

PREDIKSI PRODUKTIVITAS KOMODITAS JAGUNG BERDASARKAN POLA CURAH HUJAN DI NUSA TENGGARA TIMUR MENGGUNAKAN ALGORITMA RANDOM FOREST

Inocentus Albertus Naimnanu¹, Tri Ana Setyarini², Skolastika S. Igon³

^{1,2,3}STIKOM Uyelindo Kupang

inonaimnanu@gmail.com¹, setyarinitriana@gmail.com²,

igon5kolastika@gmail.com³

ABSTRAK

Produktivitas jagung di Nusa Tenggara Timur (NTT) mengalami fluktuasi yang tidak stabil akibat ketergantungan pada curah hujan musiman dan dampak fenomena El Niño, sehingga sulit diprediksi secara akurat. Ketiadaan model prediksi yang andal menyebabkan penanganan produksi jagung cenderung reaktif dan berpotensi memperburuk ketahanan pangan lokal. Penelitian ini bertujuan membangun sistem prediksi produktivitas jagung di NTT berbasis pola curah hujan menggunakan algoritma Random Forest. Data yang digunakan bersumber dari BPS Provinsi NTT mencakup periode 2010–2023 dengan 308 data dari 22 kabupaten/kota. Model dibangun menggunakan algoritma Random Forest Regressor dengan 100 pohon keputusan, pembagian data kronologis 80/20 (*shuffle=False*) untuk mencegah kebocoran informasi dari data masa depan ke dalam proses pelatihan, dan divalidasi menggunakan Time Series Cross Validation (TSCV) dengan 5 lipatan. Antarmuka sistem dibangun menggunakan framework Streamlit dan diuji dengan metode *black box testing*. Hasil pengujian TSCV menunjukkan rata-rata R^2 sebesar 78,29%, MAE sebesar 9.472.27 Ton, MAPE sebesar 43,92%, dan RMSE sebesar 16.039,48 Ton. Seluruh komponen sistem dinyatakan valid. Sistem ini terbukti mampu memprediksi produktivitas jagung di NTT dengan tingkat akurasi yang memadai dan dapat digunakan sebagai alat bantu pengambilan keputusan pertanian berbasis data.

Kata Kunci: Curah hujan, Nusa Tenggara Timur, Produktivitas Jagung, *Random Forest*.

ABSTRACT

Corn productivity in East Nusa Tenggara (NTT) experiences unstable fluctuations due to dependence on seasonal rainfall and the impact of the El Niño phenomenon, making it difficult to predict accurately. The lack of a reliable prediction model means that corn production management tends to be reactive and has the potential to worsen local food security. This study aims to develop a corn productivity prediction system in NTT based on rainfall patterns using the Random Forest algorithm. The data used is sourced from the NTT Provincial Statistics Agency (BPS) covering the period 2010–2023 with 308 data points from 22 regencies/cities. The model was built using the Random Forest Regressor algorithm with 100 decision trees, an 80/20 chronological data split (*shuffle=False*) to prevent information leakage from future data into the training process, and validated using Time Series Cross Validation (TSCV) with 5 folds. The system interface was built

using the Streamlit framework and tested using black-box testing methods. TSCV test results showed an average R^2 of 78.29%, MAE of 9,472.27 tons, MAPE of 43.92%, and RMSE of 16,039.48 tons. All system components were deemed valid. This system has proven capable of predicting corn productivity in NTT with a sufficient level of accuracy and can be used as a data-driven agricultural decision-making tool.

Keywords: *Corn productivity, East Nusa Tenggara, Rainfall, Random Forest.*

A. PENDAHULUAN

Jagung (*Zea mays L.*) merupakan salah satu komoditas pangan strategis di Indonesia setelah padi, karena memiliki peran ganda baik sebagai sumber pangan langsung maupun bahan baku industri, khususnya pakan ternak. Di Provinsi Nusa Tenggara Timur (NTT), jagung bahkan menempati posisi utama sebagai makanan pokok masyarakat pedesaan, menjadikannya komoditas vital bagi ketahanan pangan Hudoyo dan Nurmayasari (2019). Namun, produktivitas di wilayah ini menunjukkan fluktuasi yang signifikan dari tahun ke tahun. Berdasarkan data resmi dari Badan Pusat Statistik, tren produktivitas jagung di Nusa Tenggara Timur (NTT) selama beberapa tahun terakhir menunjukkan volatilitas yang nyata: tercatat sebesar 26,33 Ku/Ha pada tahun 2019, menurun pada tahun 2020 menjadi 24,07 Ku/Ha, kemudian naik lagi menjadi 25,81 Ku/Ha pada tahun 2021, menurun kembali menjadi 24,88 Ku/Ha pada tahun 2022, sedikit meningkat menjadi 24,94 Ku/Ha pada tahun 2023, kemudian kembali meningkat menjadi 26,93 Ku/Ha pada tahun 2024.

Data ini memperlihatkan bahwa produktivitas jagung di NTT tidak stabil dan sulit diprediksi, yang berdampak langsung pada ketersediaan pangan lokal. Ketidakstabilan ini berakar pada ketergantungan yang tinggi pada sistem pertanian tadah hujan yang sangat rentan terhadap pola curah hujan yang tidak menentu, sebagai faktor penentu utama hasil panen, berdasarkan data rilis dari BMKG Stasiun Klimatologi Nusa Tenggara Timur (2024), akibat fenomena El Niño pada tahun 2023 telah menyebabkan awal musim hujan 2023/2024 di sebagian besar wilayah NTT mengalami kemunduran (terlambat) dan sifat hujan diperkirakan berada pada kategori Bawah Normal hingga Normal. Kondisi ini menyebabkan ketidakpastian hasil panen, karena pergeseran awal musim hujan dan lamanya periode kering berdampak langsung terhadap produktivitas jagung, mengakibatkan produktivitas jagung menjadi sangat sulit diprediksi secara akurat.

Ketiadaan model prediksi yang *reliabel* menjadikan langkah-langkah yang diambil cenderung reaktif, sehingga rawan memperburuk kerentanan ekonomi dan krisis pangan di tingkat lokal (Herlina dan Prasetyorini, 2020). Masalah mendasar yang kemudian muncul adalah bagaimana cara mengetahui hasil prediksi produktivitas secara akurat berdasarkan pola curah hujan yang tidak menentu.

Pada penelitian terdahulu menunjukkan bahwa algoritma *machine learning*, khususnya Random Forest, mampu memberikan solusi efektif pada berbagai permasalahan prediksi berbasis data. Saputra et al. (2024) berhasil menerapkan Random Forest untuk memprediksi risiko kredit macet dengan akurasi mencapai 94,8%, membuktikan kemampuannya dalam menangkap pola kompleks dari data historis. Relevansi lain ditunjukkan oleh Fadellia Azzahra et al. (2024) yang menggunakan Random Forest untuk memprediksi kasus stunting, di mana validasi silang memastikan model tetap stabil dan konsisten dalam berbagai skenario data. Dalam konteks agroklimatologi, Random Forest mampu membangun model prediktif yang kuat dengan menggabungkan banyak pohon keputusan (*decision tree*), sehingga menghasilkan prediksi yang lebih stabil dan akurat. Bahkan, penelitian oleh Nursalman dan Mustikasari (2020) menegaskan bahwa melalui optimasi *hyperparameter*, performa Random Forest dapat meningkat hingga akurasi 95,4%. Berbagai bukti ini mengindikasikan bahwa Random Forest menjadi solusi potensial untuk membangun model produktivitas jagung berdasarkan data curah hujan.

