature tersebut meningkat, cenderung diikuti oleh peningkatan nilai variabel target atau Pulverized kiln (481PW01A01F01). Beberapa variabel feature menunjukkan korelasi negatif dengan Pulverized kiln (481PW01A01F01), seperti Alumina modulus KF (QCX_KF_ALM), Press. Inlet kiln (441KL01MD02N01P01), SO3 clinker (QCX_Clinker_SO3), dan Temp. inlet kiln (441CI01N03T01). Artinya, ketika nilai variabel-variabel feature ini meningkat, cenderung diikuti oleh penurunan nilai variabel Pulverized kiln (481PW01A01F01). Sebagian besar variabel feature lainnya memiliki korelasi yang lemah (nilai koefisien korelasi mendekati 0) terhadap variabel Pulverized kiln (481PW01A01F01), menunjukkan hubungan yang minimal atau tidak ada.

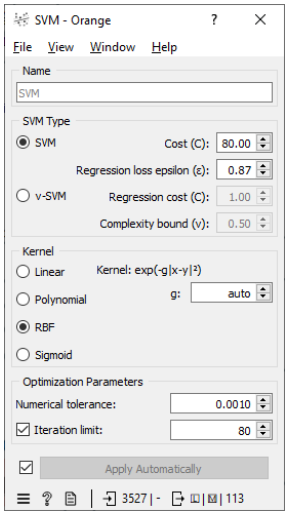


Gambar 12. Hasil Pearson Correlation

Pemodelan Algoritma

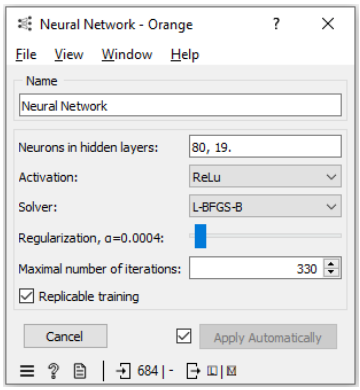
Pada penelitian ini akan digunakan 3 (tiga) jenis algoritma untuk melakukan pemodelan, yaitu Support Vector Machine (SVM), Gradient Boosting, dan Neural Network. Pemodelan dilakukan secara paralel dengan menggunakan dataset yang sama untuk ketiga algoritma pemodelan tersebut. Dataset total akhir yang didapatkan setelah melalui proses pemilahan dan cleaning adalah sebanyak 3607 instances. Dataset tersebut dibagi menjadi 2 dengan prosentase 97.8 % untuk data training dan 2.2 % untuk data test. Data training digunakan untuk membentuk model algoritma pada Support Vector Machine (SVM), Gradient Boosting, dan Neural Network. Sedangkan Data test digunakan untuk menguji model berdasarkan nilai feature yang belum pernah digunakan untuk pembentukan model algoritma sebelumnya.

Pada masing-masing algoritma (Support Vector Machine (SVM), Gradient Boosting, dan Neural Network) untuk awal pemodelan pada penelitian ini menggunakan setting default untuk nilai parameter fungsi masing-masing algoritma. Jika didapatkan nilai evaluasi yang belum memenuhi requirement, maka pada Orange datamining telah diberikan tools untuk bisa melakukan perubahan pada nilai parameter fungsi algoritma tersebut. Berikut tersaji pada Gambar 4.8, Gambar 4.9, dan Gambar 4.10 adalah hasil optimasi nilai parameter fungsi yang dilakukan untuk mendapatkan nilai evaluasi pemodelan yang terbaik.



