

PREDIKSI PELABELAN RATING AC EFISIENSI ENERGI MENGUNAKAN PEMODELAN MACHINE LEARNING

Desmarita Leni

Teknik Mesin, Fakultas Teknik, Universitas Muhammadiyah Sumatera Barat
Jl. Pasir Jambak No.4, Pasie Nan Tigo, Kec. Koto Tangah,
Kota Padang, Sumatera Barat 25586
*Email: desmaritaleni@gmail.com

Abstrak

Penelitian ini merancang sebuah pemodelan machine learning dengan menggunakan algoritma decision tree. Tujuan penelitian ini adalah untuk memprediksi tingkat efisiensi energi AC yang dijual dan tersedia dipasaran, Data yang digunakan pada penelitian diperoleh dari Direktorat Jenderal Energi Baru, Terbarukan, dan Konservasi Energi (EBTKE). Data tersebut mencakup informasi seperti daya, kapasitas pendingin, efisiensi, konsumsi energi tahunan, dan biaya listrik. Visualisasi data dilakukan dengan menggunakan boxplot dan regresi linear untuk mengevaluasi hubungan antara variabel dependen (rating) dan variabel independen. Hasil analisis mengindikasikan bahwa variabel efisiensi memiliki kontribusi yang signifikan terhadap rating, sebagaimana ditunjukkan oleh koefisien regresi linear sebesar 0.75. Pemodelan machine learning kemudian diuji dengan K-fold validation dengan jumlah 5 fold. Hasil evaluasi model menunjukkan nilai MAE sebesar 0.2, MSE sebesar 0.4, RMSE sebesar 0.63, dan tingkat akurasi sebesar 0.9. Berdasarkan hasil tersebut, dapat disimpulkan bahwa pemodelan machine learning menggunakan algoritma decision tree mampu memprediksi tingkat efisiensi AC dengan tingkat akurasi yang memuaskan.

Kata kunci: AC, efisiensi, energi, machine learning, pemodelan.

PENDAHULUAN

Peningkatan konsumsi energi untuk pendingin ruangan atau air conditioner (AC) di seluruh dunia menjadi salah satu faktor utama menyebabkan permintaan energi meningkat. Menurut prediksi dari Badan Energi Internasional (IEA), konsumsi energi untuk AC diperkirakan akan menyumbang sekitar 35% terhadap pertumbuhan permintaan energi pada tahun 2050 terutama di Asia Tenggara termasuk Indonesia (IEA, 2019). Prediksi ini sama dengan penelitian (Isaac, Morna; Van Vuuren, Detlef P, 2009) yang menyatakan bahwa, penggunaan AC mengakibatkan peningkatan emisi CO₂ sehingga mengakibatkan pemanasan iklim global, hal ini akan memicu peningkatan penggunaan AC terutama di Asia Tenggara sehingga mengakibatkan permintaan listrik yang besar. Badan Pusat Statistik (BPS) pada Maret tahun 2017 menyatakan bahwa 7,98% rumah tangga di Indonesia memiliki AC, angka tersebut diperkirakan akan terus meningkat di masa mendatang (BPS, 2017), Penggunaan AC yang tidak efisien akan menyebabkan peningkatan biaya listrik yang cukup signifikan bagi rumah tangga, serta memberikan dampak

negatif terhadap lingkungan. Banyaknya AC yang beredar di pasaran dengan berbagai merek dan tipe membuat konsumen tidak memperhatikan bahwa AC yang dibeli tidak efisiensi energi, jika hal ini terus dibiarkan maka akan berakibat fatal bagi konsumsi listrik di Indonesia.

Direktorat Jenderal Energi Baru, Terbarukan, dan Konservasi Energi (EBTKE) di Indonesia memperkenalkan sistem pelabelan energi untuk AC. Sistem ini bertujuan untuk membantu konsumen dalam memilih AC yang efisien, serta memberikan informasi yang transparan tentang tingkat efisiensi setiap AC yang dijual di pasaran. Sistem ini didasarkan pada beberapa variabel, diantaranya daya, kapasitas pendinginan, efisiensi, konsumsi listrik tahunan, dan biaya listrik. Mengingat Banyaknya model dan tipe yang bermunculan setiap waktu, tentu penyediaan data Base untuk referensi melihat label AC tidaklah efektif, terutama harus mencari secara manual di dalam data Base.

