

KOMPARASI ALGORITMA *MACHINE LEARNING* PADA IDENTIFIKASI VOLUME PRODUKSI PERIKANAN TANGKAP LAUT DI INDONESIA

*(Comparison of Machine Learning Algorithms on Identification of Marine Capture
Fisheries Production Volume in Indonesia)*

Hikmal Akbar Fitransyah*, Iqlima Yasmin Fadhilah, dan Najwa Elzahra

Universitas Pendidikan Indonesia, Jl. Dr. Setiabudi No. 229, Isola, Kec. Sukasari, Kota
Bandung, Jawa Barat 40154, Indonesia
e-mail: hikmal2601@upi.edu

ABSTRACT

This study compares three Machine Learning algorithms - Linear Regression, Random Forest, and Decision Tree - to identify the volume of marine capture fisheries production in Indonesia from 2020 to 2023, measuring performance using the coefficient of determination (R^2). Results showed that Decision Tree achieved the highest R^2 value consistently at 1.0 every year, indicating very high accuracy. Random Forest also performed very well with an R^2 close to 1.0, which was 0.998 in 2020, 0.995 in 2021, 0.995 in 2022, and 0.994 in 2023. Linear Regression, although with a lower R^2 value than the other two algorithms, showed quite good accuracy, namely 0.949 in 2020, 0.903 in 2021, 0.907 in 2022, and 0.871 in 2023. Based on this comparison, Decision Tree and Random Forest proved to be superior in predicting the volume of marine capture fisheries production in Indonesia, with high accuracy and robustness to data variations. These algorithms are recommended to be used in fisheries production volume analysis, which can support the development of sustainable capture fisheries management strategies.

Keywords: *Machine Learning, Fisheries, Volume.*

ABSTRAK

Penelitian ini membandingkan tiga algoritma *Machine Learning*, *Linear Regression*, *Random Forest*, dan *Decision Tree* untuk mengidentifikasi volume produksi perikanan tangkap laut di Indonesia dari tahun 2020 hingga 2023, dengan mengukur kinerja menggunakan koefisien determinasi (R^2). Hasil menunjukkan bahwa *Decision Tree* mencapai nilai R^2 tertinggi secara konsisten, yaitu 1.0 setiap tahun, mengindikasikan akurasi yang sangat tinggi. *Random Forest* juga menunjukkan performa yang sangat baik dengan R^2 mendekati 1.0, yaitu 0.998 pada tahun 2020, 0.995 pada tahun 2021, 0.995 pada tahun 2022, dan 0.994 pada tahun 2023. *Linear Regression*, meskipun dengan nilai R^2 lebih rendah dibandingkan kedua algoritma lainnya, menunjukkan ketepatan yang cukup baik, yaitu 0.949 pada tahun 2020, 0.903 pada tahun 2021, 0.907 pada tahun 2022, dan 0.871 pada tahun 2023. Berdasarkan komparasi ini, *Decision Tree* dan *Random Forest* terbukti unggul dalam prediksi volume produksi perikanan tangkap laut di Indonesia, dengan akurasi tinggi dan ketahanan terhadap variasi data. Algoritma-algoritma ini direkomendasikan untuk digunakan dalam analisis volume produksi perikanan, yang dapat mendukung pengembangan strategi pengelolaan perikanan tangkap yang berkelanjutan.

Kata kunci: *Machine Learning, Perikanan, Volume.*

PENDAHULUAN

Negara Kesatuan Republik Indonesia (NKRI) merupakan salah satu negara kepulauan terbesar di Asia serta kedua di dunia. Indonesia memiliki 17.506 pulau, baik besar maupun kecil, dengan luas wilayah lebih dari 7,7 juta km². Dari total luas tersebut, sekitar dua pertiga merupakan perairan yang mencakup lebih dari 5,8 juta km² (Maryani et al., 2019). Secara potensial, perikanan Indonesia merupakan yang terbesar di dunia, baik dalam sektor perikanan tangkap maupun budidaya (Wantimpres RI, 2017). Menurut data dari Kementerian Kelautan dan Perikanan, potensi sumber daya perikanan Indonesia diperkirakan bisa mencapai 12,01 juta ton per tahun, dengan Jumlah Tangkapan Ikan yang Diperbolehkan (JTB) sebanyak 8,6 juta ton per tahun. Data ini mencakup 11 Wilayah Pengelolaan Perikanan Negara Republik Indonesia (WPPNRI) dan dipecah menjadi sembilan kelompok sumber daya ikan, termasuk ikan demersal, ikan karang, pelagis kecil, cumi-cumi, udang penaeid, lobster, rajungan, kepiting, dan pelagis besar (Evan, 2022)

Berdasarkan metode operasional atau cara produksinya, perikanan dibagi menjadi dua kategori, yaitu perikanan tangkap (*capture fisheries*) dan perikanan budidaya (*aquaculture*), dengan potensi produksi berkelanjutan sekitar 67 juta ton per tahun. Dari jumlah tersebut, potensi produksi berkelanjutan (*Maximum Sustainable Yield = MSY*) untuk perikanan tangkap laut mencapai 9,3 juta ton per tahun, sementara perikanan tangkap di perairan darat (seperti danau, sungai, waduk, dan rawa) sekitar 0,9 juta ton per tahun, sehingga total perikanan tangkap mencapai 10,2 juta ton per tahun. Adapun sisanya, yaitu 56,8 juta ton per tahun, merupakan potensi perikanan budidaya, yang mencakup budidaya laut (*mariculture*), budidaya di perairan payau (tambak), dan budidaya di perairan tawar (darat) (Wantimpres RI, 2017).