Gambar 13. Nilai Parameter Fungsi Model SVM Setelah Optimasi

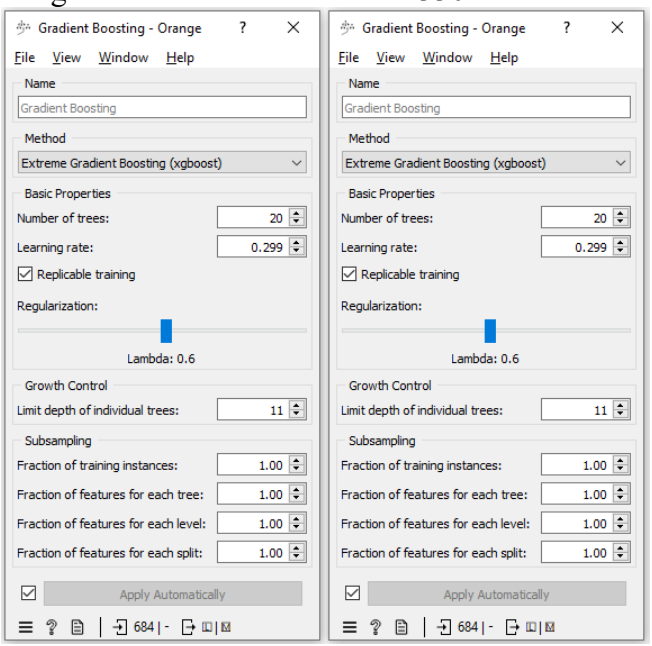
Pada Parameter Fungsi Model Support Vector Machine (SVM) (Gambar 13) terdapat fungsi Cost (C) yang diistilahkan sebagai penalti untuk kerugian dan berlaku untuk tugas klasifikasi dan regresi. Sedangkan ε merupakan parameter untuk model epsilon-SVR, berlaku untuk tugas regresi yang digunakan untuk mendefinisikan jarak dari nilai sebenarnya di mana tidak ada penalti yang terkait dengan nilai prediksi. Keduanya (C dan ε) digunakan untuk mengatur batas kesalahan pengujian. Kernel adalah fungsi yang mengubah ruang atribut menjadi ruang fitur baru agar sesuai dengan hyperplane margin maksimum, sehingga memungkinkan algoritma untuk membuat model dengan beberapa pilihan fungsi kernel yang tersedia antara lain Linear, Polinomial, RBF, dan Sigmoid. Optimasi yang didapatkan untuk penggunaan kernel RBF kami dapati menghasilkan skor kernel lebih baik daripada kernel lainnya. Fungsi g adalah untuk konstanta gamma dalam fungsi kernel (nilai yang disarankan adalah 1/k, di mana k adalah jumlah atribut, tetapi karena mungkin tidak ada set pelatihan yang diberikan kepada widget, maka defaultnya adalah 0 dan pengguna harus mengatur opsi ini secara manual). Numerical Tolerance digunakan untuk menetapkan deviasi yang diizinkan dari nilai yang diharapkan. Iteration limit digunakan untuk mengatur jumlah maksimum iterasi yang diizinkan.



Gambar 14. Nilai Parameter Fungsi Model Neural Network Setelah Optimasi

Hasil optimasi pada model Neural Network yang telah dilakukan dalam penelitian ini dapat ditunjukkan pada Gambar 14. Pada Orange datamining Algoritma Neural Network yang digunakan adalah A multi-layer perceptron (MLP) algorithm with

backpropagation. Neurons in hidden layer: didefinisikan sebagai elemen ke-i yang mewakili jumlah neuron di lapisan tersembunyi ke-i. Hasil optimasi parameter untuk penelitian ini didapatkan hasil evaluasi skor yang baik menggunakan 2 layer dengan susunan layer pertama 80 neuron dan layer kedua 19 neuron. Fungsi aktivasi untuk lapisan tersembunyi menggunakan ReLu (Rectified Linear Unit function). Adapun fungsi aktivasi lainnya yang tersedia pada aplikasi Orange Datamining adalah Identity, Logistic (logistic sigmoid function) dan tanh (hyperbolic tan function). Solver untuk optimasi Weight disediakan 3 pilihan, yaitu L-BFGS-B, SGD, dan Adam. Pada hasil optimasi kami mendapati bahwa L-BFGS-B cocok digunakan untuk penelitian ini. Alfa digunakan untuk melakukan setting parameter penalti L2 (regularisasi). Selain itu juga jumlah iterasi dapat disesuaikan, dimana digunakan maksimal iterasi di 330.