Untuk mengatasi permasalahan kompleks ini, diperlukan sebuah pendekatan yang mampu memodelkan hubungan non-linear antara variabel pola curah hujan dengan hasil panen. *Machine learning*, khususnya algoritma Random Forest, menawarkan solusi yang menjanjikan. Random Forest merupakan metode *ensemble learning* yang bekerja dengan membangun sejumlah besar pohon keputusan (*decision trees*) secara acak dan menggabungkan prediksi dari setiap pohon untuk menghasilkan prediksi akhir yang lebih akurat dan stabil (Wihardjaka et.al.,2020). Keunggulan utamanya terletak pada kemampuannya menangani data berdimensi tinggi, mengukur tingkat kepentingan setiap variabel, serta memiliki ketahanan yang baik terhadap *overfitting*, sehingga sangat sesuai untuk data iklim yang kompleks. Penelitian sebelumnya telah membuktikan keandalan Random Forest dalam berbagai kasus prediksi, seperti prediksi risiko kredit, prediksi kasus *stunting*, serta optimasi performa model hingga mencapai akurasi tinggi.

Berdasarkan uraian diatas, maka penelitian ini bertujuan untuk merancang dan membangun sebuah model prediktif menggunakan algoritma Random Forest yang mampu memperkirakan produktivitas jagung secara akurat berdasarkan variasi pola curah hujan di Provinsi Nusa Tenggara Timur. Diharapkan model yang dihasilkan tidak hanya memiliki tingkat akurasi yang tinggi, tetapi juga dapat berfungsi sebagai alat bantu pengambilan keputusan (*decision support tool*) yang praktis bagi para pemangku kepentingan. Dengan demikian, penelitian ini berupaya menjembatani antara analisis data iklim yang rumit dengan kebutuhan praktis di lapangan, guna mendukung terwujudnya pertanian yang lebih tangguh dan ketahanan pangan yang berkelanjutan di Nusa Tenggara Timur.

Rumusan Masalah

Berdasarkan uraian latar belakang diatas, rumusan masalah dalam penelitian ini adalah bagaimana memprediksi produktivitas komoditas jagung berdasarkan pola curah hujan di Nusa Tenggara Timur menggunakan algoritma Random Forest?

Tujuan Penelitian

Tujuan dari penelitian ini adalah Membangun sebuah model prediktif fungsional untuk memprediksi produktivitas komoditas jagung di Nusa Tenggara Timur berdasarkan pola curah hujan menggunakan algoritma Random Forest.

Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat baik dalam kontribusi pada pengembangan ilmu di bidang infotmatika, khususnya dalam penerapan *machcine learning* dan menjadi acuan bagi peneliti sejenis dimasa mendatang, serta model yang dihasilkan dapat berfungsi sebagai alat bantu pengambilan keputusan berbasis data bagi pemerintah daerah dalam menyusun strategi adaptasi pertanian terhadap variabilitas iklim.

Ruang Lingkup Penelitian

Ruang lingkup penelitian ini dibatasi pada aspek-aspek untuk memastikan fokus penelitian dan pencapaian tujuan yang optimal:

1. Variabel yang dianalisis mencakup data curah hujan bulanan, serta data produktivitas jagung tahunan.

2. Metode analisis menggunakan algoritma Random Forest untuk membangun model prediksi produktivitas jagung berdasarkan pola curah hujan.
3. *Tools* yang digunakan mencakup *Microsoft Office* untuk proses pengolahan data, serta *Visual Studio Code* sebagai lingkungan pengembangan teks dan skrip.

TINJAUAN PUSTAKA

1. Penelitian Terdahulu

Penyusunan penelitian ini perlu mengkaji penelitian-penelitian terdahulu yang relevan dengan topik. Kajian ini menjadi fondasi untuk mengidentifikasi perkembangan metode, hasil penelitian, serta peluang pengembangan penelitian yang masih tersedia. Beberapa penelitian yang terkait adalah:

Ferdiansyah et al., (2024) Penelitian dengan judul Memprediksi Pilihan Sarapan Pagi menggunakan Random Forest berdasarkan Pola Cuaca. Mengangkat mengenai variasi pilihan sarapan pagi yang dipengaruhi oleh kondisi cuaca sehingga tujuannya adalah untuk mengembangkan model prediktif yang dapat meramalkan pilihan sarapan seseorang secara akurat berdasarkan parameter cuaca seperti suhu dan kelembapan. Metode yang digunakan melibatkan pengumpulan data cuaca harian dari *OpenWeatherMap API* dan data pilihan sarapan melalui survei selama tujuh bulan, yang kemudian dianalisis menggunakan algoritma Random Forest di aplikasi RapidMiner. Hasil akhirnya menunjukkan bahwa model Random Forest yang dibangun berhasil mencapai tingkat akurasi yang sangat tinggi, yaitu 94,84%, dan mengidentifikasi bahwa suhu serta kelembapan adalah faktor paling penting dalam memprediksi menu sarapan.

Andrianof et al., (2023) menerapkan algoritma Random Forest dalam memprediksi kelulusan mahasiswa berdasarkan data akademik di perguruan tinggi Indonesia, masalah utama yang diangkat adalah rendahnya tingkat kelulusan tepat waktu mahasiswa di Indonesia, yang rata-rata hanya mencapai 45% secara nasional. Oleh karena itu, tujuan penelitian ini adalah menerapkan algoritma Random Forest untuk membangun model yang mampu memprediksi status kelulusan mahasiswa berdasarkan data akademik mereka, sehingga dapat mengidentifikasi mahasiswa yang berisiko terlambat lulus. Metode yang dijalankan adalah dengan mengimplementasikan model Random Forest pada dataset yang berisi seribu catatan mahasiswa, dengan menganalisis faktor-faktor seperti Indeks Prestasi Kumulatif (IPK), tingkat kehadiran, dan nilai mata kuliah inti.

Hasil akhir dari penelitian ini adalah sebuah model prediktif dengan tingkat akurasi 87,5%, presisi 86,3%, dan recall 85,9%, di mana ditemukan bahwa IPK dan tingkat kehadiran merupakan faktor paling penting yang memengaruhi kelulusan.