Meskipun memiliki potensi yang besar, pengumpulan dan analisis data produksi perikanan di Indonesia masih menghadapi kendala. Salah satu contohnya di Pelabuhan Perikanan Samudera Nizam Zachman Jakarta yang jumlah hasil tangkapannya belum didukung oleh sistem pendataan hasil tangkapan yang efektif. Sekitar 80% dari data hasil tangkapan tidak sesuai dengan data sebenarnya di lapangan (Taufik, 2019). Ketidakakuratan dalam data hasil tangkapan tersebut disebabkan oleh terbatasnya sumber daya manusia, kurangnya keterampilan dan ketelitian dalam mendata ikan hasil tangkapan (Faisal, 2019). Berkaca dari lemahnya sistem pengelolaan data tersebut, dibutuhkan suatu peningkatan sistem pengumpulan data untuk mendukung pengelolaan yang lebih baik. Dengan kemajuan teknologi informasi dan metode analisis data, *Machine Learning* dapat dijadikan solusi untuk

meningkatkan kualitas pengumpulan dan analisis data kelautan khususnya dalam mengidentifikasi volume produksi perikanan tangkap laut di Indonesia.

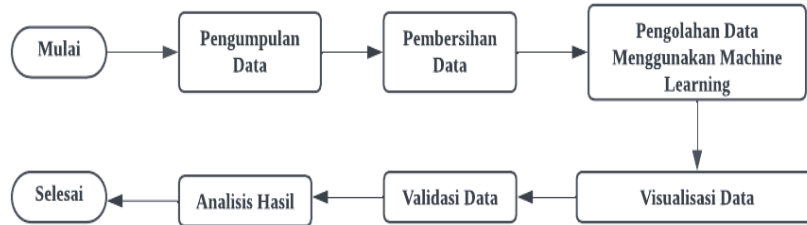
Machine Learning mengacu pada kumpulan algoritma komputer yang dirancang untuk meningkatkan kemampuan sistem melalui analisis pola dari data yang ada. Fokus utama dari *Machine Learning* adalah kemampuannya untuk menyesuaikan dan beradaptasi dengan perubahan guna merespons transformasi yang terjadi (Pambudi et al., 2020). *Machine Learning* sendiri termasuk kecerdasan yang ada dalam *Data Mining* (Rizky et al., 2024). *Machine Learning* juga merupakan sebuah teknik dalam ilmu komputer yang bertujuan mengekstraksi pola tersembunyi untuk menghasilkan wawasan baru dari sekumpulan data (Mai et al., 2022). Beberapa algoritma yang populer di antaranya adalah *Random Forest*, *Decision Tree*, dan *Linear Regression*, yang akan digunakan dalam penelitian ini untuk membandingkan tingkat akurasi masing-masing.

Decision Tree adalah salah satu metode klasifikasi yang sangat populer karena mudah dipahami dan diinterpretasikan oleh manusia (Ariawan et al., 2021). Metode ini efektif untuk mengenali pola serta termasuk dalam teknik pengenalan pola statistik (Patlisan, & Rusdah. 2023). *Decision Tree* berfungsi sebagai model prediktif yang dapat digunakan baik untuk klasifikasi maupun regresi dalam proses pengambilan keputusan (Rokach, L., & Maimon, O., 2007). Selain itu, algoritma *Random Forest* juga sering digunakan untuk mengklasifikasikan objek dengan tingkat akurasi yang tinggi (Ariawan, I. et al., 2021). Di sisi lain, *Linear Regression* merupakan salah satu model prediksi yang banyak digunakan dalam *Data Mining* untuk menganalisis hubungan antara variabel bebas (independen) dan variabel tidak bebas (dependen) (Sholeh, M., Nurnawati, E. K., & Lestari, U. 2023). *Linear Regression* menggunakan garis lurus untuk menggambarkan hubungan antara dua variabel, di mana satu variabel berperan sebagai variabel respons (tergantung) dan variabel lainnya sebagai variabel penjelas (independen) (Rahma, C. A. 2022). Penelitian ini bertujuan menerapkan ketiga algoritma tersebut untuk mengidentifikasi volume produksi perikanan tangkap di Indonesia. Hasil evaluasi akurasi model akan disajikan melalui pengukuran *Mean Squared Error* (MSE) dan *R-Squared* (R^2), serta dilengkapi dengan visualisasi prediksi untuk memperjelas performa masing-masing model yang diuji.

METODE PENELITIAN

Penelitian ini secara langsung menerapkan berbagai algoritma *Machine Learning* untuk mengidentifikasi volume produksi perikanan tangkap, khususnya di Indonesia dalam

beberapa tahun terakhir. Tahapan dalam penelitian ini mencakup pengumpulan data, pemrosesan data, hingga evaluasi performa algoritma melalui eksperimen untuk mendapatkan model yang paling efektif dalam mengidentifikasi volume produksi perikanan di Indonesia.



Gambar 1. Tahapan penelitian

Atribut variabel pada data ini antara lain adalah Provinsi, Tahun, dan Volume Produksi. Algoritma yang digunakan, termasuk *Linear Regression*, *Decision Tree*, dan *Random Forest*, akan dilatih dan diuji untuk memprediksi volume produksi, dengan evaluasi akurasi menggunakan *Mean Squared Error* (MSE) dan *R-Squared* (R^2). Hasil evaluasi tiap model dibandingkan untuk menentukan model terbaik. Visualisasi prediksi juga dilakukan untuk melihat performa model.

Pengumpulan Data

Dataset diperoleh dari Statistik KKP, *website* resmi data kelautan dan perikanan milik Kementerian Kelautan dan Perikanan yang mencakup volume produksi perikanan tangkap laut per provinsi di Indonesia untuk periode tahun 2020 hingga 2023.