Gambar 15. Nilai Parameter Fungsi Model Gradient Boosting Setelah Optimasi

Gradient Boosting adalah teknik pembelajaran mesin untuk masalah regresi dan klasifikasi, yang menghasilkan model prediksi dalam bentuk ensemble (penggabungan) model prediksi yang lemah, biasanya berupa decision trees (Shi et al., 2022; R. Zhang et al., 2021). Ada beberapa pilihan metode Gradient Boosting yang digunakan dalam widget aplikasi Orange Datamining, yaitu Gradient Boosting (scikit-learn), Extreme Gradient Boosting (xgboost), Extreme Gradient Boosting Random Forest (xgboost), Gradient Boosting (catboost). Penelitian ini memilih untuk menggunakan XGboost. XGBoost adalah pustaka peningkatan gradien terdistribusi yang dioptimalkan yang dirancang agar sangat efisien, fleksibel, dan portabel. Library ini mengimplementasikan algoritma pembelajaran mesin di bawah kerangka kerja Gradient Boosting. XGBoost menyediakan peningkatan pohon paralel (juga dikenal sebagai GBDT, GBM) yang memecahkan banyak masalah ilmu data dengan cara yang cepat dan akurat. Adapun Basic Properties yang disediakan antara lain:

1. Number of Trees (Jumlah pohon): menentukan berapa banyak pohon yang ditingkatkan gradiennya yang akan disertakan. Jumlah yang banyak biasanya menghasilkan performa yang lebih baik.
2. Learning rate (Laju pembelajaran): menentukan laju pembelajaran yang ditingkatkan. Laju pembelajaran mengecilkan kontribusi setiap pohon.

3. Replicable training (Pelatihan yang dapat direplikasi): memperbaiki benih acak, yang memungkinkan replikasi hasil.
4. Regularization: menentukan jangka waktu regularisasi L2. Hanya tersedia untuk metode xgboost dan catboost.

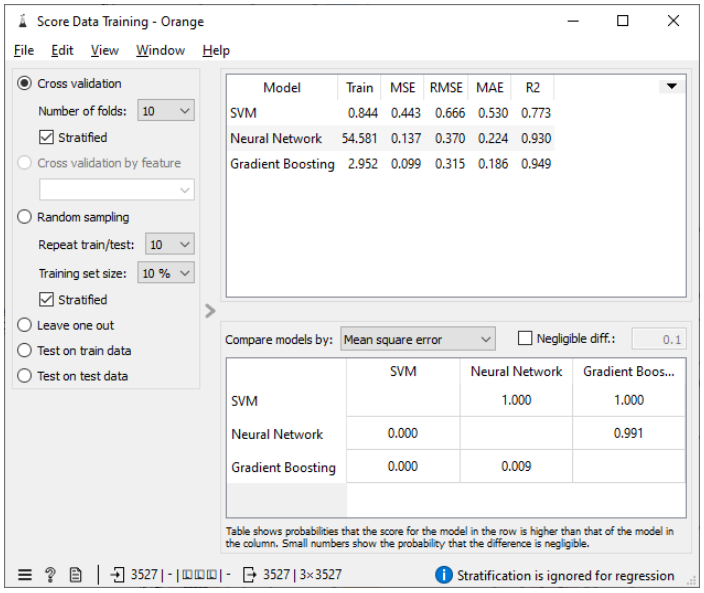
Growth Control dilakukan untuk mengendalikan kedalaman trees pada algoritma. Sedangkan Subsampling digunakan untuk menentukan persentase fitur yang akan digunakan saat membangun setiap pohon.

Evaluasi Pemodelan Algoritma

Pada tahap evaluasi model, penelitian ini menggunakan K-fold cross validation untuk mendapatkan nilai MSE, RMSE, MAE, dan R2 dengan jumlah folds sebanyak 10. K-fold cross validation adalah teknik evaluasi model machine learning yang umum digunakan untuk mengukur performa model secara lebih akurat dan objektif. Jika dibandingkan dengan metode evaluasi lainnya, K-fold cross validation memiliki beberapa kelebihan, yaitu mencegah overfitting, memperkirakan performa model yang lebih akurat, analisis generalisasi model lebih bagus. Berdasarkan kelebihan tersebut K-fold cross validation lebih cocok digunakan untuk evaluasi algoritma machine learning dengan kebutuhan prediktif. Evaluasi model di aplikasi Orange Datamining dapat dilakukan dengan menggunakan widget: test & score (Okoji et al., 2022).