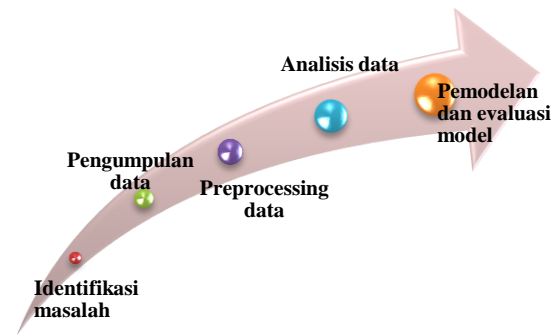
Di era digitalisasi saat ini terdapat banyak data dan informasi yang dapat diakses secara mudah dan cepat, hal ini membuat

machine learning menjadi metode yang populer dan efektif untuk melakukan pelabelan Energi (Tso, Geoffrey KF; Yau, Kelvin KW., 2007). *Machine learning* merupakan salah satu cabang dari kecerdasan buatan yang berfokus pada pembelajaran dari data, tanpa harus dilatih secara eksplisit. *Machine learning* memungkinkan sistem untuk belajar dan meningkatkan performa secara otomatis tanpa perlu dilatih secara manual (Rocha, Felipe, et al., 2021.) Banyak metode machine learning yang digunakan untuk memprediksi konsumsi energi serta pelabelan peralatan elektronik termasuk AC seperti (Wang, Weiqi; Zhou, Zixuan; Lu, Zhongming, 2022) menggunakan metode *machine learning* berbasis data untuk mendeteksi efisiensi penggunaan AC berdasarkan ruangan, (Duarte, Grasiela Regina, et al., 2017) membandingkan berbagai algoritma *machine learning* untuk memprediksi konsumsi energi berdasarkan arsitektur bangunan, (Hettinga, Sanne; Van't Veer, Rein; Boter, Jaap, 2023) menggunakan kumpulan data terbuka untuk pemodelan *machine learning* sebagai pelabelan bangunan, industri, dan daerah efisiensi energi.

Upaya mengembangkan konteks pelabelan Energi, *machine learning* dapat digunakan untuk mengidentifikasi pola dan hubungan dalam data yang dikumpulkan (Khayatian, Fazel, et al., 2016) Penggunaan *machine learning* dapat menjadi solusi yang efektif dalam melakukan pelabelan efisiensi energi AC, salah satu Metode *machine learning* yang dapat digunakan adalah algoritma *decision trees*, dimana algoritma ini merupakan salah satu metode pembelajaran mesin yang menggunakan struktur tree untuk membuat keputusan sehingga membantu dalam menganalisis data dengan mengelompokkan data ke dalam kelompok yang memiliki karakteristik yang sama (Yu, Zhun, et al, 2010). Berdasarkan permasalahan yang telah diuraikan sebelumnya, penelitian ini merancang sebuah pemodelan machine learning menggunakan algoritma *decision tree* untuk memprediksi Rating AC dan menganalisis korelasi antara rating AC dengan variabel-variabel seperti daya, kapasitas pendinginan, efisiensi, konsumsi energi tahunan, dan biaya listrik. Pemodelan *machine learning* dalam penelitian ini dibuat menggunakan bahasa pemrograman Python.

METODE PENELITIAN

Metode Penelitian dari permasalahan yang telah diuraikan di bagian pendahuluan, maka untuk memperoleh solusi dari masalah tersebut agar tujuan tercapai disusunlah skema dan tahapan penelitian seperti Gambar 1.



Gambar 1. Diagram alir Penelitian

Penelitian ini bertujuan untuk mengidentifikasi dan memprediksi rating AC berdasarkan efisiensi energi dari beberapa variabel seperti daya, kapasitas pendinginan, efisiensi, konsumsi listrik tahunan, dan biaya listrik. Data penelitian ini diperoleh dari Direktorat Jenderal Energi Baru, Terbarukan, dan Konservasi Energi (EBTKE) pada bulan Juli 2022 dan terdiri dari 400 data AC non-inverter yang berbagai model dan tipe. Metode yang digunakan adalah pemodelan machine learning menggunakan algoritma *decision tree*.

Tahap *preprocessing* data meliputi pembersihan data, perubahan data menjadi bentuk yang sesuai dengan kebutuhan analisis, dan penghapusan data yang tidak valid. Analisis data dilakukan dengan menggunakan linear regresi dan korelasi Pearson untuk melihat korelasi antara variabel dependen (rating AC) dengan variabel independen seperti Daya, Kapasitas Pendinginan, Efisiensi, Konsumsi Energi tahunan, dan biaya listrik. Setelah analisis data, dilakukan pemodelan machine learning menggunakan algoritma *decision tree*. Model di evaluasi dengan *k-fold validation* dan hasil

prediksi machine learning di evaluasi dengan melihat nilai MAE, MSE, RMSE, dan Accuracy.

Hasil analisis dan pemodelan diinterpretasi untuk melihat pola data, korelasi antar variabel, dan hasil prediksi model. Hasil analisis kemudian divalidasi dengan data dari sumber lain. Berdasarkan hasil analisis dan interpretasi data, dapat ditarik kesimpulan mengenai pelabelan energi bagi AC berdasarkan variabel-variabel yang telah ditentukan.

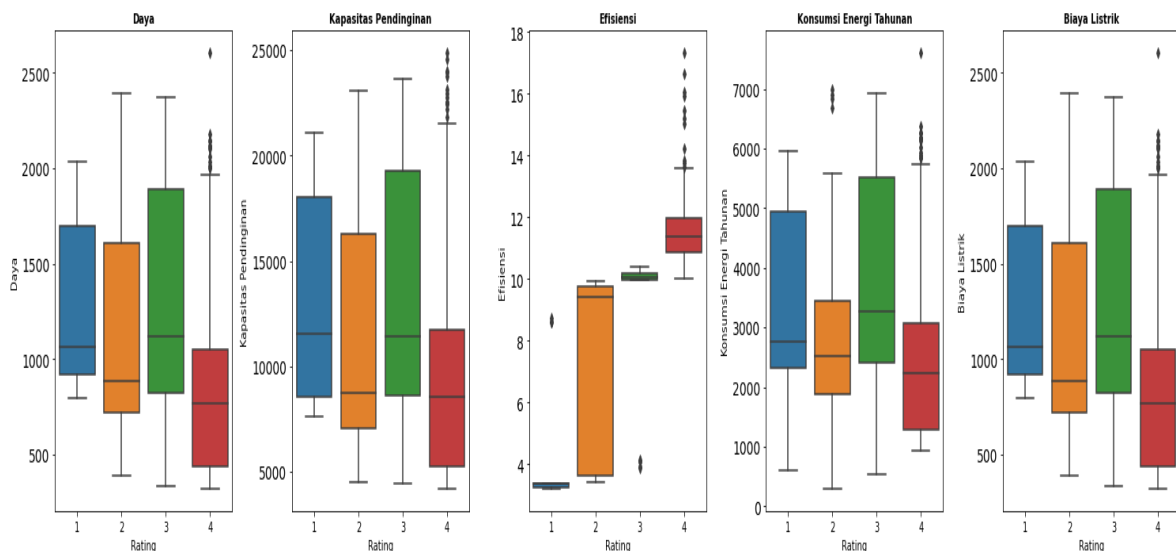
Data penelitian ini diperoleh dari database Direktorat Jenderal Energi Baru, Terbarukan, dan Konservasi Energi (EBTKE). Data yang digunakan dalam penelitian ini terdiri dari berbagai model dan tipe AC yang tersedia di pasaran. Data ini berjumlah 400 data dan setelah dilakukan *preprocessing* data tersebut berkurang menjadi 396 data, hal ini karena data tersebut memiliki biaya listrik tahunan 0, sehingga dihilangkan dari dataset. Dataset pelabelan rating efisiensi AC ini terdiri dari Daya, Kapasitas Pendinginan, Efisiensi, Konsumsi Energi Tahunan, Biaya Listrik, dan Rating.

Setelah data dibersihkan, langkah selanjutnya adalah analisis data, untuk mempermudah melakukan analisis data maka setiap variabel pada dataset divisualisasikan seperti variabel dependen (Rating) dengan setiap variabel independen (Daya, Kapasitas Pendinginan, Efisiensi, Konsumsi Energi

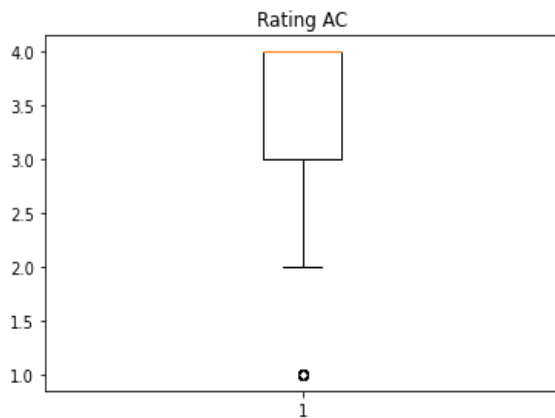
Tahunan, dan Biaya Listrik). Penggunaan visualisasi *boxplot* dapat membantu dalam mengetahui distribusi data serta mencari titik outlier yang terjadi. *Boxplot* juga dapat digunakan untuk membandingkan distribusi data dari beberapa kelompok data. Menggunakan *boxplot* dapat memberikan informasi mengenai nilai minimum, nilai maksimum, nilai median, dan nilai terpusat lainnya yang terdapat dalam data (Thirumalai, Chandrasegar; kanimozhi, r. vaishnavi, B, 2017), hal ini dapat membantu dalam mengidentifikasi pola dan trend yang terjadi pada data yang dianalisis, seperti yang terlihat pada Gambar 2 dan Gambar 3.

HASIL DAN PEMBAHASAN

Setelah melihat distribusi data pada masing-masing variabel dataset maka langkah selanjutnya adalah melihat korelasi antara variabel dependen dengan variabel independen menggunakan regresi linear. Persamaan koefesien linear regresi (atau biasa disebut sebagai persamaan regresi linear) adalah sebuah model matematika yang menggambarkan hubungan linier antara satu variabel dependen (Y) dengan satu atau lebih variabel independen (X).



Gambar 2. pola dan trend data rating AC



Gambar 3. Visualisasi variabel dependen menggunakan *boxplot*

Persamaan regresi linear dapat dituliskan sebagai berikut (Yan, Xin; Su, Xiaogang, 2009)

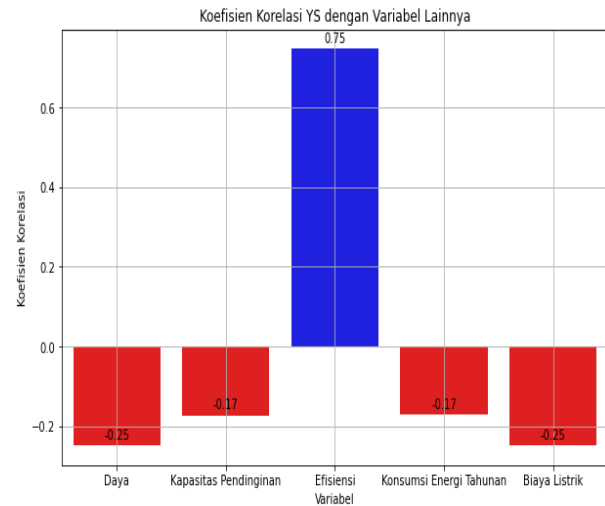
$$Y = a + bX \quad (1)$$

Variabel dependen (Rating) Y
 X=Variabel independen (Variabel yang mempengaruhi rating)

a = *Intercept* (nilai Y ketika X = 0)

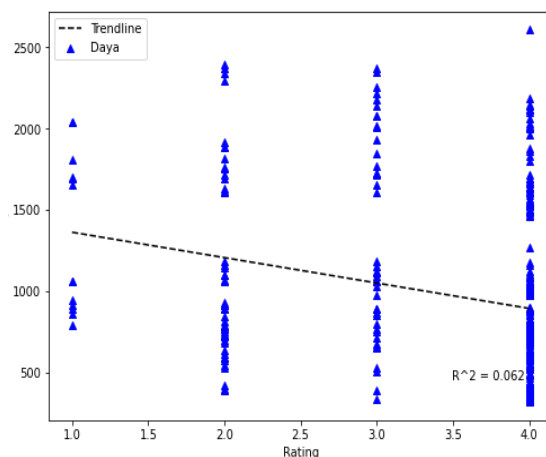
b = Slope atau koefisien regresi (perubahan Y per satuan perubahan X)

Dari persamaan di atas, dapat dilihat bahwa hubungan linier antara Y dengan X dapat ditunjukkan dengan koefisien regresi (b). Nilai koefisien regresi positif menunjukkan bahwa Y akan meningkat seiring dengan peningkatan X, sedangkan nilai koefisien regresi negatif menunjukkan bahwa Y akan menurun seiring dengan peningkatan X. Menggunakan persamaan regresi linear tersebut didapatkan hubungan antara rating dengan variabel lainya seperti yang terlihat pada Gambar 4. Berdasarkan Gambar 4, dapat dilihat bahwa koefesien regresi linear untuk variabel Efisiensi adalah 0.75, artinya jika Efisiensi bertambah 1 unit, Rating akan bertambah 0.75 unit. Sedangkan untuk variabel Daya, Kapasitas Pendinginan, Konsumsi Energi Tahunan, dan Biaya Listrik, koefesien regresi linearnya bernilai negatif. Ini berarti jika salah satu dari variabel tersebut bertambah 1 unit, Rating akan berkurang sebesar nilai koefesien regresi linear tersebut. Koefesien regresi linear negatif ini menunjukkan bahwa terdapat hubungan negatif antara variabel tersebut dengan Rating.

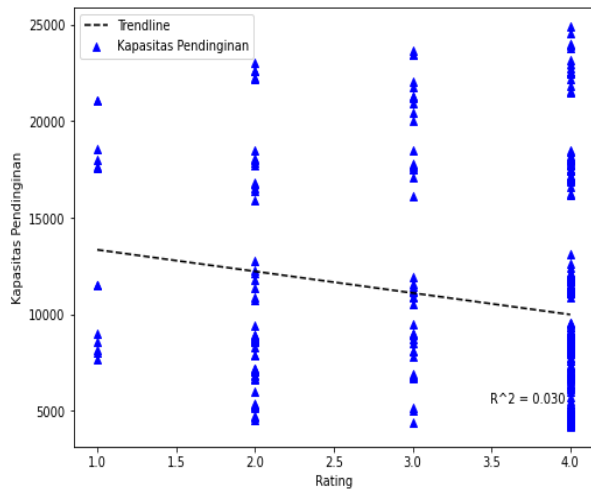


Gambar 4. Grafik batang koefesien regresi linear

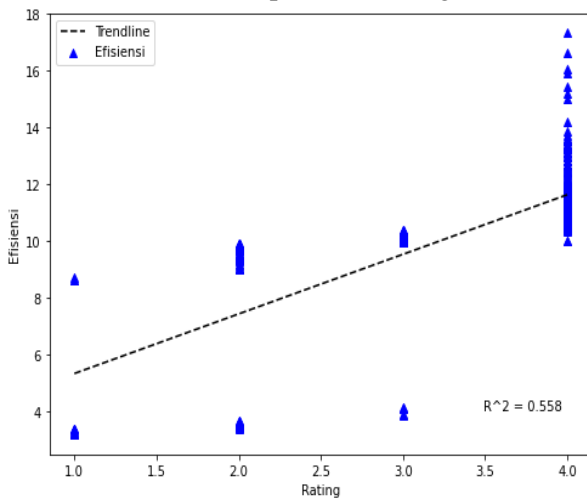
Berdasarkan hasil koefesien linear regresi, kemudian dilakukan visualisasi antara variabel dependen (Rating) dengan variabel independen (Daya, Kapasitas Pendinginan, Efisiensi, Konsumsi Energi Tahunan, dan Biaya Listrik) menggunakan *scatter plot*, untuk melihat seberapa besar pengaruh dari masing-masing variabel independen terhadap variabel dependen, serta seberapa baik model regresi linear mampu menjelaskan variasi dari variabel dependen tersebut. Visualisasi menggunakan scatterplot, dapat memperoleh informasi tentang korelasi antara variabel dependen dan independen, serta tingkat konsistensi atau ke tidak konsistenan data.



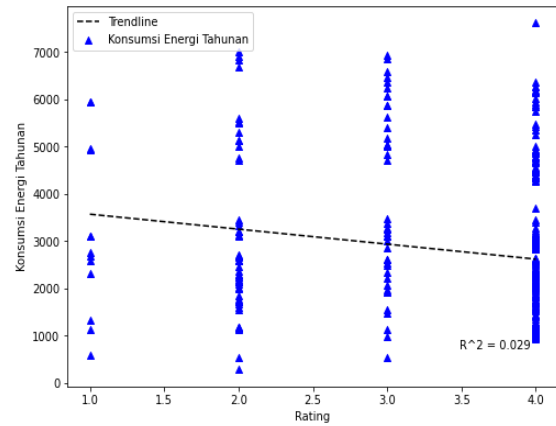
(a) Daya



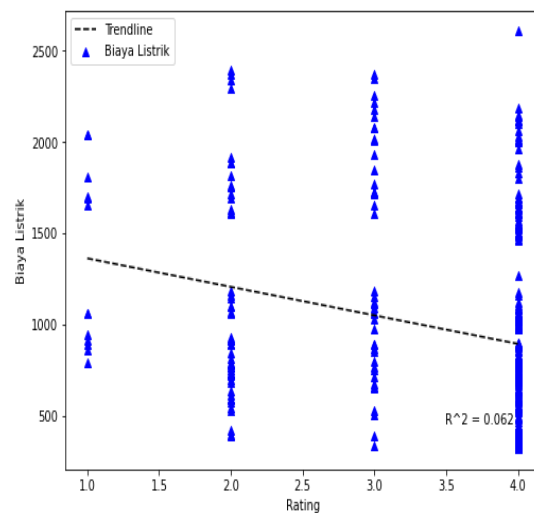
(b) Kapasitas Pendinginan



(c) Efisiensi



(d) Kosumsi Energi Tahunan



(e) Biaya Listrik

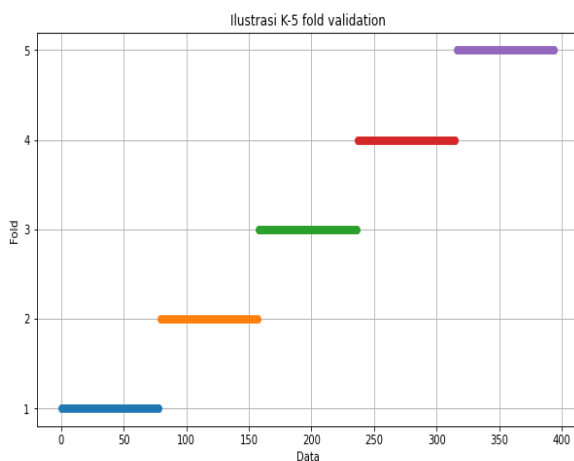
Gambar 5. Visualisasi menggunakan scatter plot dan R-squared (a,b,c,d,e)

Jika terdapat korelasi yang kuat antara variabel dependen dan independen, maka kita dapat menyimpulkan bahwa model regresi linear tersebut mampu menjelaskan variasi dari variabel dependen dengan baik. Sebaliknya, jika korelasi yang terjadi kurang kuat atau tidak konsisten, maka model regresi linear tersebut tidak dapat diandalkan untuk menjelaskan variasi dari variabel dependen tersebut. Pengamatan yang terlihat pada Gambar 5, sama halnya dengan hasil koefisien linear regresi dimana variabel Efisiensi memiliki nilai *R-squared* sebesar 0.558, yang menunjukkan bahwa variabel Efisiensi memiliki pengaruh yang cukup kuat terhadap variabel Rating, sedangkan variabel lainnya memiliki pengaruh kecil terhadap rating.

Pemodelan dan Evaluasi model

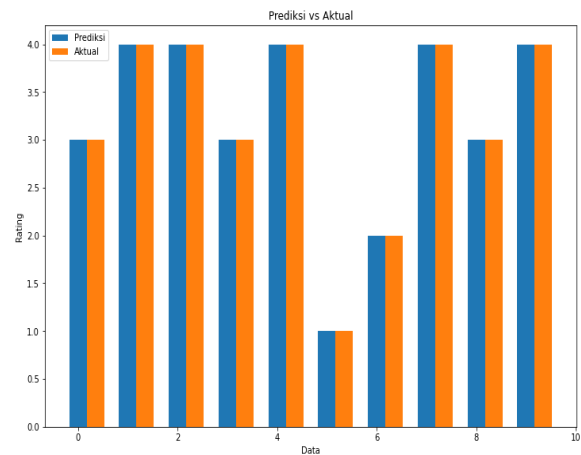
Setelah melihat hubungan antara data, langkah selanjutnya adalah membuat pemodelan *machine learning* menggunakan algoritma *decision tree*. Pemodelan ini dibuat dengan menggunakan bahasa pemrograman python yang dijalankan di *google colab*. Pemodelan *machine learning* memiliki lima variabel input yang terdiri dari daya, kapasitas pendinginan, efisiensi, konsumsi energi tahunan dan biaya listrik, dan memiliki satu output yaitu Rating AC.

Pemodelan ini menggunakan *K-fold validation* dengan jumlah fold sebanyak 5 dilakukan dengan cara membagi data menjadi 5 bagian yang disebut *fold*. Setiap *fold* kemudian digunakan sebagai data validasi sementara sisanya digunakan sebagai data training (Leni D, Sumiati R., 2022). Proses ini diulang sebanyak 5 kali seperti pada Gambar 6, dengan setiap *fold* digunakan sebagai data validasi pada setiap iterasinya. Proses ini dilakukan beberapa kali hingga semua *fold* digunakan sebagai data uji. Akhirnya, hasil evaluasi dari setiap fold dijumlahkan dan dibagi dengan jumlah *fold* yang ada untuk mendapatkan hasil evaluasi akhir, dengan menggunakan *K-fold validation*, model akan lebih terlatih dengan berbagai data latih yang berbeda sehingga hasil evaluasinya akan lebih akurat.



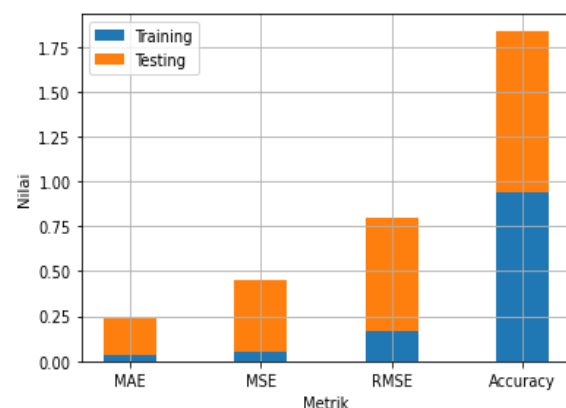
Gambar 6. Ilustrasi penggunaan K-5 fold validation

Pemodelan *machine learning* menggunakan algoritma *decision tree*, untuk prediksi Rating AC ini di uji dengan 10 data yang tidak ada didalam dataset. Perlu diketahui nilai target atau variabel rating memiliki tipe data ordinal dimana, tipe data ordinal adalah tipe data yang memiliki urutan atau ranking, tetapi tidak memiliki jarak yang sama antar nilai, dalam hal ini nilai 1, 2, 3, dan 4 merupakan tipe data ordinal karena memiliki urutan atau ranking, tetapi tidak memiliki jarak yang sama antar nilai sehingga sehingga hasil prediksi yang diperoleh selalu bilangan bulat. Hasil prediksi menggunakan data uji dapat dilihat pada Gambar 7.



Gambar 7. Grafik Hasil prediksi dengan nilai sebenarnya

Data uji ini di inputkan kedalam model dan di evaluasi dengan menggunakan metrik evaluasi seperti MAE, MSE, RMSE, dan *Accuracy*. Hasil perbandingan evaluasi data training dan data testing dapat dilihat pada Gambar 6, dimana evaluasi model dengan menggunakan data training diperoleh nilai MAE 0.033, MSE 0.0482, RMSE 0.1553, dan *Accuracy*: 0.9492 sedangkan data testing menghasilkan nilai MAE 0.2, MSE 0.4, RMSE 0.63, dan *Accuracy*: 0.9. Berdasarkan hasil tersebut, dapat disimpulkan bahwa model yang telah dibangun memiliki kemampuan yang baik dalam memprediksi nilai Rating dengan menggunakan data training, diikuti oleh data testing.



Gambar 8. perbandingan nilai metrik evaluasi data training dan data testing.

Hal ini ditunjukkan oleh nilai MAE, MSE, RMSE, dan Accuracy yang relatif kecil pada data training, dan sedikit lebih besar pada data testing. Namun demikian, nilai-nilai tersebut masih dapat dikatakan cukup baik untuk tujuan pemodelan yang telah ditetapkan.

Metrik-metrik tersebut digunakan untuk mengukur seberapa besar *error* yang terjadi antara nilai prediksi dengan nilai aktual seperti, MAE (*Mean Absolute Error*) merupakan metrik evaluasi yang menghitung rata-rata dari selisih antara nilai prediksi dan nilai sebenarnya. Nilai MAE semakin kecil, semakin baik kualitas model. MSE (*Mean Squared Error*) merupakan metrik evaluasi yang menghitung rata-rata dari kuadrat selisih antara nilai prediksi dan nilai sebenarnya. Nilai MSE semakin kecil, semakin baik kualitas model. RMSE (*Root Mean Squared Error*) merupakan metrik evaluasi yang merupakan akar dari MSE. Nilai RMSE semakin kecil, semakin baik kualitas model. RMSE lebih sering digunakan karena skalanya lebih sesuai dengan skala data sebenarnya, sedangkan *Accuracy* adalah sebuah metrik evaluasi yang digunakan untuk mengukur seberapa baik suatu model dalam memprediksi data (Ağbulut, Ümit, Gürel, Ali Etem; Biçen, Yunus., 2021).

PENUTUP

Kesimpulan

Hasil dari penelitian ini, didapatkan bahwa model machine learning yang digunakan untuk memprediksi rating AC berdasarkan variabel-variabel independen seperti daya, kapasitas pendinginan, efisiensi, konsumsi energi tahunan, dan biaya listrik memberikan hasil evaluasi yang cukup baik. Hal ini dapat dilihat dari hasil evaluasi model menggunakan metrik MAE, MSE, RMSE, dan accuracy yang menunjukkan bahwa model tersebut mampu memberikan prediksi yang akurat dengan nilai yang cukup kecil. Selain itu, hasil analisis regresi linear menunjukkan bahwa variabel efisiensi memiliki hubungan yang paling kuat dengan rating AC, dimana semakin tinggi efisiensi AC maka semakin tinggi pula rating yang diperoleh. Namun, terdapat beberapa variabel lainnya seperti daya dan biaya listrik yang memiliki hubungan negatif dengan rating AC, sehingga semakin tinggi nilai kedua variabel tersebut maka semakin rendah pula rating yang diperoleh. Dari hasil tersebut, dapat disimpulkan bahwa variabel efisiensi merupakan faktor yang

paling berpengaruh dalam menentukan rating AC.

DAFTAR PUSTAKA

- Ağbulut, Ümit, Gürel, Ali Etem; Biçen, Yunus. (2021). Prediction of daily global solar radiation using different machine learning algorithms: Evaluation and comparison. *Renewable and Sustainable Energy Reviews*, 135: 110114.
- BPS. (2017). Diambil kembali dari <https://lokadata.beritagar.id/chart/preview/rumah-tangga-memiliki-ac-2017-1531737620>
- Duarte, Grasiela Regina, et al. (2017). Comparison of machine learning techniques for predicting energy loads in buildings. *Ambiente Construído*, 17.3: 103-115.
- Hettinga, Sanne; Van't Veer, Rein; Boter, Jaap. (2023). Large scale energy labelling with models: The Eu Tabula model versus machine learning with open data. *Energy*, 264: 126175.
- IEA. (2019). *the-future-of-cooling-in-southeast-asia*. Diambil kembali dari <https://www.iea.org/reports/the-future-of-cooling-in-southeast-asia>
- IEA. (2022). *future-of-cooling-in-southeast-asia*. Diambil kembali dari <https://www.iea.org/reports/the-future-of-cooling-in-southeast-asia>
- Isaac, Morna; Van Vuuren, Detlef P. (2009). Modeling global residential sector energy demand for heating and air conditioning in the context of climate change. *Energy policy*, 37.2: 507-521.
- Khayatian, Fazel, et al. (2016). Application of neural networks for evaluating energy performance certificates of

- residential buildings. *Energy and Buildings*, 125: 45-54.
- Leni D, Sumiati R. (2022). Perbandingan Algoritma Machine Learning Untuk Prediksi Sifat Mekanik Pada Baja Paduan Rendah. *Jurnal Rekayasa Material, Manufaktur dan Energi*, 5(2):167-74.
- Rocha, Felipe, et al. (2021.). Evaluating Machine Learning Classifiers for Prediction in an IoT-based Smart Building System. *In: 2021 IEEE 7th World Forum on Internet of Things (WF-IoT). IEEE*, (hal. 563-568).
- Thirumalai, Chandrasegar; kanimozhi, r. vaishnavi, B. (2017). Data analysis using box plot on electricity consumption. *Communication and Aerospace Technology (ICECA). IEEE*, (hal. 598-600).
- Tso, Geoffrey KF; Yau, Kelvin KW. (2007). Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks. *Energy*, 32.9: 1761-1768.
- Wang, Weiqi; Zhou, Zixuan; Lu, Zhongming. (2022). Data-driven assessment of room air conditioner efficiency for saving energy. *Journal of Cleaner Production*, 338: 130615.
- Yan, Xin; Su, Xiaogang. (2009). *Linear regression analysis: theory and computing*.
- Yu, Zhun, et al. (2010). A decision tree method for building energy demand modeling. *Energy and Buildings*, 42.10: 1637-1646.