Tabel 1. Dataset produksi perikanan tangkap laut per provinsi

No.	Provinsi	Volume (Ton)	Tahun
1.	Aceh	304283,44	2020
2.	Bengkulu	68070,41	2020
3.	Jambi	47489,28	2020
4.	Kepulauan Bangka Belitung	225558	2020
5.	Kepulauan Riau	332175,96	2020
...	...	...	...
132	Sulawesi Utara	252335,18	2023
133	Maluku	506385,74	2023

134	Maluku Utara	354650,3	2023
135	Papua	181390,92	2023
136	Papua Barat	162795,25	2023

Pembersihan Data

Tahap selanjutnya adalah pembersihan data atau yang sering disebut sebagai *Data Cleaning*, suatu proses penting dalam *Data Mining* yang melibatkan identifikasi, koreksi, dan penghapusan anomali ataupun ketidaksesuaian dalam dataset (Putri, 2024). Melalui proses ini, data yang telah terkumpul akan melewati tahap pembersihan dengan menghapus variabel-variabel yang tidak relevan, sehingga data siap untuk dianalisis secara optimal. Proses ini bertujuan untuk meminimalisir potensi *error* yang dapat muncul akibat adanya nilai kosong atau data yang tidak lengkap saat tahap pengolahan. Dengan kualitas data yang akan mendukung pengambilan keputusan yang tepat, sementara *data warehouse* memerlukan integrasi data berkualitas yang konsisten (Meilina, 2015).

Pengolahan Data Menggunakan Metode *Machine Learning*

Dalam analisis untuk mengidentifikasi volume produksi perikanan tangkap laut di Indonesia, digunakan tiga metode regresi, yaitu *Decision Tree*, *Random Forest*, dan *Linear Regression*. Hasil dari ketiga algoritma tersebut kemudian akan dibandingkan untuk menentukan metode yang memiliki tingkat akurasi tertinggi.

Visualisasi Data

Visualisasi data adalah proses merepresentasikan informasi atau data menggunakan elemen grafis seperti grafik, diagram, dan tabel untuk membantu memahami pola, tren, dan hubungan dalam data (Putri et al., 2023). Data yang telah diolah disajikan dalam bentuk visual yang mudah dipahami, seperti grafik, tabel, dan diagram, untuk memudahkan pemahaman dan penyajian hasil penelitian (Devanda et al., 2024). Tujuan utama visualisasi data adalah menyederhanakan data agar lebih mudah dipahami, memudahkan penemuan wawasan, serta mendukung pengambilan keputusan yang lebih baik (Guntara et al., 2023).

Validasi Data

Validasi data digunakan untuk memastikan bahwa hasilnya konsisten ketika dilakukan secara terpisah. Proses ini melibatkan perbandingan nilai akurasi prediksi dari setiap algoritma untuk mengidentifikasi algoritma yang paling akurat (Saifudin, A. 2018).

Analisis Hasil

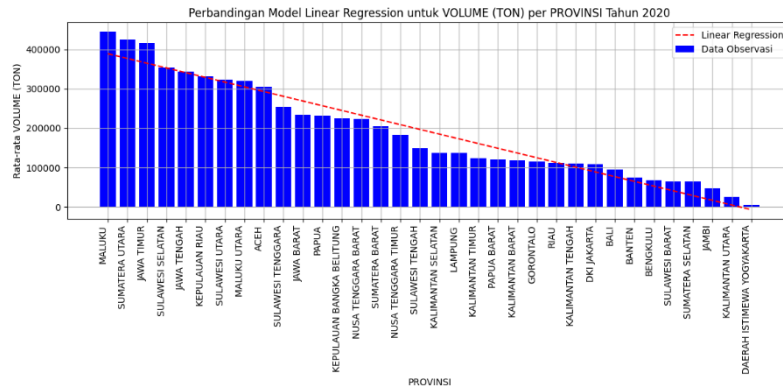
Setelah melalui tahap evaluasi, data hasil evaluasi akan dibandingkan performanya masing-masing, yang diukur melalui nilai R^2 . Hasil ini akan menunjukkan tingkat akurasi dan kemampuan masing-masing model dalam menangani variasi pada data produksi perikanan. Algoritma yang memiliki nilai R^2 tertinggi dianggap paling akurat dan andal dalam memprediksi volume produksi. Untuk memperjelas perbedaan performa di antara algoritma, hasil prediksi divisualisasikan dalam bentuk grafik, diagram, atau tabel. Visualisasi ini memudahkan untuk memahami pola prediksi tiap model dari tahun ke tahun, serta memperlihatkan bagaimana suatu algoritma dikatakan lebih unggul dibandingkan dengan algoritma lain.

HASIL DAN PEMBAHASAN

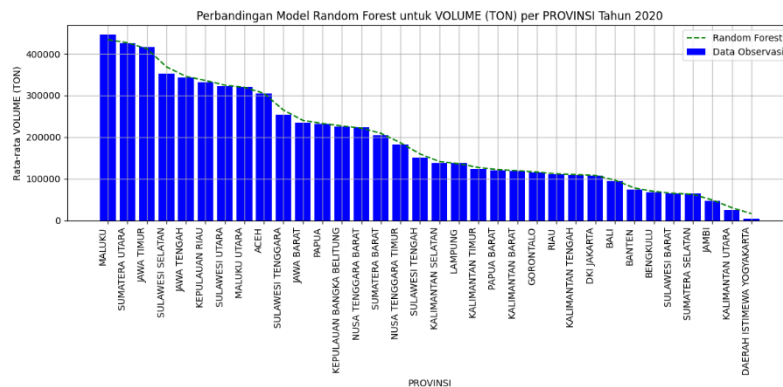
Penelitian ini menggunakan tiga algoritma *Machine Learning* yaitu *Decision Tree*, *Random Forest*, dan *Linear Regression* untuk mengidentifikasi volume produksi perikanan tangkap laut di Indonesia selama periode 2020 hingga 2023. Evaluasi kinerja model dilakukan menggunakan koefisien determinasi (R^2), yang mengindikasikan sejauh mana model mampu memprediksi data aktual.