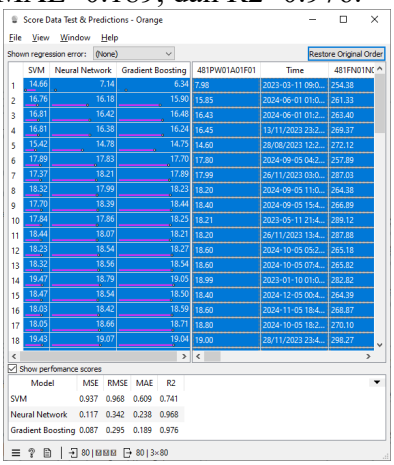
Pada Gambar 16 merupakan hasil skor evaluasi model menggunakan data training dimana terlihat bahwa model dengan algoritma Gradient Boosting secara umum memiliki nilai MSE, RMSE, MAE lebih rendah dan R2 yang lebih tinggi daripada model algoritma Support Vector Machine (SVM) dan Neural Network. Semakin rendah nilai MSE, RMSE, MAE menunjukkan bahwa error antara nilai target dan prediktif semakin kecil. R-squared (R2) adalah salah satu metrik yang paling umum digunakan dalam evaluasi model regresi pada machine learning. Metrik ini memberikan kita gambaran tentang seberapa baik model kita mampu menjelaskan variabilitas dalam data. Nilai R2 yang tinggi menunjukkan bahwa model kita sangat baik dalam menjelaskan data. Semakin tinggi nilai R2 (maksimal 1), semakin baik model kita dalam memprediksi nilai variabel dependen.

Hasil Skor akurasi (Gambar 16) untuk Gradient Boosting adalah MSE=0.099, RMSE=0.315, MAE=0.186, dan R2=0.949. Sedangkan untuk model Neural Network dengan skor akurasi yang tidak jauh berbeda dari model Gradient Boosting berada diurutan kedua dengan nilai MSE=0.137, RMSE=0.370, MAE=0.224, dan R2=0.930. Namun, Neural Network memiliki waktu Train yang lebih lama 18 kali jika dibandingkan dengan model algoritma Gradient Boosting. Model Support Vector Machine (SVM) memiliki skor akurasi yang paling buruk diantara ketiga pemodelan algoritma yang diujikan dengan skor akurasi MSE=0.443, RMSE=0.666, MAE=0.530, dan R2=0.773. Hal tersebut menunjukkan bahwa model dengan algoritma Gradient Boosting memiliki kemampuan akurasi prediktif lebih baik dibandingkan kedua model algoritma lainnya.



Gambar 16. Hasil Skor Evaluasi Model Menggunakan Data Training

Evaluasi pemodelan dengan menggunakan data test juga didapatkan hasil yang mendekati seperti pada uji evaluasi dengan data training. Evaluasi model pada data test di aplikasi Orange Datamining menggunakan widget: Predictions. Pada Gambar 17 merupakan hasil evaluasi model dengan menggunakan data test dimana terlihat bahwa secara umum model dengan algoritma Gradient Boosting memiliki nilai MSE, RMSE, MAE lebih rendah dan R2 yang lebih tinggi daripada model algoritma Support Vector Machine (SVM) dan Neural Network. Hal tersebut menunjukkan bahwa Gradient Boosting secara konsisten memiliki akurasi prediktif yang lebih baik daripada Support Vector Machine (SVM) dan Neural Network. Data test yang digunakan pada uji ini belum pernah dipakai untuk pembentukan pola model, jadi bisa dikatakan bahwa evaluasi pemodelan ini menggambarkan sifat generalisasi model algoritma untuk data baru. Pada uji data test ini untuk model algoritma Gradient Boosting mendapatkan skor akurasi MSE=0.087, RMSE=0.295, MAE=0.189, dan R2=0.976.



