

OPTIMASI PERTANIAN DI BEKASI UTARA: PREDIKSI CURAH HUJAN DAN REKOMENDASI TANAMAN DENGAN MENGGUNAKAN MODEL REGRESI LINIER

OPTIMIZING AGRICULTURE IN NORTH BEKASI: RAINFALL PREDICTION AND CROP RECOMMENDATIONS USING LINEAR REGRESSION MODELS

Matias Indra Pangestu, s32200095@student.ubm.ac.id¹⁾, Jusia Amanda Ginting, gintingjusia2@gmail.com²⁾, I Gusti Ngurah Suryantara, gusti@bundamulia.ac.id³⁾
Matthew Marvelino, matthew.212523@gmail.com⁴⁾

¹⁾²⁾³⁾ Program Studi Informatika, Universitas Bunda Mulia

Diterima 2 September 2024 / Disetujui 16 September 2024

ABSTRACT

Uncertainty and variability in rainfall pose significant challenges for farmers in determining the optimal planting time and the most suitable crop types for fluctuating weather conditions. The system developed in this study utilizes historical rainfall data, climate parameters, and crop variety information collected from the North Bekasi region. Data collection was conducted using Landsat 8 satellite imagery and CHIRPS rainfall datasets, which were then processed using the Python programming language. Machine learning methods were applied to process and analyze the data to build an accurate predictive model. The results of the study indicate that the developed rainfall prediction system is capable of providing reasonably accurate rainfall forecasts. Additionally, the system can also offer recommendations for optimal crops based on climate conditions and rainfall predictions, which is expected to assist farmers in improving crop yields and food security in the region. The study concludes that the integration of satellite data, climate analysis, and machine learning methods can be an effective tool in supporting better agricultural decision-making. The developed system also has the potential to be implemented in other regions with similar conditions, after adjustments for local data are made.

Keywords: CHIRPS, Regresi Linear, Machine Learning, Data Satelit, Lansat 8

ABSTRAK

Ketidakpastian dan variabilitas curah hujan merupakan tantangan utama bagi petani dalam menentukan waktu tanam dan jenis tanaman yang paling sesuai untuk kondisi cuaca yang berubah-ubah. Sistem yang dibangun dalam penelitian ini menggunakan data historis curah hujan, parameter iklim, dan informasi varietas tanaman yang dikumpulkan dari wilayah Bekasi Utara. Proses pengumpulan data dilakukan dengan menggunakan citra satelit Landsat 8 dan dataset curah hujan CHIRPS yang kemudian diolah menggunakan bahasa pemrograman Python. Metode machine learning diterapkan untuk memproses dan menganalisis data tersebut guna membangun model prediksi yang akurat. Hasil penelitian menunjukkan bahwa sistem prediksi curah hujan yang dikembangkan mampu memberikan prakiraan curah hujan yang cukup akurat. Selain itu, sistem ini juga dapat memberikan rekomendasi tanaman optimal berdasarkan kondisi iklim dan prediksi curah hujan, yang diharapkan dapat membantu petani dalam meningkatkan hasil panen dan ketahanan pangan di wilayah tersebut. Penelitian ini menyimpulkan bahwa integrasi antara data satelit, analisis iklim, dan metode *machine learning* dapat menjadi alat yang efektif dalam membantu pengambilan keputusan pertanian yang lebih baik. Sistem yang dikembangkan juga memiliki potensi untuk diimplementasikan di wilayah lain dengan kondisi serupa, setelah dilakukan penyesuaian data lokal.

Kata Kunci: CHIRPS, Regresi Linear, Machine Learning, Data Satelit, Lansat 8

*Korespondensi Penulis:

E-mail: gintingjusia2@gmail.com

PENDAHULUAN

Indonesia sebagai negara maritim-kontinental memiliki kekayaan uap air karena terletak diantara dua samudra utama, yakni Samudera Pasifik dan Samudera Hindia. Letak geografis Indonesia yang berada pada lintang rendah dan dekat dengan khatulistiwa memberikan energi panas dan sinar matahari yang cukup besar untuk mengangkat uap air ke atmosfer. Hal ini menciptakan variasi cuaca yang beragam di berbagai wilayah Indonesia. Akses yang cepat dan akurat terhadap prakiraan cuaca menjadi sangat penting karena cuaca memiliki peran penting dalam berbagai aktivitas manusia, seperti penentuan waktu tanam (pertanian) dan kesiapan penerbangan (transportasi udara). Curah hujan merupakan salah satu faktor utama yang menentukan hasil panen. Daerah Bekasi Utara memiliki curah hujan yang tinggi yang dapat menyebabkan gagal panen, sedangkan curah hujan yang rendah dapat menyebabkan kekeringan dan hama tanaman. Penting bagi petani di Bekasi Utara untuk mengetahui prediksi curah hujan dan memilih tanaman yang optimal untuk ditanam berdasarkan kondisi cuaca.[1] Sistem prediksi curah hujan dan rekomendasi tanaman dapat membantu petani di Bekasi Utara dalam membuat keputusan yang tepat terkait waktu tanam dan jenis tanaman. Sistem ini dapat memprediksi curah hujan di masa depan dan merekomendasikan tanaman optimal yang sesuai dengan kondisi cuaca tersebut.[2] Regresi linear adalah metode statistik yang dapat digunakan untuk memprediksi nilai variabel dependen berdasarkan satu atau lebih variabel independen. Dalam penelitian ini, regresi linear akan digunakan untuk memprediksi curah hujan dan merekomendasikan tanaman optimal berdasarkan data curah hujan dan data tanaman di masa lalu.[3] Tujuan utamanya adalah membantu petani dalam membuat keputusan yang tepat terkait waktu tanam dan jenis tanaman yang cocok.

TINJUAN PUSTAKA

Curah hujan adalah jumlah air hujan yang jatuh di suatu tempat dalam periode waktu tertentu, biasanya diukur dalam milimeter (mm) atau inci (in) per satuan waktu (jam, hari, bulan, atau tahun). Curah hujan dapat diukur dengan menggunakan alat ukur curah hujan, seperti rain gauge atau pluviometer.[4] Faktor-Faktor yang Mempengaruhi Curah Hujan Curah hujan dipengaruhi oleh beberapa faktor antara lain:

1. Suhu Udara
Semakin tinggi suhu udara, semakin banyak air yang dapat diuapkan. Hal ini menyebabkan peningkatan curah hujan.
2. Kelembaban Udara
Udara yang lembab mengandung banyak uap air. Ketika udara lembab naik ke atmosfer yang lebih dingin, uap air mengembun dan membentuk awan. Awan yang tebal dan basah dapat menghasilkan curah hujan yang tinggi.
3. Tekanan Udara
Tekanan udara yang rendah menyebabkan udara naik ke atmosfer yang lebih dingin. Udara yang naik ini mengembun dan membentuk awan. Awan yang tebal dan basah dapat menghasilkan curah hujan yang tinggi.
4. Kecepatan Angin
Angin yang kencang dapat membawa uap air dari laut ke daratan. Hal ini menyebabkan peningkatan curah hujan di daerah daratan.
5. Arah Angin
Arah angin dapat menentukan dari mana uap air berasal. Arah angin yang membawa uap air dari laut ke daratan dapat menyebabkan peningkatan curah hujan di daerah daratan.
6. Kondisi Geografis

Kondisi geografis, seperti gunung dan pegunungan, dapat mempengaruhi curah hujan. Gunung dan pegunungan dapat menghalangi angin yang membawa uap air, sehingga menyebabkan curah hujan di daerah lereng gunung lebih tinggi dibandingkan daerah lainnya.

Penelitian ini menggunakan data curah hujan di kota Bekasi yang diperoleh dari: Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) dan Tropical Rainfall Measuring Mission (TRMM)

Penelitian ini juga membutuhkan data *Land Surface Temperature(LST)* di Kota Bekasi akan diperoleh dari Landsat 8. Landsat 8 adalah satelit observasi Bumi yang diluncurkan oleh NASA pada 11 Februari 2013, sebagai bagian dari Misi Kontinuitas Data Landsat (LDCM). Satelit ini melanjutkan warisan program Landsat, yang telah menyediakan data kontinu tentang permukaan Bumi sejak 1972. Landsat 8 membawa dua instrumen utama: Operational Land Imager (OLI) dan Thermal Infrared Sensor (TIRS). [5]