Algoritma *Decision Tree* secara konsisten mencapai nilai R^2 sebesar 1.0 setiap tahun, dari 2020 hingga 2023. Nilai ini menunjukkan tingkat akurasi yang sangat tinggi serta kemampuan model untuk beradaptasi dengan variasi dalam data. Struktur model *Decision Tree* yang mampu menangani variasi kompleks dalam dataset menjadikannya unggul dalam analisis ini. Di sisi lain, *Random Forest* juga menunjukkan performa yang sangat baik, dengan nilai R^2 mendekati 1.0 di setiap tahun: 0.998 pada 2020, 0.995 pada 2021, 0.995 pada 2022, dan 0.994 pada 2023. Meskipun sedikit lebih rendah dibandingkan *Decision Tree*, algoritma ini tetap memberikan akurasi yang tinggi. Keunggulan *Random Forest* berasal dari kemampuannya menggabungkan beberapa *tree* keputusan, yang mengurangi risiko *overfitting* serta meningkatkan konsistensi prediksi (Putri et al., 2023).

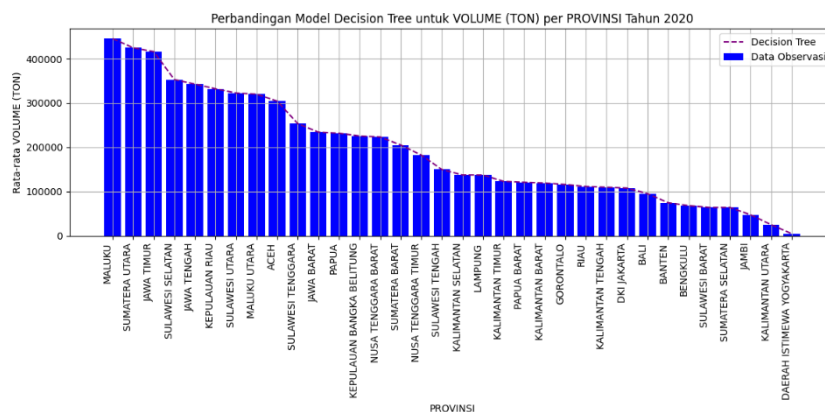
Sebaliknya, *Linear Regression* memiliki nilai R^2 yang lebih rendah dibandingkan kedua algoritma lainnya, yaitu 0.949 pada 2020, 0.903 pada 2021, 0.907 pada 2022, dan 0.871 pada 2023. Meskipun menunjukkan akurasi yang cukup baik, tren penurunan nilai R^2 seiring waktu mengindikasikan bahwa *Linear Regression* kurang optimal untuk dataset dengan pola yang kompleks atau non-linear. Visualisasi per tahun menunjukkan pola prediksi untuk setiap model dan menguatkan bahwa *Decision Tree* dan *Random Forest* lebih cocok untuk data aktual daripada *Linear Regression*.



Gambar 2. Grafik Visualisasi Model *Linear Regression* Tahun 2020



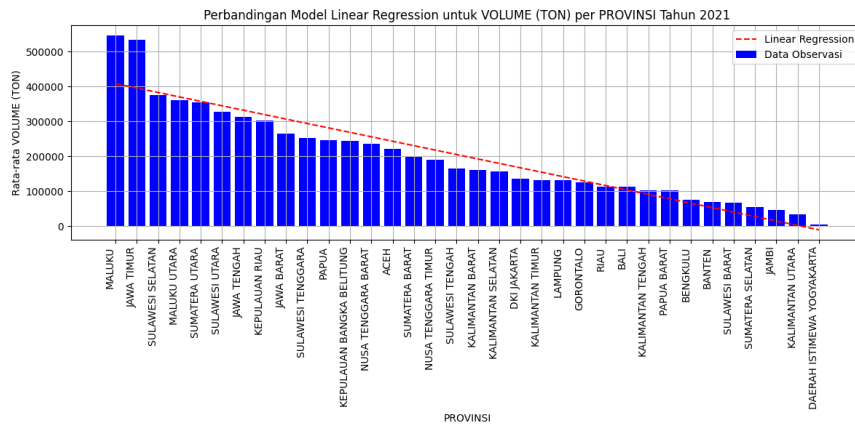
Gambar 3. Grafik Visualisasi Model *Random Forest* Tahun 2020



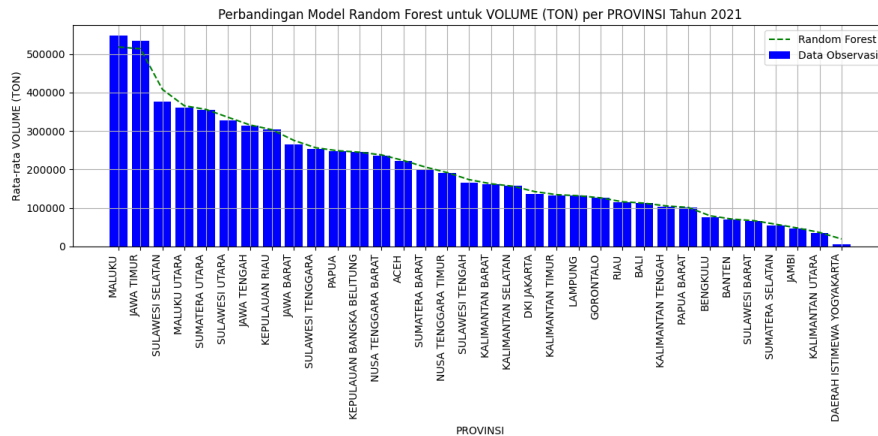
Gambar 4. Grafik Visualisasi Model *Decision Tree* Tahun 2020