Romahdoni et al., (2025) menerapkan Algoritma Random Forest untuk memprediksi keberhasilan Usaha mikro, kecil dan menengah (UMKM). membahas masalah banyaknya Usaha Mikro, Kecil, dan Menengah (UMKM) yang tidak mampu bertahan di tengah persaingan bisnis yang ketat, sering kali disebabkan oleh faktor internal seperti manajemen keuangan yang lemah dan strategi pemasaran yang tidak optimal. Tujuan dari penelitian ini adalah menerapkan algoritma Random Forest untuk membangun sebuah sistem yang dapat memprediksi peluang keberhasilan sebuah UMKM berdasarkan berbagai variabel bisnis. Untuk itu, metode yang digunakan adalah menganalisis dataset dari Kaggle yang berisi 250 sampel dengan 13 variabel menggunakan software Orange Data Mining, di mana model Random Forest diimplementasikan dengan parameter 10 pohon keputusan dan dievaluasi menggunakan Cross Validation. Hasil akhirnya adalah model prediksi yang menunjukkan kinerja sangat baik dengan tingkat akurasi 86,4%, F1-score 0.862, dan recall 0.864, membuktikan bahwa Random Forest adalah metode yang andal untuk memprediksi keberhasilan UMKM.

Penelitian -penelitian terdahulu menunjukkan bahwa algoritma Random Forest memiliki kemampuan yang andal dan konsisten dalam membangun model prediktif di berbagai bidang, mulai dari perilaku konsumsi, performa akademik, hingga keberhasilan usaha.

2. Prediksi

Prediksi adalah suatu proses penalaran yang bertujuan untuk memperkirakan hal yang kemungkinan besar akan terjadi di masa depan dengan memanfaatkan data dan informasi yang telah dikumpulkan dari masa lalu dan masa kini. Proses ini tidak selalu menghasilkan jawaban secara pasti, melainkan memberikan gambaran yang paling mendekati berdasarkan data yang tersedia. Dalam Kamus Besar Bahasa Indonesia, prediksi diartikan sebagai hasil ramalan atau perkiraan mengenai kejadian di waktu mendatang berdasarkan pemanfaatan data historis dan informasi aktual, sehingga prediksi sering juga disebut sebagai prakiraan atau peramalan. (Hengky dan Rusman, 2022)

3. Komoditas jagung

Nilai strategis jagung terlihat dari luasan panen yang mendominasi beberapa kabupaten sentra, seperti Timor Tengah Selatan, Belu, serta Kupang, dimana total luasan panen setingkat provinsi mencapai ratusan ribu hektar menurut data BPS. Produktivitas jagung di NTT dipengaruhi oleh faktor lingkungan fisik khususnya curah hujan, teknik budidaya, penggunaan varietas lokal, pemupukan, dan penyuluhan pertanian yang berkesinambungan. Lahan kering yang mendominasi bentang alam NTT menuntut adaptasi teknologi dan manajemen tanam spesifik, termasuk pengembangan varietas toleran kekeringan dan penggunaan pupuk organik (Pallo, 2020). Kendala lain seperti ketersediaan benih unggul, keterbatasan modal, serta akses pasar juga mempengaruhi stabilitas dan perkembangan agribisnis jagung daerah ini.

4. Pola curah hujan

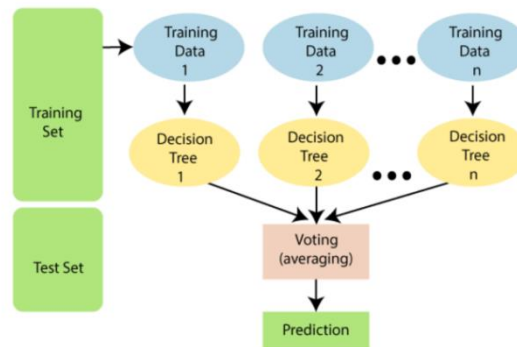
Pola curah hujan merupakan karakteristik distribusi presipitasi di suatu kawasan dalam periode tertentu yang dipengaruhi oleh faktor klimatologis dan geografis, terutama dinamika suhu permukaan air laut. Ketika suhu air laut di Pasifik tengah menurun, pembentukan awan konvektif terbatas sehingga intensitas presipitasi di Indonesia meningkat, fenomena ini disebut La Nina. Sebaliknya, peningkatan suhu air laut mengurangi aktivitas konveksi atmosfer yang membentuk awan hujan sehingga volume presipitasi di Indonesia menurun, kondisi ini disebut El Nino. (Pradana et al., 2023).

5. Algoritma Random Forest

Algoritma Random Forest termasuk dalam jenis algoritma ensemble yang bekerja dengan menggabungkan banyak pohon keputusan (*Decision Tree*) untuk menghasilkan prediksi. Dalam kasus regresi, algoritma ini membangun sejumlah *Decision Tree* berdasarkan subset data yang diambil secara acak melalui teknik bootstrap sampling. Setiap pohon memberikan nilai prediksi tersendiri, kemudian seluruh hasil tersebut dirata-ratakan untuk menghasilkan *output* akhir. (Jollyta et al., 2023)

Random Forest Regression merupakan salah satu metode dalam machine learning yang berfungsi untuk memprediksi variabel bernilai kontinu. Teknik ini termasuk ke dalam kelompok algoritma ensemble, yaitu algoritma yang menggabungkan sejumlah model *Decision Tree* untuk meningkatkan ketepatan hasil prediksi serta meminimalkan

risiko overfitting. Jika dibandingkan dengan metode regresi klasik seperti Regresi Linear, Random Forest Regression memiliki fleksibilitas lebih tinggi dan mampu menangani data dengan pola non-linear.(Widarto Rachbini, 2025)



Gambar 1. Diagram Struktur Random Forest

Random Forest Regression menghasilkan prediksi dengan membangun sejumlah pohon keputusan secara independen, lalu menggabungkan keluarannya. Inti mekanismenya mencakup beberapa konsep berikut:

1. *Bootstrap sampling*

Bootstrap sampling adalah prosedur pemilihan subset data dari himpunan asli secara acak dengan pengembalian, sehingga setiap sampel dapat terpilih lebih dari sekali.

2. Pembentukan *decision tree*

Setiap pohon dibangun dengan memilih subset fitur secara acak, kemudian menentukan titik belah terbaik memakai kriteria tertentu. Menggunakan MSE (*Mean Squared Error*) atau varians. Pohon mencari belahan di mana error-nya paling kecil.

$$MSE_n = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{Y}_i)^2 \dots\dots\dots(1)$$

Dimana:

MSE_n = menyatakan galat kuadrat rata-rata pada pohon ke-n

N = Jumlah sampel pada pohon ke-n

X_i = adalah nilai sampel ke-i pada pohon ke-n

\bar{Y}_i = merupakan nilai rata-rata sampel pada pohon ke-n.

3. Prediksi akhir

Dalam Random Forest Regression, nilai prediksi final diperoleh dengan merata-ratakan output dari seluruh pohon keputusan dalam ensemble:

$$\hat{y} = \frac{1}{T} \sum_{t=1}^T \hat{y}_t \dots\dots\dots(2)$$

Dimana:

\hat{y} = adalah prediksi akhir model.

T = menyatakan jumlah pohon dalam hutan.

\hat{y}_t = adalah prediksi yang dihasilkan oleh pohon ke-t

Beberapa keunggulan utama dari Random Forest Regression antara:

1. Mengurangi *Overfitting*

Dengan menggabungkan hasil rata-rata dari banyak pohon keputusan, model menjadi lebih stabil serta tidak mudah terpengaruh oleh data ekstrem atau outlier.

2. Meningkatkan Akurasi

Kombinasi banyak *Decision Tree* memungkinkan model mengenali pola data yang lebih rumit, sehingga ketepatan prediksi meningkat.

3. Menangani hubungan non-linear

Berbeda dari regresi linear yang mensyaratkan asumsi keterkaitan linier antar variabel, Random Forest mampu berkinerja baik pada dataset yang lebih kompleks dengan pola non-linear.

4. Mendukung data kategorikal dan numerik

Algoritma ini dapat mengolah beragam tipe fitur, baik kategorikal maupun numerik, tanpa tuntutan pra-pemrosesan yang berat.

Evaluasi Model

Evaluasi model regresi bertujuan menilai sejauh mana model mampu memprediksi variabel dependen secara akurat berdasarkan data yang tersedia. Beragam metrik digunakan untuk mengukur kinerja, masing-masing dengan keunggulan dan keterbatasan yang bergantung pada tujuan analisis serta karakteristik data. (Kusrini, et al., 2025).

1. *Mean Absolute Error (MAE)*

Mean Absolute Error (MAE) mengukur rata-rata besar kesalahan prediksi dengan menghitung nilai rata-rata dari selisih absolut antara prediksi model dan nilai aktual.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \dots\dots\dots(3)$$

Dimana:

n = jumlah data

y_i = nilai aktual pengamatan ke- i

\hat{y}_i = nilai prediksi model pengamatan ke- i

2. *Root Mean Squared Error (RMSE)*

Root Mean Squared Error (RMSE) merupakan akar kuadrat dari MSE, sehingga mengembalikan satuan kesalahan ke skala yang sama dengan variabel aslinya dan memudahkan interpretasi. Secara umum, RMSE dihitung dengan mengambil akar dari rata-rata kuadrat selisih antara nilai aktual dan prediksi.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|^2} \dots\dots\dots (4)$$

Dimana:

n = jumlah data

y_i = nilai aktual pengamatan ke- i

\hat{y}_i = nilai prediksi model pengamatan ke- i

3. R^2 (*koefisien determinasi*)

Koefisien determinasi (R^2) mengukur proporsi variasi pada variabel dependen yang dapat dijelaskan oleh variabel independen dalam model. Nilai R^2 berada pada rentang 0 hingga 1; semakin mendekati 1 berarti model semakin baik dalam menjelaskan keragaman data.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \dots\dots\dots(5)$$

Dimana:

n = jumlah data

y_i = nilai aktual pengamatan ke- i

\hat{y}_i = nilai prediksi model pengamatan ke- i

\bar{y} = nilai rata-rata dari semua y_i

4. *Mean Absolute Percentage Error (Mape)*

Rata-rata persentase kesalahan prediksi secara keseluruhan dihitung menggunakan metrik MAPE (*Mean Absolute Percentage Error*), yang mengukur seberapa besar selisih antara nilai prediksi dan nilai aktual dalam bentuk persentase.

$$MAPE = \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \times 100 \dots \dots \dots (6)$$

- n = jumlah data
- y = nilai hasil aktual
- \hat{y} = nilai hasil pendugaan

Visual Studio Code

Visual Studio Code (VS Code) adalah aplikasi editor teks yang ringan namun sangat andal, dikembangkan oleh *Microsoft* untuk mendukung sistem operasi multiplatform, termasuk *Linux*, *Mac*, dan *Windows*. Editor ini secara langsung dapat menjalankan berbagai bahasa pemrograman seperti *JavaScript*, *Typescript*, dan *Node.js*. Selain itu, bahasa pemrograman lain seperti *C++*, *C#*, *Python*, *Go*, dan *Java* juga dapat digunakan dengan bantuan *plugin* yang tersedia di *marketplace Visual Studio Code*.(Ramadhan dan Saputra, 2020).

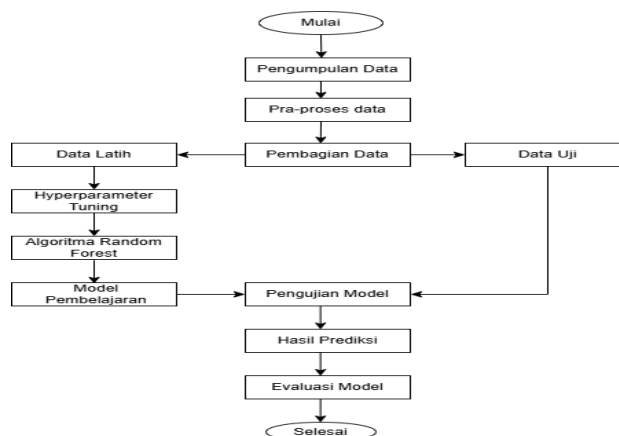
B. METODE PENELITIAN

Data Penelitian

Bahan penelitian yang digunakan dalam penelitian ini berupa data produktivitas jagung, luas tanam, dan data pola curah hujan di Provinsi Nusa Tenggara Timur (NTT) selama periode 2010 – 2023. Data produktivitas jagung mencakup hasil produksi dalam satuan ton per hektar (ton/ha) serta data curah hujan yang diperoleh dari Badan Pusat Statistik (BPS).

Prosedur Pelitian

Berikut ini adalah prosedur dalam penelitian ini yakni sebagai berikut:



Gambar 2. Flowchart Penelitian

1. Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari instansi resmi pemerintah.

a. Data Produktivitas Jagung

Data produktivitas jagung dan Data luas lahan diperoleh dari Badan Pusat Statistik (BPS) Provinsi Nusa Tenggara Timur. Data ini mencakup informasi hasil panen jagung dalam satuan ton per hektar (ton/ha) pada tingkat kabupaten/kota selama periode 2010–2023.

Data ini digunakan untuk menggambarkan tingkat hasil produksi jagung di setiap wilayah dan tahun tertentu.

b. Data Pola Curah Hujan.

Data curah hujan diperoleh dari Badan Badan Pusat Statistik (BPS) Provinsi Nusa Tenggara Timur, yang menyediakan data historis curah hujan bulanan (mm) di wilayah NTT. Data curah hujan digunakan untuk membentuk pola curah hujan berdasarkan nilai rata-rata, maksimum, dan minimum per bulan maupun per musim tanam.

2. Pra-proses data

Data mentah yang telah dikumpulkan kemudian melalui tahap praproses agar siap digunakan dalam analisis. Proses ini mencakup pembersihan data dengan menghapus data duplikat dan menangani nilai yang hilang (missing value), normalisasi skala data agar

seragam, serta pembuatan variabel baru seperti rata-rata curah hujan per musim tanam. Tahapan ini bertujuan untuk memastikan bahwa data yang digunakan memiliki kualitas yang baik dan relevan terhadap tujuan penelitian.

3. Pembagian Data

Dataset dibagi menjadi dua bagian yaitu data latih (training data) dan data uji (testing data). Data latih digunakan untuk melatih model, sedangkan data uji digunakan untuk mengukur kemampuan model dalam melakukan prediksi terhadap data baru.

4. *Hyperparameter Tuning*

Sebelum melatih model utama, dilakukan proses pencarian parameter terbaik untuk mengoptimalkan kinerja algoritma. Menggunakan metode seperti *Grid Search*, penelitian ini akan menguji berbagai kombinasi parameter Random Forest (seperti jumlah pohon atau *n_estimators* dan kedalaman pohon atau *max_depth*) untuk menemukan kombinasi mana yang menghasilkan *error* paling kecil.

5. Penerapan Algoritma Random Forest

penerapan algoritma Random Forest untuk membangun model prediksi. Algoritma ini bekerja dengan membuat sejumlah pohon keputusan (*decision trees*) berdasarkan data latih yang diambil secara acak. Setiap pohon akan menghasilkan prediksi tersendiri dan hasil akhir diperoleh melalui rata-rata dari seluruh pohon tersebut.

6. Model pembelajaran

Model yang dihasilkan dari algoritma Random Forest merupakan hasil proses pembelajaran dari data latih. Model ini telah mempelajari pola hubungan antara curah hujan, terhadap hasil produksi jagung. Model pembelajaran ini kemudian digunakan untuk melakukan prediksi pada data uji.

7. Pengujian Model

Model yang telah dilatih kemudian diuji menggunakan data uji untuk mengevaluasi tingkat akurasi dan keandalannya.

8. Hasil Prediksi

Hasil yang akan dicapai dari penelitian ini adalah terbentuknya model prediksi produksi jagung berbasis pola curah hujan. Dari hasil pengujian, diperoleh prediksi produksi jagung di masa mendatang berdasarkan kondisi curah hujan yang menjadi inputannya.

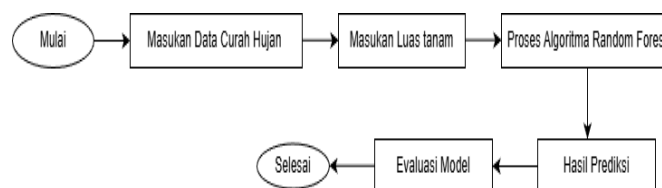
9. Evaluasi Model

Evaluasi dilakukan dengan menghitung metrik seperti *Mean Absolute Error (MAE)*, *Root Mean Square Error (RMSE)*, dan R^2 (*koefisien determinasi*). Nilai-nilai ini digunakan untuk mengukur sejauh mana model mampu memberikan hasil prediksi yang mendekati nilai aktual produksi jagung.

Perancangan Sistem

1. Alur Proses Sistem

Proses kerja sistem digambarkan sebagai berikut:



Gambar 3. Alur kerja Sistem.

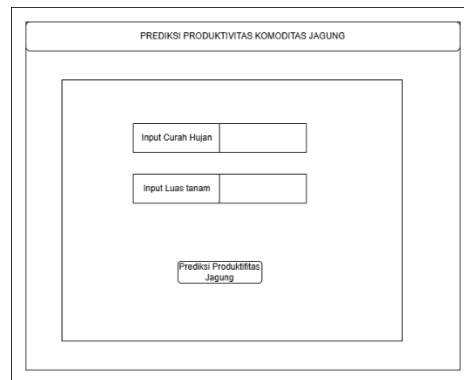
Perancangan Sistem

Antarmuka sistem dikembangkan menggunakan bahasa pemrograman Python dengan pustaka *Streamlit* untuk menampilkan tampilan grafis berbasis web. Model prediksi Random Forest yang telah dilatih diintegrasikan ke dalam sistem sehingga pengguna dapat memasukkan data curah hujan dan luas tanam untuk memperoleh hasil prediksi produktivitas jagung secara interaktif.

1. Halaman Inputan

Komponen utama antarmuka terdiri atas dua kolom input, yaitu curah hujan (mm/tahun) dan luas tanam jagung (ha), serta satu kolom output berupa hasil prediksi produktivitas jagung (ton/ha). Hasil perhitungan model Random Forest ditampilkan secara langsung kepada pengguna setelah tombol “Prediksi” ditekan. Desain antarmuka

ini bertujuan untuk mempermudah pengguna non-teknis dalam memanfaatkan sistem prediksi berbasis data iklim dan agronomis.

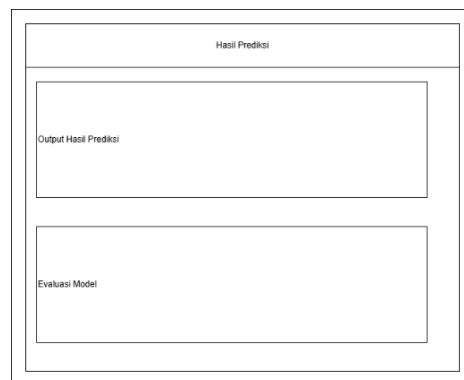


The image shows a web interface for predicting corn productivity. At the top, it says "PREDIKSI PRODUKTIVITAS KOMODITAS JAGUNG". Below this, there are two input fields: "Input Curah Hujan" and "Input Luas tanam". At the bottom, there is a button labeled "Prediksi Produktifitas Jagung".

Gambar 4. Halaman inputan data

2. Halaman Output

Pada halaman output yang ditampilkan sistem adalah hasil prediksi (ton/ha) dan evaluasi model.



The image shows a web interface for displaying prediction results. At the top, it says "Hasil Prediksi". Below this, there are two main sections: "Output Hasil Prediksi" and "Evaluasi Model".

Gambar 5. Halaman hasil

C. HASIL DAN PEMBAHASAN

Impelentasi Sistem

Implementasi sistem ini menggunakan bahasa pemrograman Python dengan *framework Streamlit* sebagai antarmuka berbasis web. Sistem ini mengintegrasikan model *machine learning* Random Forest Regressi untuk memprediksi produktivitas jagung di Provinsi NTT.

Tahap Pre-processing Data

Tahap pra-proses bertujuan untuk memastikan data mentah yang diperoleh dari Badan Pusat Statistik (BPS) memiliki kualitas yang optimal sebelum digunakan dalam proses pelatihan model.

Tabel 1. Dataset Produktivitas jagung dan rata-rata curah hujan

Tahun	Kode_kabupaten	Kabupaten	Curah_Hujan	Luas_Tanam	Produktivitas
2010	1	Sumba Barat	156	4.838	9.987
2010	2	Sumba Timur	91	19.084	24.589
2010	3	Kupang	322	19.995	70.034
2010	4	Timor Tengah Selatan	230	64.080	151.879
2010	5	Timor Tengah Utara	79	12.605	55.285
2010	6	Belu	228	23.589	77.124
2010	7	Alor	136	4.385	14.242
2010	8	Lembata	0	9.548	16.712
2010	9	Flores Timur	100	15 672	41.602
2010	10	Sikka	91	11.820	41.268
2010	11	Ende	130	4.563	10.150
2010	12	Ngada	295	9.555	14.956
2010	13	Manggarai	382	5.109	11.989
2010	14	Rote Ndao	152	4.270	10.365
2010	15	Manggarai Barat	131	6.407	14.808
2010	16	Sumba Tengah	0	4.353	11.079
...
...

...
2023	20	Sabu Raijua	163	2.819	11.613
2023	21	Malaka	98	620	6.214
2023	22	Kupang	113	11.647	49.377

Langkah-langkah teknis yang dilakukan adalah sebagai berikut:

1. Pembersihan Nama Kolom

Sistem melakukan normalisasi pada judul kolom Excel dengan menghapus spasi di awal dan akhir (*trailing spaces*) menggunakan fungsi *.strip()*. Langkah ini mencegah kesalahan pemanggilan variabel akibat perbedaan karakter tersembunyi dalam nama kolom.

2. Penangan *Missing Values*

Untuk variabel Tahun, dilakukan pengisian otomatis menggunakan metode *Forward Fill (ffill)* guna memastikan kontinuitas data kronologis. Untuk variabel Curah Hujan, nilai kosong atau bernilai nol diimputasi menggunakan nilai rata-rata curah hujan spesifik per kabupaten melalui fungsi *groupby*. Pendekatan imputasi berbasis kelompok wilayah ini dipilih karena setiap kabupaten memiliki karakteristik curah hujan yang berbeda.

3. Normalisasi Fitur

Normalisasi diterapkan menggunakan *MinMaxScaler* pada variabel masukan, yaitu Kode Kabupaten, Curah Hujan, dan Luas Tanam, untuk menyamakan skala data ke dalam rentang 0 hingga 1. Penyeragaman skala ini penting agar tidak ada satu variabel pun yang mendominasi proses pembelajaran model hanya karena perbedaan satuan pengukurannya. Meskipun algoritma Random Forest secara teknis tidak memerlukan normalisasi karena berbasis pohon keputusan, tahap ini ditetapkan sebagai bentuk standarisasi pre-processing data agar dataset berada dalam kondisi siap pakai.

Implementasi Model *Machine Learning*

Implementasi inti dari kecerdasan buatan dalam penelitian ini menggunakan algoritma Random Forest Regresi, dengan rincian sebagai berikut:

1. Pembagian Dataset

Model menggunakan fungsi (*train_test_split*) dengan parameter *shuffle=False* untuk pembagian data dilakukan secara kronologis (tanpa pengacakan) untuk menjaga urutan waktu dan mencegah kebocoran data (data leakage) sebelum membaginya. 80% dari total data dialokasikan sebagai data latih yang berfungsi sebagai basis pengetahuan bagi algoritma untuk mengenali serta mempelajari pola hubungan antara variabel curah hujan, luas lahan, dan produktivitas jagung secara mendalam. Sementara itu, 20% sisanya difungsikan sebagai data uji yang berperan sebagai parameter evaluasi independen untuk mengukur sejauh mana model mampu memberikan estimasi yang akurat pada informasi baru yang belum pernah dipelajari sebelumnya, sehingga mencegah terjadinya fenomena *overfitting* atau kondisi di mana model hanya mampu menghafal data tanpa bisa melakukan generalisasi.

2. Konfigurasi Parameter Model

Model Random Forest dikonfigurasi dengan parameter *n_estimators* sebesar 100 pohon keputusan dan *random_state* bernilai 42 untuk memastikan reproduibilitas hasil. Penggunaan 100 pohon bertujuan membangun seratus jalur logika independen yang masing-masing memproses sampel data secara acak. Nilai prediksi akhir diperoleh melalui rata-rata seluruh luaran pohon guna meminimalkan variansi dan menstabilkan akurasi model secara kolektif.

3. Persitensi Model

Setelah proses pelatihan selesai, model disimpan dalam format .pkl ke dalam direktori models/. Penyimpanan ini memungkinkan aplikasi antarmuka memanggil fungsi prediksi secara instan tanpa perlu mengulangi proses pelatihan setiap kali sistem dijalankan, sehingga efisiensi operasional sistem terjaga.

Hasil Uji Performa Model

Pengujian model dilakukan terhadap 61 sampel data uji (*y_true*) produktivitas jagung di NTT dan estimasi yang dihasilkan (*y_pred*) oleh algoritma Random Forest mencerminkan kemampuan model dalam melakukan generalisasi terhadap pola historis. Evaluasi menunjukkan bahwa algoritma Random Forest mampu memetakan pola historis

produktivitas jagung di NTT secara akurat. Kedekatan antara nilai riil (y_{true}) dan hasil estimasi (y_{pred}) membuktikan keberhasilan model dalam mengekstraksi fitur dominan dari variabel curah hujan dan luas lahan. Deviasi minimal pada berbagai sampel uji tersebut menegaskan adanya presisi tinggi dalam menangkap karakteristik data untuk memprediksi hasil panen di masa mendatang.

Tabel 2. Perbandingan data aktual dan data prediksi

No	Data Aktual (Ton)	Hasil Prediksi (Ton)	Selisih (Ton)
1	31,975.00	50,576.63	18,601.63
2	23,995.00	21,448.24	2,546.76
3	30,191.00	26,811.61	3,379.39
4	20,118.00	20,688.12	570.12
5	30,344.00	25,969.71	4,374.29
6	14,938.00	11,890.66	3,047.34
7	16,938.00	21,272.89	4,334.89
8	10,316.00	7,648.72	2,667.28
9	6,895.00	16,049.32	9,154.32
10	10,078.00	8,391.51	1,686.49
...
61	1,247.00	949.,42	279.58

Tabel 3. Evaluasi Model

R ²	MAE	MAPE	RMSE
0,9172	6.459,92	36,33%	9.965,99

Evaluasi terhadap 61 sampel data uji menunjukkan bahwa model Random Forest memperoleh nilai koefisien determinasi (R^2) sebesar 0,9172, yang mengindikasikan bahwa model mampu menjelaskan hampir 92% variansi produktivitas jagung pada periode data uji. Nilai MAE sebesar 6.459,92 Ton dan RMSE sebesar 9.965,99 Ton menunjukkan adanya deviasi yang cukup besar pada beberapa titik data. Adapun nilai MAPE sebesar 36,33% menunjukkan bahwa secara rata-rata terdapat kesalahan relatif sebesar 36,33% antara nilai prediksi dan nilai aktual, mengindikasikan masih adanya ruang optimasi yang perlu ditangani melalui penambahan variabel prediktor dan penyesuaian *hyperparameter* model.

Validasi Stabilitas Model (*Time series cross validation*)

Setelah melakukan pengujian pada model utama, langkah selanjutnya adalah menguji stabilitas dan reliabilitas algoritma melalui metode *time series cross validation* (TSCV) dengan parameter $k = 5$. Penggunaan validasi silang berbasis deret waktu ini sangat krusial dalam penelitian agroklimatologi karena data produktivitas jagung memiliki ketergantungan temporal yang kuat. Berbeda dengan K-Fold *cross validation* standar yang mengacak data secara acak, TSCV menjaga urutan kronologis data untuk mencegah fenomena data *leakage* atau kebocoran informasi masa depan ke dalam proses pelatihan masa lalu. Tujuan utama *time series cross validation* adalah membuktikan bahwa model konsisten di berbagai periode waktu, bukan hanya bagus pada satu pengujian saja. Jika hanya menampilkan 1 fold, tidak ada cara untuk mengetahui apakah hasil tersebut kebetulan baik atau memang mencerminkan kemampuan model yang sesungguhnya. Performa model pada tiap lipatan data (*fold*) secara berturut-turut adalah sebagai berikut:

Tabel 4. Rata-rata performa *Time series cross validation*

Fold	Data Latih	Data Uji	R^2	MAE (Ton)	MAPE	RMSE (Ton)
1	51 baris	50 baris	75.29%	10,863.74	66.76%	18,424.74

2	101 baris	50 baris	78.52%	8,363.74	36.23%	15,649.03
3	151 baris	50 baris	68.20%	10,602.62	26.05%	19,693.10
4	201 baris	50 baris	76.78%	11,050.54	53.47%	16,207.51
5	251 baris	50 baris	92.67%	6,480.72	37.10%	10,223.03
Rata-rata	-	-	78.29%	9,472.27	43.92%	16,039.48

Kelima fold ditampilkan karena masing-masing merepresentasikan kemampuan model pada periode waktu yang berbeda, di mana setiap fold melatih model secara independen dengan data latih yang terus bertambah secara progresif. Rata-rata dari kelima fold itulah yang digunakan sebagai acuan akurasi utama, karena nilai tunggal dari satu fold saja tidak cukup untuk membuktikan bahwa model mampu melakukan generalisasi secara konsisten terhadap data yang belum pernah dipelajari sebelumnya.

Nilai rata-rata R^2 dari TSCV sebesar 78,29% dengan variasi antar lipatan berkisar antara 75,29% hingga 92,67% mencerminkan keberagaman pola data produksi jagung antar periode akibat variabilitas iklim di NTT. Sebagai acuan pelaporan, nilai rata-rata TSCV digunakan sebagai hasil akurasi utama karena lebih merepresentasikan kemampuan generalisasi model secara lintas-temporal, sedangkan hasil evaluasi 80/20 berfungsi sebagai referensi performa pada periode data terkini.

Implementasi Antarmuka Sistem

Halaman Input Data merupakan halaman utama sistem Prediksi Produktivitas Jagung NTT. Pada halaman ini pengguna dapat memasukkan nilai Rata-rata Curah Hujan (mm) dan Luas Tanam NTT (Ha) sesuai data yang tersedia, kemudian menekan tombol Prediksi produktivitas jagung untuk memproses dan menampilkan hasil prediksi produksi jagung secara otomatis.



Gambar 6. Halaman Inputan



Gambar 7. Halaman Output

Pengujian Sistem

Metode yang diterapkan dalam pengujian ini adalah *black box testing*. Pemilihan metode ini didasarkan pada kebutuhan untuk mengamati fungsionalitas sistem dari sudut pandang pengguna akhir (*end-user*). Pengujian ini berfokus sepenuhnya pada input dan output sistem tanpa melibatkan pemeriksaan struktur kode *internal*.

Tabel 5. Pengujian Sistem

Fitur / Komponen	Skenario Pengujian	Hasil yang Diharapkan	Status
Input Curah Hujan	Memasukkan angka desimal (titik) pada kolom curah hujan.	Sistem menerima angka desimal sebagai parameter input.	Valid
Input Luas Tanam	Memasukkan angka ribuan (tanpa	Sistem menerima angka tersebut sebagai nilai numerik yang utuh.	Valid

Fitur / Komponen	Skenario Pengujian	Hasil yang Diharapkan	Status
	pemisah) pada kolom luas tanam.		
Input Tidak Valid	Membiarkan kolom input kosong lalu menekan tombol prediksi.	Sistem menampilkan peringatan agar pengguna mengisi data terlebih dahulu.	Valid
Input Karakter Teks	Mencoba memasukan karakter teks pada kolom angka.	Sistem menolak input dan hanya menerima nilai numerik.	Valid
Tombol Prediksi	Menekan tombol “Prediksi Produktivitas Jagung” setelah data terisi.	Sistem memproses data melalui model dan mengalihkan ke tampilan hasil.	Valid
Matrik Akurasi (R2)	Menampilkan nilai koefisien determinasi hasil pengujian model.	Muncul angka akurasi dalam format desimal dan persentase.	Valid
Metrik MAE, MAPE dan RMSE	Menampilkan rata-rata kesalahan prediksi (Error) pada layar.	Muncul angka MAE dalam satuan Ton dan MAPE dalam satuan persen (%).	Valid
Visualisasi Grafik	Menampilkan grafik perbandingan data asli dan prediksi.	Muncul grafik garis yang memetakan fluktuasi data secara akurat.	Valid

Fitur / Komponen	Skenario Pengujian	Hasil yang Diharapkan	Status
Responsivitas UI	Mengubah ukuran jendela browser atau perangkat.	Tampilan tabel dan grafik menyesuaikan dengan ukuran layar <i>(Responsive)</i> .	Valid

Analisis Kelebihan dan Kekurangan Sistem

Setelah dilakukan implementasi dan pengujian terhadap sistem prediksi produktivitas jagung menggunakan algoritma Random Forest, ditemukan beberapa kelebihan dan kekurangan sebagai berikut:

1. Kelebihan Sistem

Keunggulan utama sistem ini terletak pada kemampuan algoritma Random Forest dalam menghasilkan estimasi produktivitas yang presisi serta didukung oleh kecepatan pemrosesan data secara real-time tepat setelah masukan diberikan. Selain itu terdapat fitur visualisasi interaktif yang memudahkan pengguna dalam memahami fluktuasi panen serta didukung mekanisme penyaringan data otomatis yang secara mandiri mampu menjaga validitas luaran sistem dari gangguan data yang tidak valid.

2. Kekurangan Sistem

Sistem ini masih memiliki keterbatasan fungsional karena hanya bergantung pada parameter cuaca dan luas lahan tanpa melibatkan variabel agronomis penting lainnya seperti serangan hama atau ketersediaan pupuk. Terdapat pula rigiditas pada format pengisian angka yang mewajibkan standar desimal internasional serta adanya ketergantungan yang sangat tinggi terhadap integritas data masa lalu dan cakupan wilayah operasional yang saat ini masih terbatas pada karakteristik geografis Nusa Tenggara Timur saja.

D. KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, dapat disimpulkan bahwa sistem prediksi produktivitas jagung di Provinsi Nusa Tenggara Timur

berhasil dibangun menggunakan algoritma Random Forest dengan antarmuka berbasis web Streamlit, memanfaatkan dua variabel input yaitu curah hujan dan luas tanam. Pengujian model menggunakan metode Time Series Cross Validation dengan 5 lipatan menghasilkan rata-rata akurasi (R^2) sebesar 78,29%, rata-rata MAE sebesar 9,472,27 Ton, rata-rata MAPE sebesar 43,92%, dan rata-rata RMSE sebesar 16.039,48 Ton, yang menunjukkan bahwa model mampu memprediksi produktivitas jagung dengan tingkat akurasi yang cukup memadai meskipun masih terdapat ruang untuk peningkatan. Seluruh komponen sistem telah diuji menggunakan metode black box testing dan dinyatakan berfungsi sesuai spesifikasi yang dirancang.

Saran

Berdasarkan keterbatasan yang ditemukan dalam penelitian ini, saran untuk pengembangan selanjutnya memperluas cakupan variabel luas panen sebagai pelengkap atau pengganti luas tanam untuk memberikan gambaran hasil riil yang lebih akurat, serta penambahan data mengenai jenis varietas benih yang digunakan oleh petani karena setiap varietas memiliki potensi hasil yang berbeda, selain itu, diperlukan peningkatan kuantitas data historis untuk memperkuat stabilitas prediksi model terhadap anomali cuaca di masa mendatang.

DAFTAR PUSTAKA

- Diyah Ruswanti. (2020). Pengukuran Performa Support Vector Machine dan Neutral Netwok dalam Meramalkan Tingkat Curah Hujan. *Gaung Informatika*, 13(1), 66–75. <https://www.academia.edu/108183614/>
- Dwi Ferdiansyah, Ervan Satrya Perdana, Gugun Guntara, Satria Winata, S. P. (2024). Memprediksi Pilihan Sarapan Pagi Menggunakan Random Forest Berdasarkan Pola Cuaca. *Seminar Teknologi Majalengka (Stima)*, 8, 77–83. <https://doi.org/10.31949/stima.v8i0.1212>
- Fadellia Azzahra, Suarna, N., & Arie Wijaya, Y. (2024). Penerapan Algoritma Random Forest Dan Cross Validation Untuk Prediksi Data Stunting. *Kopertip : Jurnal Ilmiah Manajemen Informatika Dan Komputer*, 8(1), 1–6. https://www.researchgate.net/publication/379297824_

- Harkamsyah Andrianof, Aggy Pramana Gusman, O. A. P. (2023). Jurnal Sains Informatika Terapan (JSIT). Implementasi Algoritma Random Forest Untuk Prediksi Kelulusan Mahasiswa Berdasarkan Data Akademik: Studi Kasus Di Perguruan Tinggi Indonesia, 2(1), 16–20. :
<https://rcfindonesia.org/home/index.php/jsit/article/view/464>
- Henni Kumaladewi Hengky, A. D. P. R. (2022). Model Prediksi Stunting. Penerbit NEM.
<https://books.google.co.id/books?id=IGimEAAAQBAJ>
- Herlina, N., & Prasetyorini, A. (2020). Effect of Climate Change on Planting Season and Productivity of Maize (*Zea mays* L.) in Malang Regency. Jurnal Ilmu Pertanian Indonesia, 25(1), 118–128. <https://doi.org/10.18343/jipi.25.1.118>
- Hudoyo, A., & Nurmayasari, I. (2019). Peningkatan Produktivitas Jagung di. In Indonesia Indonesian Journal of Socio Economics (Vol. 1, Issue 2), 102–108.
<https://jurnal.fp.unila.ac.id/index.php/IJSE/article/view/3627>
- Jollyta, D., Hajjah, A., Haerani, E., & Siddik, M. (2023). Algoritma Klasifikasi untuk Pemula Solusi Python dan RapidMiner. Deepublish
<https://books.google.co.id/books?id=y84TEQAAQBAJ>
- Kusrini, K., Agastya, I. M. A., Shiddiq, M. F. A., Setyanto, A., Kartikakirana, R. A., Adninda, G. B., Yuana, K. A., Nucifera, F., Johari, M. Z. F., Sanjaya, H., Rohman, A. N., Samponu, Y. B., Wahyuni, D., Rizaldi, Z., & Hartato, B. P. (2025). Data science geospasial. Andi – Computers, 978-623-001-472-6.
https://www.google.co.id/books/edition/DATA_SCIENCE_GEOSPASIAL/LNBBEQAAQBAJ?hl=id&gbpv=0
- Pallo, M. (2020). Analisis Deskriptif Produksi Jagung di Nusa Tenggara Timur. Partner, 25(1), 1311. <https://doi.org/10.35726/jp.v25i1.452>
- Pallo, M., Djunina, H., Se'u, V. E., & Mulik, Y. (2023). Eksplorasi Trend Produksi Jagung di Nusa Tenggara Timur dan Prediksinya. Produksi Tanaman, 011(10), 757–762
<https://doi.org/10.21776/ub.protan.2023.011.10.04>
- Ramadhan, N., & Saputra, M. H. K. (2020). Buku teknis pengantar aplikasi helpdesk internal berbasis website dan mobile. Kreatif.
<https://books.google.co.id/books?id=GZz6DwAAQBAJ>

- Romahdoni, M. R., Sormin, R. D., & Budiono, D. (2025). Penerapan Machine Learning Algoritma Random Forest untuk Prediksi Keberhasilan UMKM. *Journal of Software Engineering and Technology.*, 5(2), 28.
<https://doi.org/10.69769/seat.v5i2.245>
- Saputra, D. B., Atina, V., & Nastiti, F. E. (2024). Penerapan Model Crisp-Dm Pada Prediksi Nasabah Kredit Menggunakan Algoritma Random Forest. *IDEALIS : InDonEsiA Journal Information System*, 7(2), 240–247.
<https://doi.org/10.36080/idealis.v7i2.3244>
- Vichario Indra Pradana , Yuana Sukmawaty, N. S. (2023). RAGAM : PENERAPAN MODEL REGRESI PANEL KOMPONEN DUA ARAH PADA POLA CURAH HUJAN PROVINSI KALIMANTAN TENGAH. 02, 1–10.
<https://doi.org/10.20527/ragam.v2i1.8303>
- Widarto Rachbini, Z. (2025). DATA SCIENCE untuk Pemula dengan RapidMiner. CV. AA. Rizky, 9786234054330.
<https://books.google.co.id/books?id=95lMEQAAQBAJ>
- Wihardjaka, A., Pramono, A., & Sutriadi, M. T. (2020). Peningkatan Produktivitas Padi Sawah Tadah Hujan Melalui Penerapan Teknologi Adaptif Dampak Perubahan Iklim. *Jurnal Sumberdaya Lahan*, 14(1), 25.
<https://doi.org/10.21082/jsdl.v14n1.2020.25-36>
- SALMAN, N., & SARI, M. (2020). PENGARUH PENYETELAN HYPERPARAMETER TERHADAP KINERJA PREDIKSI RANDOM FOREST PADA PENDETEKSIAN SPAM. *Jurnal INSTEK (Informatika Sains Dan Teknologi)*, 5(2), 149–158. <https://doi.org/10.24252/instek.v5i2.15617>.