Fitur utama landsat 8 :

1. Operational Land Imager (OLI)

Spektrum Band: OLI menangkap data dalam sembilan spektrum band atau Kanal, termasuk band cahaya tampak, inframerah-dekat, dan inframerah-pendek. Ini memungkinkan pengamatan detail dari berbagai jenis penutup lahan, kesehatan vegetasi, dan kualitas air. Resolusi Spasial: OLI menyediakan data dengan resolusi spasial 30 meter untuk sebagian besar band, dengan resolusi 15 meter untuk 7 band pankromatik, yang memungkinkan pemetaan dan analisis yang detail.

2. Thermal Infrared Sensor (TIRS)

Band Termal: TIRS menangkap data dalam dua band inframerah termal, yang penting untuk mengukur suhu permukaan dan fluks panas. Ini sangat berguna untuk mempelajari pulau panas perkotaan, praktik pertanian, dan manajemen sumber daya air.

Korelasi antara data curah hujan dan data Landsat 8 terkait dengan Suhu Permukaan Tanah (LST) yang didapatkan dari citra Landsat 8. LST dipengaruhi oleh berbagai faktor, termasuk tutupan vegetasi, kelembaban tanah, dan kondisi atmosfer. Curah hujan, pada gilirannya, dapat memengaruhi faktor-faktor ini dan akibatnya berdampak pada LST.[6]

Regresi linear adalah model statistik yang digunakan untuk memprediksi nilai variabel dependen (Y) berdasarkan nilai variabel independen (X). Model ini diasumsikan memiliki hubungan linear antara variabel dependen dan independen.[2] Persamaan regresi Linear dapat dilihat pada Rumus (1)

$$Y = \beta_0 + \beta_1 X + \varepsilon \dots\dots\dots (1)$$

Keterangan:

- Y: variabel dependen
- X: variabel independen
- β_0 : konstanta
- β_1 : koefisien regresi
- ε : error

Asumsi regresi linear adalah kondisi yang harus dipenuhi agar model regresi linear dapat memberikan hasil yang akurat dan valid. Berikut adalah penjelasan lebih lanjut mengenai keempat asumsi utama regresi linear:[7]

1. Linearitas:

Hubungan antara variabel dependen (Y) dan variabel independen (X) diasumsikan linear. Artinya, jika X diplot terhadap Y, maka titik-titik data akan membentuk garis lurus.

2. Homoskedastisitas:

Varians error (ε) diasumsikan sama untuk semua nilai X. Artinya, penyebaran error di sekitar garis regresi sama di semua nilai X.

3. Normalitas:

Error (ε) diasumsikan terdistribusi normal. Artinya, error mengikuti kurva lonceng simetris.

4. Independensi:

Error antar observasi diasumsikan tidak saling terkait. Artinya, error pada satu observasi tidak mempengaruhi error pada observasi lain.

5. Non-multikolinearitas:

Variabel independen diasumsikan tidak saling berkorelasi. Artinya, tidak ada hubungan linear antara dua variabel independen

Penerapan Regresi Linear untuk Prediksi Curah Hujan. Regresi linear dapat digunakan untuk memprediksi curah hujan berdasarkan variabel lain seperti:

1. Suhu: Suhu udara dapat mempengaruhi curah hujan. Semakin tinggi suhu, semakin banyak air yang dapat menguap dari permukaan bumi, sehingga meningkatkan kemungkinan terjadinya hujan.
2. Kelembaban: Kelembaban udara menunjukkan jumlah uap air di atmosfer. Semakin tinggi kelembaban, semakin banyak air yang tersedia untuk turun sebagai hujan.
3. Tekanan udara: Tekanan udara dapat mempengaruhi pola cuaca. Tekanan udara yang rendah dikaitkan dengan peningkatan kemungkinan terjadinya hujan.
4. Arah angin: Arah angin dapat membawa uap air dari satu tempat ke tempat lain. Arah angin yang dominan dapat mempengaruhi lokasi terjadinya hujan.