Hasil grafik pada tahun 2020 menunjukkan perbedaan yang signifikan dalam kemampuan prediksi di antara ketiga algoritma. Model *Linear Regression* terlihat kurang mampu menangkap pola data yang kompleks, menghasilkan garis prediksi yang linier dan menyimpang pada beberapa titik, terutama di provinsi dengan volume produksi sangat tinggi atau rendah. Sebaliknya, *Random Forest* menunjukkan performa yang jauh lebih baik dengan prediksi yang presisi dan konsisten. *Decision Tree* memberikan hasil terbaik dengan prediksi

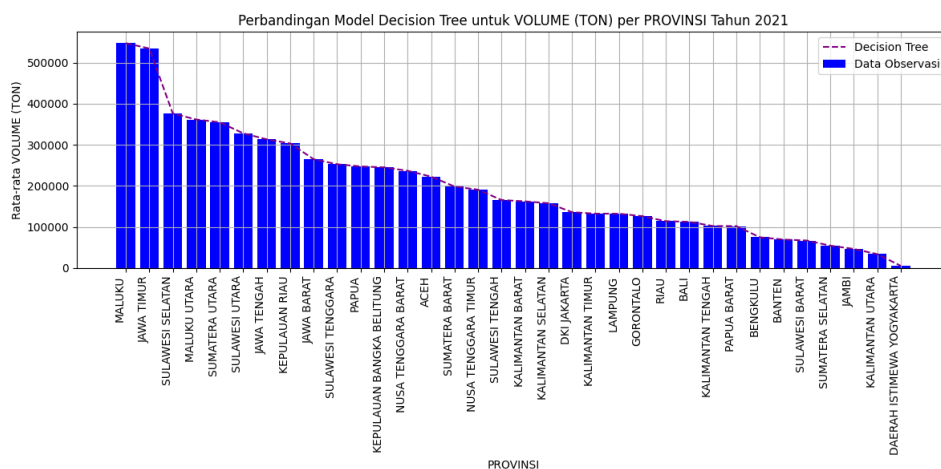
yang sangat akurat dan nilai R^2 sempurna (1.0), mencerminkan kemampuannya untuk merepresentasikan pola data secara tepat.