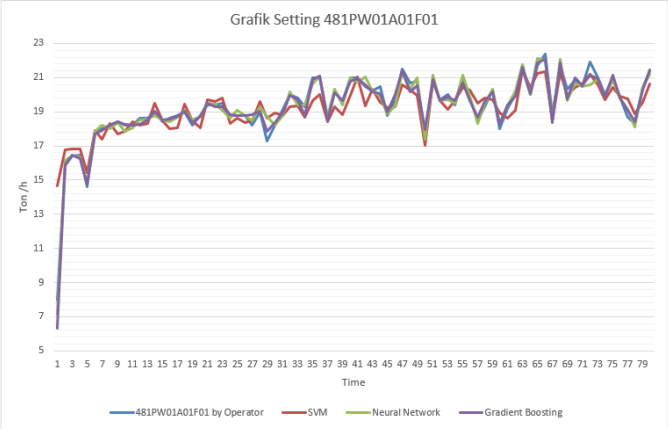
Gambar 17. Hasil Skor Evaluasi Model Menggunakan Data Test

Berdasarkan hasil evaluasi model algoritma diatas dengan menggunakan data training maupun data test maka dapat disimpulkan bahwa model machine learning dengan algoritma Gradient Boosting secara umum memiliki skor akurasi prediksi yang lebih baik

dibandingkan pemodelan dengan menggunakan algoritma Support Vector Machine (SVM) dan Neural Network.

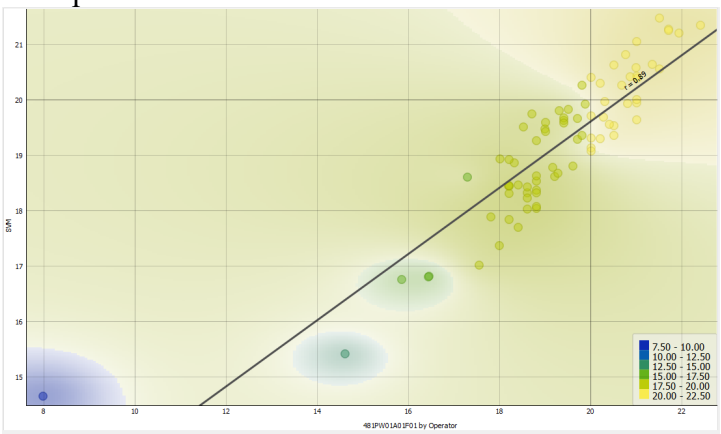
Implementasi Model Algoritma Machine Learning Untuk Penghematan Biaya Energi dan Reduksi Emisi Karbondioksida

Berdasarkan hasil score predictions pada data test, kita dapat membandingkan antara nilai setting Pulverized kiln (481PW01A01F01) yang dilakukan input oleh operator dengan nilai hasil prediksi masing-masing model algoritma. Hasil score prediction dilakukan eksport dahulu ke dalam file Microsoft excel dengan menggunakan widget: Save Data. Selanjutnya file export tersebut dilakukan analisa potensi penghematan dan nilai reduksi emisi karbondioksida.