Machine Learning (ML) adalah cabang dari kecerdasan buatan yang memungkinkan sistem komputer untuk belajar dari data dan membuat prediksi atau keputusan tanpa diprogram secara eksplisit. Dalam konteks prediksi curah hujan dan rekomendasi tanaman, ML digunakan untuk mengolah dan menganalisis data historis serta faktor-faktor lingkungan lainnya guna menghasilkan model yang akurat.[5]

2.5. Data Statistik Pertanian Kota Bekasi

Pada Januari 2023, luas lahan pertanian di Kota Bekasi tersisa 307 hektar, yang tersebar di daerah Musitikaya dan Bekasi Utara. Luas lahan pertanian ini mengalami penurunan signifikan dalam beberapa tahun terakhir akibat urbanisasi dan perubahan penggunaan lahan.[8] dapat dilihat pada Gambar 1.

Wilayah Jawa Barat	Luas Panen Tanaman Padi (Ha) (Hektar)				
	Luas Panen (Ha)				
	2023	2022	2021	2020	2019
Kota Bekasi	307	454.07	477.98	483.64	518.27

Gambar 1. Luas Panen

Kebutuhan air tanaman di Kota Bekasi bervariasi tergantung pada jenis tanaman dan kondisi iklim setempat. Tanaman padi, sebagai tanaman utama, membutuhkan air dalam jumlah besar terutama selama fase pertumbuhan vegetatif hingga pembentukan bulir. Rata-rata, tanaman padi memerlukan sekitar 1.200 mm air selama satu musim tanam.

Selain padi, tanaman hortikultura seperti cabai, tomat, dan bayam juga memerlukan irigasi yang memadai. Cabai dan tomat membutuhkan sekitar 600-800 mm air per musim tanam, sedangkan bayam membutuhkan sekitar 300-400 mm air.

Tabel 1 Tanaman Air Rendah

Kategori	Nama	Kebutuhan Air harian	Kebutuhan Air Bulanan
Tanaman Air Rendah	Kangkung	3-5 mm/hari	90-150 mm/bulan
Tanaman Air Sedang	Kubis	5-7 mm/hari	150-210 mm./bulan
Tanaman Air Tinggi	Kedelai	8-10 mm/hari	230-300 mm/bulan

Berdasarkan curah hujan di Bekasi yang berkisar antara 200-400 mm per bulan, tanaman-tanaman yang dapat tumbuh optimal di kota tersebut adalah yang memiliki tingkat toleransi terhadap curah hujan yang bervariasi dan mampu bertahan di iklim tropis. Berdasarkan curah hujan di Bekasi yang berkisar antara 200-400 mm per bulan, tanaman-tanaman yang dapat tumbuh optimal di kota tersebut adalah yang memiliki tingkat toleransi terhadap curah hujan yang bervariasi dan mampu bertahan di iklim tropis. Kategori kebutuhan dapat dilihat pada Tabel 1.

METODOLOGI PENELITIAN

Penelitian ini mengadopsi metode pengembangan yang dikenal sebagai waterfall, sebuah pendekatan yang sering digunakan dalam penelitian deskriptif-kualitatif. Metode ini melibatkan proses pengembangan perangkat lunak secara berurutan, di mana kemajuan dianggap sebagai aliran yang terus bergerak ke bawah, mirip dengan aliran air terjun, melalui tahapan-tahapan perencanaan, pemodelan, implementasi, dan pengujian.[9]

Penelitian yang dilakukan berada di Kota Bekasi dimana peta dari kota Bekasi dapat dilihat pada gambar 2.



Gambar 2. Denah Kota Bekasi

Data curah hujan diperoleh dari Badan Meteorologi, Klimatologi, Geofisika dan Lansat 8 seperti pada Tabel 2.

Tabel 2. Data Curah Hujan BPS

Bulan	Curah Hujan		
	2020	2022	2023
Januari	1089.00	106.60	211.00
Februari	1898.00	150.30	254.00
Maret	655.00	113.20	326.00

April	655.00	316.60	313.00
Mei	433.00	228.50	294.00
Juni	190.00	463.70	321.00
Juli	45.00	358.10	134.00
Agustus	87.00	384.90	145.00
September	217.00	353.70	62.00
Oktober	574.00	492.30	62.00
November	422.00	321.00	1051.00
Desember	407.00	224.10	563.00