Gambar 5. Grafik Visualisasi Model *Linear Regression* Tahun 2021

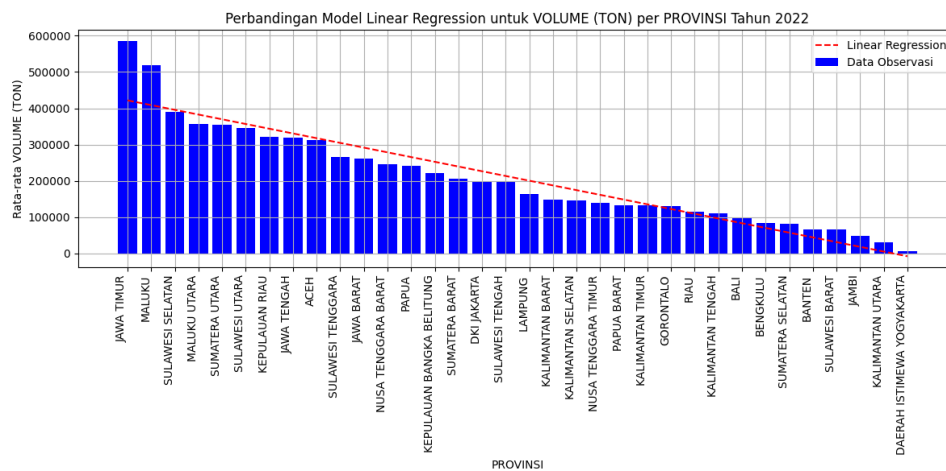


Gambar 6. Grafik Visualisasi Model *Random Forest* Tahun 2021

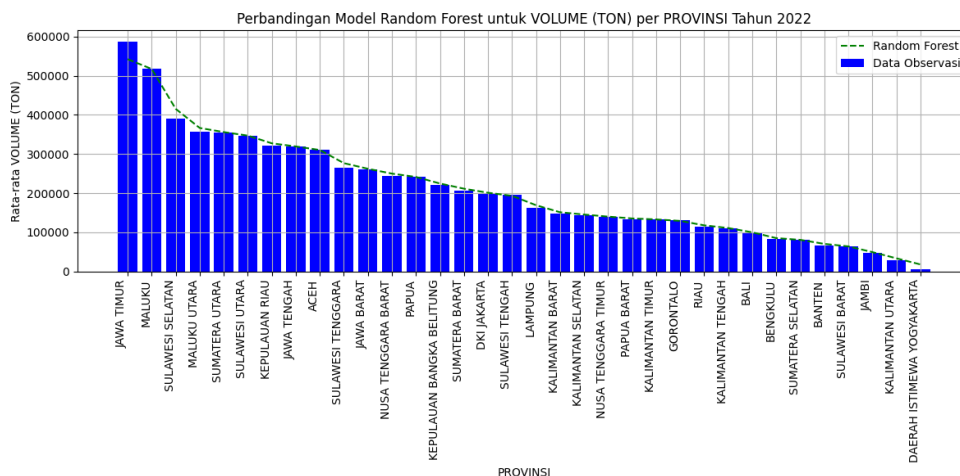


Gambar 7. Grafik Visualisasi Model *Decision Tree* Tahun 2021

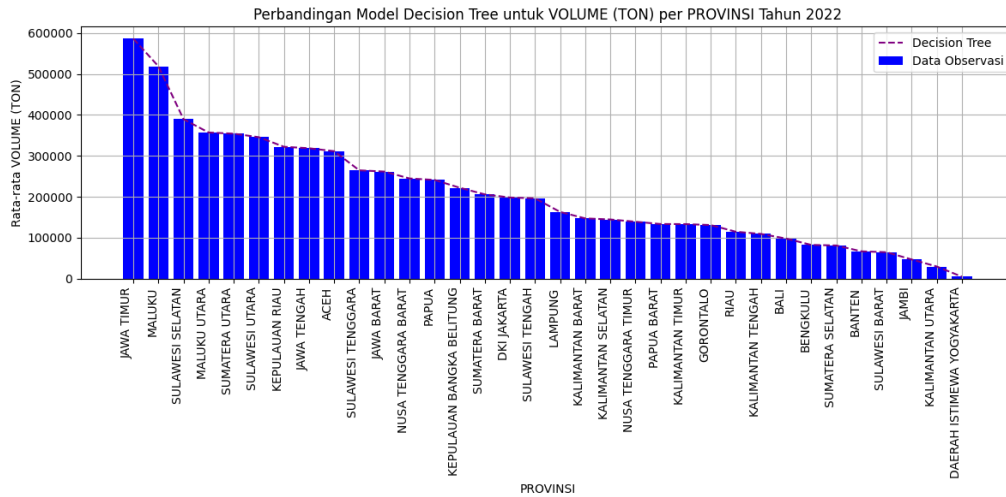
Linear Regression tidak sebaik *Random Forest* atau *Decision Tree*, terutama di daerah yang produksinya banyak seperti Jawa Timur dan daerah yang produksinya lebih sedikit seperti Maluku Utara atau Gorontalo. Garis perkiraan dalam *Linear Regression* mungkin menyimpang, terutama pada kumpulan data yang besar, karena garis perkiraan tersebut linier dan kurang mampu menangkap hubungan non-linier dalam informasi. Dibandingkan dengan model Regresi, model *Random Forest* menampilkan hasil prediksi yang lebih tepat, yang berhubungan dengan pemeriksaan data dengan akurasi dan ketahanan yang tinggi serta pola non-linier. *Decision Tree* juga cukup handal dalam prediksinya dan hampir sama efektifnya dengan *Random Forest*, meskipun lebih sederhana dan kurang toleran terhadap perubahan data. Model *Random Forest* sangat berhasil dalam memprediksi berapa banyak ikan yang ditangkap Indonesia pada tahun 2021.