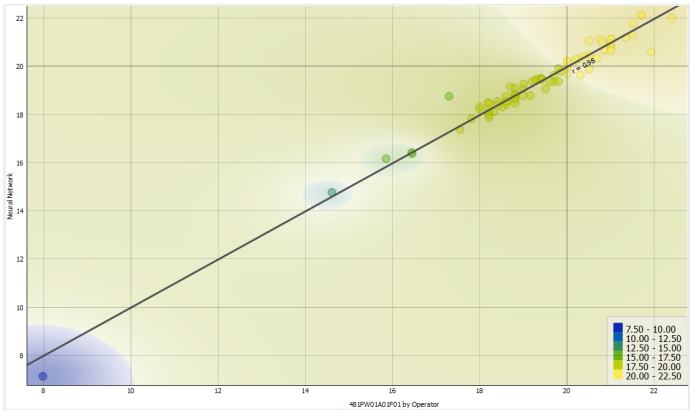


Gambar 18. Grafik Perbandingan Setting Nilai 481PW01A01F01

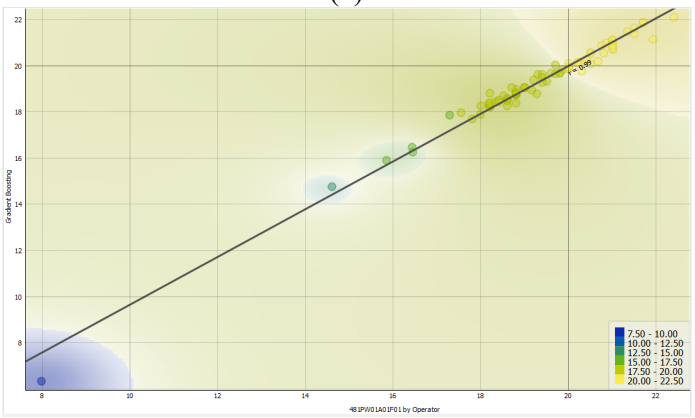
Pada gambar 18 dapat dilihat bahwa trending nilai setting 481PW01A01F01 yang dilakukan oleh operator, hasil prediksi model SVM, Neural Network, dan Gradient Boosting tidak terdapat perbedaan yang cukup jauh. Namun jika kita melihat nilai rata-rata nya, maka nilai rata-rata setting yang dimasukkan oleh operator adalah 19.28 ton/h. Sedangkan nilai rata-rata jika menggunakan nilai setting hasil prediksi pemodelan algoritma SVM, Neural network, dan Gradient Boosting adalah sebesar 19.19 ton/h, 19.26 ton/h, dan 19.25 ton/h. Berdasarkan hasil rata-rata nilai prediksi tersebut, terlihat bahwa ketiga model algoritma menghasilkan rata-rata nilai prediksi setting lebih rendah daripada input dari operator. Model SVM cenderung memiliki hasil nilai yang jauh lebih rendah dibandingkan kedua model algoritma lainnya. Namun hasil tersebut masih perlu dilakukan analisa terhadap korelasi / kedekatan terhadap intuisi operator dengan menggunakan scatter plot.



(a)



(b)



(c)

Gambar 19. Grafik Scatter Plot Nilai Setting 481PW01A01F01 oleh Operato Dibandingkan SVM (a), Neural Network (b), Gradient Boosting (c)

Pada Gambar 19 dapat terlihat bahwa model algoritma Gradient Boosting mampu mendekati nilai setting operator lebih baik daripada 2 model lainnya dengan nilai r = 0.99. Model Neural Network juga memiliki kedekatan yang cukup bagus dengan nilai sebesar r = 0.98. Sedangkan untuk model SVM hanya mampu mencapai r = 0.89. Analisa scatter plot ini diperlukan untuk mendukung hasil nilai prediksi pada Gambar 19 diatas. Walaupun model SVM memiliki nilai rata-rata prediksi paling rendah, namun untuk kedekatan terhadap intuisi operator masih jauh dibandingkan dengan kedua model algoritma lainnya. Model Gradient Boosting lebih direkomendasikan untuk digunakan dikarenakan memiliki nilai r=0.99 dan nilai rata-rata prediktif yang juga lebih rendah dibandingkan nilai setting operator dan Neural Network.

Tabel 2. Perbandingan Kalkulasi Biaya Penghematan dan Reduksi Emisi CO2 Dampak Penggunaan Machine Learning Dalam Penentuan Nilai Umpan Setting 481PW01A01F01

Parameter	*SVM*	*Neural Network*	*Gradient Boosting*
Selisih rata–rata nilai prediksi model terhadap rata-rata nilai input oleh operator(ton/h)	0.10	0.03	0.04
Penghematan biaya batubara per jam (Rp)	219,431	63,010	87,880
Penghematan biaya batubara setahun (Rp)	1,441,659,617	413,976,263	577,372,520

Parameter	*SVM*	*Neural Network*	*Gradient Boosting*
Reduksi emisi CO2 (kgCO2/year)	1,218,155	349,796	487,861
Potensi pengurangan pajak karbon setahun (Rp)	36,544,658	10,493,892	14,635,827
Potensi total penghematan biaya setahun (Rp)	1,478,204,275	424,470,155	592,008,347