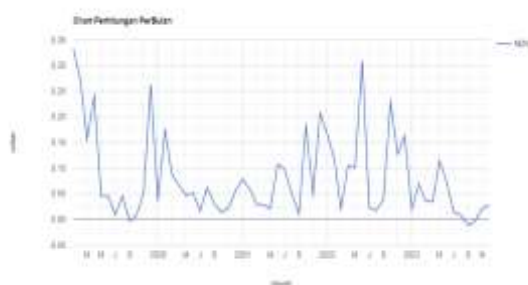
Tabel 3. Data Kelembapan

Bulan	Jumlah Hujan		
	2020	2022	2023
Januari	387.18	271.04	214.75
Februari	446.45	240.58	386.98
Maret	226.10	208.11	218.58
April	222.16	185.46	163.29
Mei	218.60	206.47	132.86
Juni	56.98	116.54	95.80
Juli	89.18	116.88	60.37
Agustus	68.65	113.19	36.25
September	148.70	167.03	8.18
Oktober	173.50	286.21	33.09
November	247.14	378.21	269.19
Desember	215.98	205.64	82.70

Data Landsat 8 yang digunakan dalam penelitian ini adalah citra multispectral. Dapat dilihat pada Gambar 3 dan 4.



Gambar 3. Data LST Lansat 8



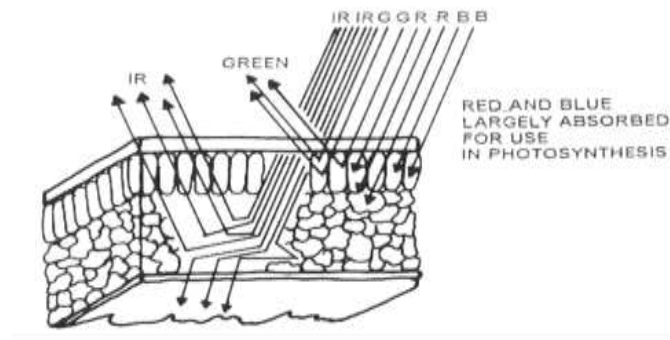
Gambar 4. Data NDVI Lansat 8

Citra multispectral terdiri dari beberapa band atau Kanal dengan panjang gelombang yang berbeda, yaitu:

- Kanal 1: Biru (450-510 nm)
- Kanal 2: Hijau (520-600 nm)
- Kanal 3: Merah (630-690 nm)
- Kanal 4: Near Infrared (NIR) (770-850 nm)
- Kanal 5: Shortwave Infrared 1 (SWIR1) (1560-1660 nm)
- Kanal 6: Thermal Infrared (TIR) (1040-1120 nm)
- Kanal 7: Shortwave Infrared 2 (SWIR2) (2100-2300 nm)
- Kanal 8: Panchromatic (500-680 nm)

Citra Landsat 8 memiliki resolusi spasial 30 meter untuk band 1-7 dan 15 meter untuk band 8. Pemilihan data dari Landsat 8 menggunakan Google Earth Engine. Landsat 8 adalah satelit observasi Bumi yang menyediakan data citra satelit dengan berbagai pita spektral yang dapat digunakan untuk berbagai analisis, termasuk pemantauan cuaca dan iklim. Setelah mengakses data Landsat 8, fokus berpindah ke pemilihan area studi di wilayah tertentu, dalam konteks ini, wilayah Bekasi Utara. Penggunaan Landsat 8 memungkinkan pemantauan yang berkelanjutan selama periode lima tahun untuk memperoleh pemahaman yang komprehensif tentang pola curah hujan, kelembaban, dan suhu tanah di wilayah tersebut. Indeks vegetasi adalah ukuran kehijauan vegetasi yang diperoleh dari pengolahan data digital nilai kecerahan (brightness) dari beberapa kanal data sensor satelit. Untuk memantau vegetasi, dilakukan perbandingan antara tingkat kecerahan kanal cahaya merah (red) dan kanal cahaya inframerah dekat (near infrared). Fenomena ini terjadi karena klorofil menyerap cahaya merah sementara jaringan mesofil daun memantulkan cahaya inframerah dekat, menyebabkan perbedaan besar dalam nilai kecerahan yang diterima sensor satelit pada kedua kanal tersebut. Pada wilayah non-vegetasi, seperti perairan, pemukiman, tanah kosong terbuka, dan area dengan vegetasi yang rusak, nilai rasio ini akan rendah (minimum). Sebaliknya, di area dengan vegetasi yang sangat rapat dan sehat, perbandingan kedua

kanal ini akan sangat tinggi (maksimum), dengan pantulan yang lebih besar pada kanal inframerah dekat dan lebih kecil pada kanal merah. Pola pantulan spektral air menurun pada kedua kanal inframerah dan merah.[10]



Gambar 5. Struktur Tanaman[10][12]

Indeks vegetasi yang digunakan dalam penelitian ini adalah Normalized Difference Vegetation Index (NDVI) , Enhanced Vegetation Index (EVI) dan Normalized Difference Water Index (NDWI)

NDVI:

NDVI adalah salah satu indeks vegetasi yang paling umum digunakan untuk mengukur kesehatan dan vigor vegetasi. NDVI dihitung dengan rumus (2)

$$NDVI = (NIR - RED) / (NIR + RED) \dots\dots\dots (2)$$

Keterangan:

- NIR: Near Infrared (Kanal 5 pada Landsat 8)
- RED: Red (Kanal 4 pada Landsat 8)

Nilai NDVI berkisar antara -1 dan 1. Nilai NDVI yang tinggi menunjukkan vegetasi yang lebat dan sehat, sedangkan nilai NDVI yang rendah menunjukkan vegetasi yang jarang dan tidak sehat.

EVI adalah indeks vegetasi yang lebih canggih daripada NDVI. EVI dihitung dengan rumus (3)

$$EVI = 2.5 * (NIR - RED) / (NIR + 6 * RED - 7.5 * BLUE + 1) \dots\dots\dots (3)$$

Keterangan:

- NIR: Near Infrared (Kanal 5 pada Landsat 8)
- RED: Red (Kanal 4 pada Landsat 8)
- BLUE: Blue (Kanal 1 pada Landsat 8)

Nilai EVI berkisar antara -1 dan 1. Nilai EVI yang tinggi menunjukkan vegetasi yang lebat dan sehat, sedangkan nilai EVI yang rendah menunjukkan vegetasi yang jarang dan tidak sehat.

Perbandingan NDVI dan EVI: EVI memiliki beberapa keuntungan dibandingkan NDVI, yaitu: EVI lebih sensitif terhadap perubahan vegetasi yang lebat dan EVI lebih tahan terhadap pengaruh atmosfer. Selain itu EVI dapat digunakan untuk mengukur vegetasi di daerah dengan tutupan awan yang tinggi.

Normalized Difference Water Index (NDWI) adalah indeks yang digunakan untuk mengukur keberadaan air atau kelembaban dalam vegetasi dan tanah berdasarkan perbedaan reflektansi antara dua wilayah spektral di dalam spektrum elektromagnetik, yaitu wilayah dekat inframerah (NIR) dan reflektansi cahaya di wilayah infrared pendek (SWIR) NDWI sering digunakan dalam pemantauan sumber daya air, seperti penentuan perubahan luas dan kondisi vegetasi di daerah rawa, sungai, dan danau, serta penilaian kelembaban tanah dalam pertanian.

Formula umum untuk menghitung NDWI dapat dilihat pada rumus (4)

$$NDWI = (NIR - SWIR) / (NIR + SWIR) \dots\dots\dots (4)$$

Keterangan:

- NIR: Near Infrared (Kanal 5 pada Landsat 8)
- SWIR: spasial 30 meter untuk visible, NIR, dan SWIR (Kanal 6 pada Landsat 8)

NDWI memiliki rentang nilai dari -1 hingga 1. Nilai-nilai positif menunjukkan keberadaan air atau kelembaban, sementara nilai-nilai negatif menunjukkan daerah tanpa air.

Keunggulan NDWI termasuk sensitivitas terhadap air tanah yang dalam dan kelembaban vegetasi, serta kekurangannya adalah rentang dinamis yang sempit dan sensitivitas terhadap kondisi atmosferik seperti awan dan kabut.