Gambar 8. Grafik Visualisasi Model *Linear Regression* Tahun 2022

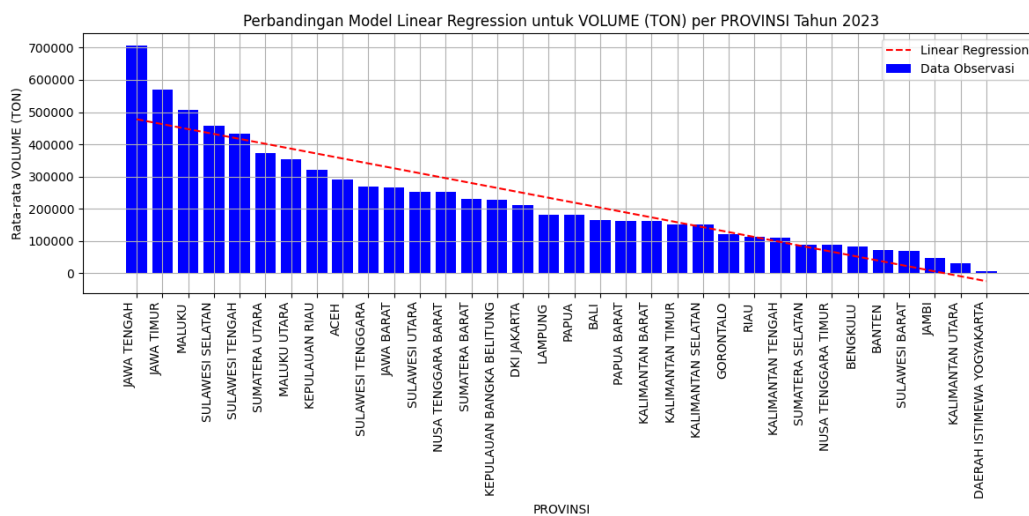


Gambar 9. Grafik Visualisasi Model *Random Forest* Tahun 2022

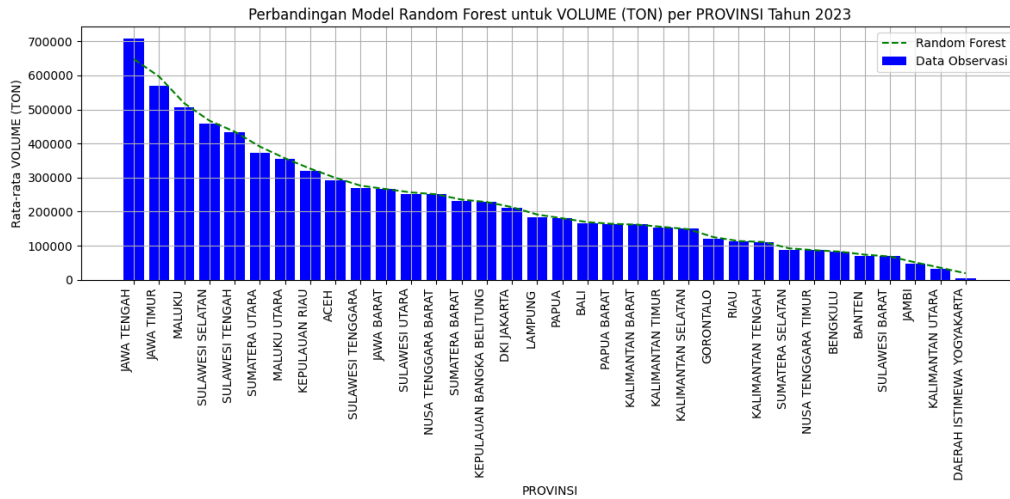


Gambar 10. Grafik Visualisasi Model *Decision Tree* Tahun 2022

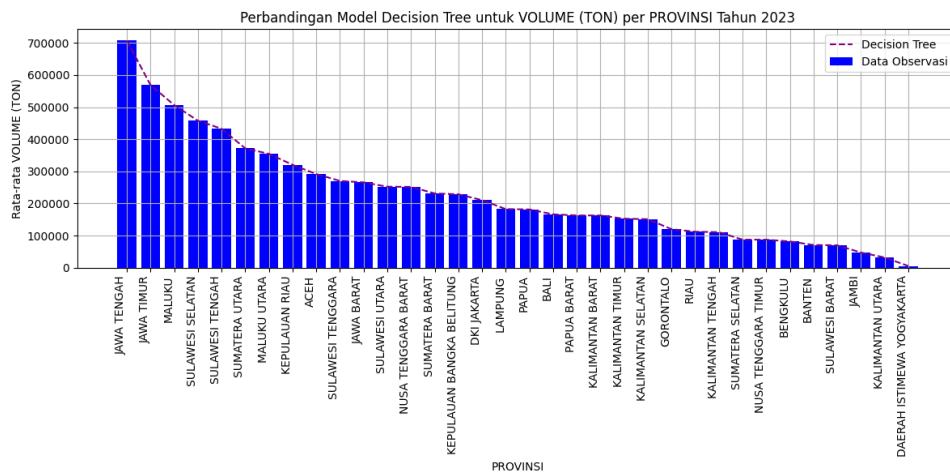
Pada tahun 2022, *Linear Regression* kembali menunjukkan keterbatasannya dalam menangkap pola data yang kompleks, menghasilkan garis prediksi linier yang menyimpang dari data aktual, terutama pada provinsi dengan volume produksi yang sangat tinggi atau rendah. *Random Forest*, di sisi lain, memberikan prediksi yang lebih akurat dengan nilai R^2 sebesar 0.995. *Decision Tree* kembali menunjukkan performa terbaik dengan nilai R^2 sempurna (1.0), yang mengindikasikan akurasi maksimal dalam merepresentasikan pola data. Algoritma ini berhasil menangkap detail variasi data dengan baik, bahkan pada provinsi dengan perbedaan produksi yang signifikan. *Random Forest* dan *Decision Tree* terus menunjukkan keunggulan mereka dalam menangani pola data produksi perikanan yang kompleks pada tahun 2022.



Gambar 11. Grafik Visualisasi Model *Linear Regression* Tahun 2023



Gambar 12. Grafik Visualisasi Model *Random Forest* Tahun 2023



Gambar 13. Grafik Visualisasi Model *Decision Tree* Tahun 2023

Hasil visualisasi tahun 2023 mempertegas pola performa dari ketiga algoritma yang diuji. *Linear Regression* kembali menunjukkan keterbatasannya, dengan garis prediksi yang linier dan tidak mampu menangkap hubungan non-linear dalam data. *Random Forest* tetap konsisten memberikan prediksi yang akurat dengan nilai R^2 sebesar 0.994, mencerminkan kemampuannya dalam menangani kompleksitas data dan variasi volume produksi antarprovinsi. *Decision Tree* kembali mencapai nilai R^2 sempurna (1.0), menunjukkan akurasi maksimal dalam memprediksi volume produksi perikanan di seluruh provinsi. Dengan kemampuannya yang unggul, *Decision Tree* tetap menjadi algoritma terbaik, diikuti oleh *Random Forest* yang juga menunjukkan kinerja sangat baik, jauh melampaui *Linear Regression*. Hasil visualisasi tersebut menegaskan kembali keandalan *Decision Tree* dan *Random Forest* dalam analisis volume produksi perikanan tangkap laut.

KESIMPULAN

Penelitian ini menunjukkan bahwa algoritma *Decision Tree* dan *Random Forest* memiliki keunggulan signifikan dibandingkan *Linear Regression* dalam mengidentifikasi volume produksi perikanan tangkap laut di Indonesia. Kedua algoritma ini berhasil memberikan akurasi prediksi yang sangat tinggi, dengan *Decision Tree* mencapai nilai R^2 sempurna (100%) secara konsisten dari tahun 2020 hingga 2023, sementara *Random Forest* mendekati nilai R^2 100% di setiap tahun, yaitu 99.8% pada tahun 2020, 99.5% pada tahun 2021 dan 2022, serta 99.4% pada tahun 2023. Dengan tingkat akurasi mencapai 100% pada metode *Decision Tree* dan hampir 100% pada *Random Forest*, hasil penelitian ini memperkuat kemampuan kedua algoritma tersebut dalam mengidentifikasi dan mengelola data produksi perikanan tangkap laut di Indonesia. Hasil ini mendukung analisis yang lebih akurat dan dapat diandalkan untuk mendukung pengembangan strategi pengelolaan perikanan yang berkelanjutan.

UCAPAN TERIMA KASIH

Terima kasih yang sebesar-besarnya kepada seluruh pihak yang telah berkontribusi dalam penelitian ini. Penulis menghargai bantuan dari Kementerian Kelautan dan Perikanan yang telah menyediakan data terkait volume produksi perikanan tangkap laut di Indonesia, yang menjadi dasar utama dalam analisis penelitian ini. Ucapan terima kasih yang tulus juga kami sampaikan kepada dosen pembimbing atas bimbingan, saran, dan masukan berharga yang diberikan pada setiap tahap penelitian. Penulis sangat mengapresiasi kerjasama dan dedikasi yang ditunjukkan oleh rekan-rekan penelitian yang telah bekerja dengan penuh semangat demi keberhasilan penelitian ini.