Pada Tabel 2 terlihat bahwa dari selisih nilai rata-rata setting umpan 481PW01A01F01 pada masing-masing model algoritma terhadap nilai setting yang diinputkan oleh operator, maka kita turunkan menjadi biaya penghematan (cost saving) dengan asumsi harga batubara 150 USD/ton dan nilai kurs 15,300 Rp/USD. Penghematan terbesar adalah menggunakan model SVM. Namun mengingat nilai r pada regresi (Gambar4.14), maka rekomendasi penggunaan model Gradient Boosting dan Neural Network lebih direkomendasikan. Penghematan satu tahun didapatkan dengan asumsi periode operasi pabrik dilakukan sekitar 9 bulan selama satu tahun operasi. Reduksi emisi CO2 didapatkan dengan menggunakan standard perhitungan mengacu pada Global Cement and Concrete Association (GCCA)/ ECRA GmbH 2020. Dimana berdasarkan standard tersebut bahwa untuk menghitung nilai emisi menggunakan indeks 96 kgCO2/GJ untuk batubara, dimana nilai rata-rata Heating Value sebesar 20.2 GJ/ton. Sedangkan untuk nilai tarif pajak emisi CO2 berdasarkan UU 7/2021 tentang Harmonisasi Peraturan Perpajakan - Pasal 13 dimana tarif pajak karbon ditetapkan lebih tinggi atau sama dengan harga karbon di pasar karbon dengan tarif paling rendah Rp30,- per kilogram karbon dioksida ekuivalen (CO2e). Nilai potensi terbesar bisa dicapai dengan menggunakan model SVM. Namun untuk mendapatkan nilai kedekatan regresi yang paling baik terhadap nilai setting operator penggunaan model Gradient Boosting lebih direkomendasikan.

Kesimpulan

Berdasarkan analisa dan hasil penelitian yang dilakukan, dapat disimpulkan bahwa Model Algoritma machine learning yang digunakan untuk melakukan prediksi umpan Pulverized kiln (481PW01A01F01) pada sistem rotary kiln pabrik semen adalah Support Vector Machine (SVM), Gradient Boosting, dan Neural Network. Secara umum, variabel Pulverized kiln (481PW01A01F01) memiliki korelasi positif dengan beberapa variabel seperti Pressure Primary (481FN01N02P01), Silica modulus KF (QCX_KF_SIM), Feed (KILN_FEED_FM), dan Silica Clinker (QCX_Clinker_SIM), yang berarti bahwa ketika nilai variabel-variabel tersebut meningkat, maka nilai variabel 481PW01A01F01 cenderung meningkat juga. Sebaliknya, variabel 481PW01A01F01 memiliki korelasi negatif dengan beberapa variabel seperti Alumina modulus KF (QCX_KF_ALM), Press. Inlet kiln (441KL01MD02N01P01), SO3 clinker (QCX_Clinker_SO3), dan Temp. inlet kiln (441CI01N03T01) yang berarti bahwa ketika nilai variabel-variabel tersebut menurun, maka nilai variabel Pulverized kiln (481PW01A01F01) cenderung meningkat, Model Algoritma machine learning yang memiliki nilai akurasi terbaik untuk melakukan prediksi umpan Pulverized kiln (481PW01A01F01) pada sistem rotary kiln pabrik semen adalah Gradient Boosting dengan skor MSE=0.087, RMSE=0.295, MAE=0.189, dan R2=0.976, dimana parameter fungsi internal yang digunakan adalah methode: xgboost, Number of trees: 20, learning

rate: 0.299, lambda regularization: 0.6, limit depth of individual trees: 11, fraction of training instances: 1, fraction of features for each tree: 1, fraction of features for each level: 1, fraction of features for each split: 1, Berdasarkan uji datatest pada penelitian ini, maka didapatkan bahwa secara rata-rata hasil nilai prediktif ketiga model algoritma menunjukkan angka yang lebih rendah dibandingkan nilai rata-rata setting umpan Pulverized kiln (481PW01A01F01) yang dilakukan oleh operator. Hal tersebut akan memiliki dampak terhadap penghematan penggunaan batubara dan pengurangan emisi CO2.