HASIL DAN PEMBAHASAN

Berdasarkan desain sistem yang telah dibuat sebelumnya, pada bagian ini akan diuraikan secara rinci tahapan-tahapan yang diperlukan untuk mengimplementasikan metode pengembangan yang telah ditetapkan. Penjelasan ini mencakup setiap langkah dari proses implementasi, mulai dari persiapan awal hingga penerapan akhir, guna memastikan bahwa semua aspek pengembangan sistem dilakukan sesuai dengan rencana dan mencapai hasil yang diinginkan.

Pengumpulan data yang dilakukan pada penelitian ini berasal dari *dataset* yang diambil dari data BPS Kota Bekasi dan Data Satelit, Data satelit yang dipakai adalah satelit *Landsat 8* dan satelit CHIRPS yang dimana data dari satelit akan diproses menjadi *dataset* CSV menggunakan *Google Earth Engine*.

Wilayah Kota Bekasi didefinisikan menggunakan ROI (Region of Interest). Koleksi data presipitasi dari UCSBCHG/CHIRPS/PENTAD difilter berdasarkan rentang tanggal 1 Januari 2019 hingga 31 Desember 2023. Setiap gambar dalam koleksi ini berisi informasi mengenai curah hujan.

Setiap gambar dalam koleksi ini dipetakan untuk menambahkan informasi bulan dan tahun. Hal ini dilakukan dengan mengakses *metadata system:time_start* dari setiap gambar dan mengubahnya menjadi bulan dan tahun yang sesuai.

Data kemudian diorganisir menjadi koleksi gambar yang disusun berdasarkan bulan dan tahun. Untuk setiap kombinasi bulan dan tahun dari 2019 hingga 2023, data curah hujan dijumlahkan dan diberi label sesuai dengan bulan, tahun, dan tanggal.

Setiap gambar dalam koleksi yang telah diorganisir tersebut dipetakan ke wilayah Bekasi. Setiap fitur dalam wilayah Bekasi diberi atribut tambahan berupa tanggal, bulan, dan tahun. Data curah hujan untuk setiap bulan kemudian dihitung rata-ratanya untuk setiap fitur di wilayah Bekasi.

Koleksi fitur yang telah diproses, yang berisi informasi tentang curah hujan rata-rata bulanan untuk setiap bagian wilayah Bekasi, diekspor ke *Google Drive* dalam format CSV untuk analisis lebih lanjut. Data ini memiliki data curah hujan yang terstruktur dengan baik untuk periode yang telah ditentukan.

Hasil Pengujian



```
def load_data():
    data = pd.read_csv('dataset.csv')
    data['date'] = pd.to_datetime(data['date'], format='%Y-%m-%d').dt.date
    return data

data = load_data().drop(labels = 'year', axis=1)

# call all model
model = pickle.load(open('model.pkl', 'rb'))
model_lst = pickle.load(open('model_lst.pkl', 'rb'))
model_ndvi = pickle.load(open('model_ndvi.pkl', 'rb'))
```

Gambar 6. Load Data

Fungsi *load_data()* pada Gambar 6, membaca file *dataset.csv* dan mengonversi kolom *date* menjadi tipe *datetime*. Data menyimpan hasil dari fungsi *load_data()* setelah menghapus kolom *year*.

Langkah berikutnya adalah memuat model yang telah disimpan menggunakan *pickle*. Persiapan Data dan Antarmuka Pengguna Menampilkan judul dan subjudul pada antarmuka *Streamlit*. Membuat rentang tanggal masa depan mulai dari 2024-01-01 sebanyak 54 periode dengan frekuensi bulanan. *future_rainfall_data* diinisialisasi dengan data historis yang sudah dibaca sebelumnya. Berdasarkan nilai curah hujan yang diprediksi, memberikan rekomendasi tanaman yang optimal.

SIMPULAN

Berdasarkan hasil penelitian dan analisis yang telah dilakukan mengenai prediksi curah hujan menggunakan variabel LST (Land Surface Temperature) dan NDVI (Normalized Difference Vegetation Index) dengan model regresi linear yang dibangun untuk memprediksi curah hujan berdasarkan LST dan NDVI menunjukkan hasil yang cukup baik. Hal ini ditunjukkan oleh nilai Mean Squared Error (MSE) sebesar 3370.697 dan nilai R^2 (Koefisien Determinasi) sebesar 0.507. Nilai R^2 ini menunjukkan bahwa model dapat menjelaskan sekitar 50.7% variasi dalam data curah hujan berdasarkan LST dan NDVI. Analisis korelasi menunjukkan adanya hubungan yang signifikan antara curah hujan, LST, dan NDVI. Meskipun demikian, kompleksitas hubungan ini menunjukkan bahwa faktor-faktor lain mungkin perlu dipertimbangkan untuk meningkatkan akurasi model prediksi.

DAFTAR PUSTAKA

- [1] A. Wulandari, M. Muliadi, dan A. Apriansyah, "Pengaruh Sebaran Uap Air terhadap Curah Hujan di Kalimantan Barat," *PRISMA FISIKA*, vol. 6, no. 3, Okt 2018, doi: 10.26418/pf.v6i3.28709.
- [2] Evi Dewi Sri Mulyani, "Prediksi Curah Hujan Di Kabupaten Majalengka Dengan Menggunakan Algoritma Regresi," *e-Jurnal JUSITI (Jurnal Sistem Informasi dan Teknologi Informasi)*, vol. 8–1, hlm. 67–77, Apr 2019, doi: 10.36774/jusiti.v8i1.602.
- [3] A. Fadholi, S. Meteorologi, dan D. Amir, "Persamaan Regresi Prediksi Curah Hujan Bulanan Menggunakan Data Suhu dan Kelembapan Udara di Ternate," 2013.
- [4] M. Fajri dkk., "Penerapan Model Regression Untuk Prediksi Cuaca Wilayah seberang Ulu 1 Palembang."
- [5] M. D. Islam dkk., "Rapid Rice Yield Estimation Using Integrated Remote Sensing and Meteorological Data and Machine Learning," *Remote Sens (Basel)*, vol. 15, no. 9, Mei 2023, doi: 10.3390/rs15092374.
- [6] W. Ullah dkk., "Analysis of the relationship among land surface temperature (LST), land use land cover (LULC), and normalized difference vegetation index (NDVI) with topographic elements in the lower Himalayan region," *Heliyon*, vol. 9, no. 2, hlm. e13322, Feb 2023, doi: 10.1016/j.heliyon.2023.e13322.
- [7] F. Shi, F. Zhang, N. Shen, dan M. Yang, "Quantifying interactions between slope gradient, slope length and rainfall intensity on sheet erosion on steep slopes using Multiple Linear Regression," *Science of The Total Environment*, vol. 908, hlm. 168090.
- [8] "Statistik Daerah Kota Bekasi 2023".
- [9] A. Amrin, M. D. Larasati, dan I. Satriadi, "Model Waterfall Untuk Pengembangan Sistem Informasi Pengolahan Nilai Pada SMP Kartika XI-3 Jakarta Timur," *Jurnal Teknik Komputer*, vol. 6, no. 1, hlm. 135–140, Jan 2020, doi: 10.31294/jtk.v6i1.6884.
- [10] V. Kevin, S. Que, Y. J. Prasetyo, dan C. Fibriani, "Analisis Perbedaan Indeks Vegetasi Normalized Difference Vegetation Index (NDVI) dan Normalized Burn Ratio (NBR) Kabupaten Pelalawan Menggunakan Citra Satelit Landsat 8"
- [11] Sagaino.S, Mulyana. T.M.S, Suryantara.I.G.N, Ginting.J.A, Adikara, F, "Natural Disaster Event Mapping in West Java Using K-Means Algorithm," *Jurnal Algoritma, Logika dan Komputasi*, vol 5, no 2, 2022
- [12] Ginting.J.A, Jadera,A.M," Analisa Indeks Vegetasi Menggunakan Citra Satelit Lansat 7 dan Lansat 8 menggunakan Metode K-Means di Kawasan Gunung Sinabung", *Indonesian Journal of Computing and Modeling*, vol 1, hal 42-48, 2018
- [13] Herdian.C, Widyanto.S, Ginting, J.A, Geasela,Y.M, Sutrisno.J," The Use of Feature Engineering and Hyperparameter Tuning for Machine Learning Accuracy Optimization: A Case Study on Heart Disease Prediction", *Springer Nature Switzerland*, hal 193-218, 2024