DAFTAR PUSTAKA

- Ariawan, I., Rosalia, A. A., Anzani, L., Arifin, W. A., Minsaris, L. O. A., & Lukman. 2021. Identifikasi Spesies Mangrove Menggunakan Algoritma Random Forest. *Jurnal Kemaritiman: Indonesian Journal of Maritime* 2(2): 87-96. <https://doi.org/10.17509/ijom.v2i2.33920>
- Arifin, WA, Ariawan, I., Rosalia, A. A., Lukman, & Tufailah, N. 2022. Data Scaling Performance on Various Machine Learning Algorithms to Identifyabalone Sex. *Jurnal Teknologi dan Sistem Komputer*, 10(1): 26-31. <https://dx.doi.org/10.14710/jtsiskom.2022.14105>
- Devanda, M., Kusuma, H., & Hidayat, S. (2024). Penerapan Model Regresi Linier dalam Prediksi Harga Mobil Bekas di India dan Visualisasi dengan Menggunakan Power BI. In *Jurnal Indonesia: Manajemen Informatika dan Komunikasi (JIMIK)* 5(2). <https://journal.stmiki.ac.id>

- Dewan Pertimbangan Presiden Republik Indonesia. (2017). <https://wantimpres.go.id/id/2017/04/potensi-perikanan-indonesia/>
- Guntara, R. G., Nugraha, M. R., Aprilia, R., Dewinta, N. A. P., Dewanto, F. B., & Nurahma, S. (2023). Pemodelan Data Menggunakan Algoritma Decision Tree untuk Visualisasi Data Hubungan Nilai Siswa dengan Makan Siang, Partisipasi Kursus Persiapan Tes, dan Tingkat Pendidikan. *Jurnal Perjuangan* 5(2). <https://e-journal.unper.ac.id/index.php/informatics>
- Putri, A. (2024). Pentingnya Data Cleaning sebelum Visualisasi: Teknik Dan Tips. *Jurnal Teknologi Pintar* 4(5).
- Kementerian Kelautan dan Perikanan. 2022. <https://kkp.go.id/news/news-detail/kkp-perbarui-data-estimasi-potensi-ikan-totalnya-1201-juta-ton-per-tahun65c1b1dfcd9b1.html>
- Kencana Putri, A., & Ichsanuddin Nur, D. (2023). Penggunaan Bahasa Python untuk Analisis an Visualisasi Data Penduduk Di Desa Sumberjo, Nganjuk. *Jurnal Pengabdian Kepada Masyarakat* 3(3). https://jurnalfkip.samawa-university.ac.id/karya_jpm/index
- Mai, P., Tarigan, S., Tata Hardinata, J., Qurniawan, H., Safii, M., Winanjaya, R., Studi, P., Informasi, S., Tunas, S., & Pematangsiantar, B. (2022). Implementasi Data Mining Menggunakan Algoritma Apriori dalam Menentukan Persediaan Barang (Studi Kasus: Toko Sinar Harahap). *Jurnal Sistem Informasi, Teknologi Informasi dan Komputer* 12(2). <https://jurnal.umj.ac.id/index.php/just-it/index>
- Meilina, P. (2015). Penerapan Data Mining dengan Metode Klasifikasi Menggunakan Decision Tree dan Regresi. *Jurnal Teknologi* 7(1).
- Maryani, H., & Nasution, A. 2019. Rekonsepsi Model Pemberantasan Illegal Fishing di Perairan Indonesia (Analisis Perspektif Hukum Internasional). *Jurnal Legislasi Indonesia*, 16(3): 379-391.
- Pambudi, H. K., Giri, P., Kusuma, A., Yulianti, F., Ahessa, K., Jurnal, J., Teknologi, I., Terapan, I., Julian, K. A., Studi, P., Logistik, T., & Rekayasa Industri, F. (2020). Prediksi Status Pengiriman Barang Menggunakan Metode Machine Learning. *Jurnal Ilmiah Teknologi Terapan* 6(2).
- Patlisan, & Rusdah. 2023. Optimasi Akurasi Model Decision Tree Menggunakan Random Forest Regression untuk Prediksi Kuantitas Pembelian Barang pada Perusahaan Manufaktur. *Jurnal SIMETRIS*, 14(2).
- Rahma, C. A. 2022. *Analisis dan Penerapan Data Mining untuk Mengestimasi Laju Pertumbuhan Penduduk Kabupaten Musi Banyuasin Menggunakan Metode Regresi Linier Berganda di Badan Pusat Statistik Musi Banyuasin*. [SKRIPSI]. Jambi: Universitas Dinamika Bangsa.
- Rizky, A. C., Muzakki, A., Rangga, M., & Kusuma, P. (2024). Prediksi Jumlah Produksi Ikan Tahun 2018-2022 pada Teluk Banten Menggunakan Support Vector Regression. *Jurnal Responsif*, 6(2), 205–213. <https://ejurnal.ars.ac.id/index.php/jti>
- Rokach, L., & Maimon, O. 2007. *Data Mining with Decision Trees – Theory and Applications*. Singapore: World Scientific Publishing.

- Saifudin, A. 2018. Metode Data Mining untuk Seleksi Calon Mahasiswa pada Penerimaan Mahasiswa Baru di Universitas Pamulang. *Jurnal Teknologi*, 10(1).
- Sholeh, M., Nurnawati, E. K., & Lestari, U. 2023. Penerapan Data Mining dengan Metode Regresi Linear untuk Memprediksi Data Nilai Hasil Ujian Menggunakan RapidMiner. *JISKA (Jurnal Informatika Sunan Kalijaga)*, 8(1): 10-21.
- Taufik, F. R. 2019. Mekanisme Penyediaan Data Produksi Perikanan Tangkap di Pelabuhan Perikanan Samudera Nizam Zachman Jakarta. [SKRIPSI]. Bogor: Institut Pertanian Bogor. 71 halaman.