BIBLIOGRAFI

Fatahi, R., Nasiri, H., Homafar, A., Khosravi, R., Siavoshi, H., & Chehreh Chelgani, S. (2023). Modeling operational cement rotary kiln variables with explainable artificial intelligence methods–a "conscious lab" development. *Particulate Science and Technology*, *41*(5). https://doi.org/10.1080/02726351.2022.2135470

Hu, W., & Mao, Z. (2024). A high accuracy method for the sintering condition recognition of rotary kiln. *Measurement: Journal of the International Measurement Confederation*, *229*. https://doi.org/10.1016/j.measurement.2024.114459

Hu, X. (2017). Support Vector Machine and Its Application to Regression and Classification. *Thesis*.

Kelleher, J. D., Namee, B. Mac, & D'Arcy, A. (2015). Fundamentals of Machine Learning for Predictive Data Analytics. In *Igarss 2014* (Issue 1).

Liu, Q., Yu, H., Wang, X., & Lu, S. (2020). Research on online monitoring of energy consumption of rotary kiln for cement clinker formation heat. *Proceedings of 2020 IEEE International Conference on Information Technology, Big Data and Artificial Intelligence, ICIBA 2020*. https://doi.org/10.1109/ICIBA50161.2020.9277235

Liu, X., Qiu, L., Rodriguez, J., Wu, W., Ma, J., Peng, Z., Wang, D., & Fang, Y. (2022). Data-Driven Neural Predictors-Based Robust MPC for Power Converters. *IEEE Transactions on Power Electronics*, *37*(10). https://doi.org/10.1109/TPEL.2022.3171100

Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in Neurorobotics*, *7*(DEC). https://doi.org/10.3389/fnbot.2013.00021

Okoji, A. I., Anozie, A. N., Omoleye, J. A., Taiwo, A. E., & Osuolale, F. N. (2022). Energetic assessment of a precalcining rotary kiln in a cement plant using process simulator and neural networks. *Alexandria Engineering Journal*, *61*(7). https://doi.org/10.1016/j.aej.2021.10.010

Qin, S. J., & Dong, Y. (2020). On data science for process systems modeling, control and operations. *IFAC-PapersOnLine*, *53*(2). https://doi.org/10.1016/j.ifacol.2020.12.540

Ramasamy, V., Kannan, R., Muralidharan, G., Sidharthan, R. K., Veerasamy, G., Venkatesh, S., & Amirtharajan, R. (2023). A comprehensive review on Advanced Process Control of cement kiln process with the focus on MPC tuning strategies. In *Journal of Process Control* (Vol. 121). https://doi.org/10.1016/j.jprocont.2022.12.002

Ramasamy, V., Sidharthan, R. K., Kannan, R., & Muralidharan, G. (2019). Optimal tuning of model predictive controller weights using genetic algorithm with

interactive decision tree for industrial cement kiln process. *Processes*, *7*(12). https://doi.org/10.3390/PR7120938

Shi, X., Huang, G., Hao, X., Yang, Y., & Li, Z. (2022). Sliding window and dual-channel CNN (SWDC-CNN): A novel method for synchronous prediction of coal and electricity consumption in cement calcination process. *Applied Soft Computing*, *129*. https://doi.org/10.1016/j.asoc.2022.109520

Zhang, L., Zhang, X., Chen, H., & Tang, H. (2019). A robust temperature prediction model of shuttle kiln based on ensemble random vector functional link network. *Applied Thermal Engineering*, *150*. https://doi.org/10.1016/j.applthermaleng.2018.12.092

Zhang, R., Lu, S., Yu, H., & Wang, X. (2021). Recognition method of cement rotary kiln burning state based on Otsu-Kmeans flame image segmentation and SVM. *Optik*, *243*. https://doi.org/10.1016/j.ijleo.2021.167418

Zhang, X., & Zhao, J. (2011). Prediction model for rotary kiln coal feed based on hybrid SVM. *Procedia Engineering*, *15*. https://doi.org/10.1016/j.proeng.2011.08.127

Zheng, J., Zhao, L., & Du, W. (2022). Hybrid model of a cement rotary kiln using an improved attention-based recurrent neural network. *ISA Transactions*, *129*. https://doi.org/10.1016/j.isatra.2022.02.018

Copyright holder:
Riduwan Maliki (2025)

First publication right:
Syntax Literate: Jurnal Ilmiah Indonesia

This article is licensed under:

