

**PENERAPAN BAYESIAN RIDGE REGRESSION UNTUK PREDIKSI
CURAH HUJAN JANGKA PENDEK**

SKRIPSI

DISUSUN OLEH

SRI NINGRUM

2209020038



UMSU

Unggul | Cerdas | Terpercaya

**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA**

MEDAN

2026

**PENERAPAN BAYESIAN RIDGE REGRESSION UNTUK PREDIKSI
CURAH HUJAN JANGKA PENDEK**

SKRIPSI

**Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer
(S.Kom) dalam Program Studi Teknologi Informasi, pada Fakultas Ilmu Komputer
dan Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara**

SRI NINGRUM

NPM. 2209020038

**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA**

MEDAN

2026

LEMBAR PENGESAHAN

Judul Skripsi : Penerapan Bayesian Ridge Regression Untuk Prediksi
Curah Hujan Jangka Pendek
Nama Mahasiswa : Sri Ningrum
NPM : 2209020038
Program Studi : Teknologi Informasi

Menyetujui
Komisi Pembimbing



(Farid Akbar Siregar, S.Kom., M.Kom.)
NIDN. 0104049401

Ketua Program Studi



(Fatma Sari Hutagalung, S.Kom, M.Kom)
NIDN. 0117019301



(Dr. Al-Rizki Hidayati, S.Kom., M.Kom.)
NIDN. 0127099201

PERNYATAAN ORISINALITAS

PENERAPAN BAYESIAN RIDGE REGRESSION UNTUK PREDIKSI CURAH HUJAN JANGKA PENDEK

SKRIPSI

Saya menyatakan bahwa karya tulis ini adalah hasil karya sendiri, kecuali beberapa kutipan dan ringkasan yang masing-masing disebutkan sumbernya.

Medan, ~~27~~ Juni 2026

Yang membuat pernyataan



Sri Ningrum

NPM. 2209020038

**PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN
AKADEMIS**

Sebagai sivitas akademika Universitas Muhammadiyah Sumatera Utara, saya bertanda tangan dibawah ini:

Nama	: Sri Ningrum
NPM	: 2209020038
Program Studi	: Teknologi Informasi
Karya Ilmiah	: Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Muhammadiyah Sumatera Utara Hak Bebas Royalti Non-Eksekutif (*Non-Exclusive Royalty free Right*) atas penelitian skripsi saya yang berjudul:

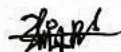
**PENERAPAN BAYESIAN RIDGE REGRESSION UNTUK PREDIKSI
CURAH HUJAN JANGKA PENDEK**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non-Eksekutif ini, Universitas Muhammadiyah Sumatera Utara berhak menyimpan, mengalih media, memformat, mengelola dalam bentuk database, merawat dan mempublikasikan Skripsi saya ini tanpa meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemegang dan atau sebagai pemilik hak cipta.

Demikian pernyataan ini dibuat dengan sebenarnya.

Medan, 24 Juni 2026

Yang membuat pernyataan



Sri Ningrum

NPM. 2209020038

RIWAYAT HIDUP

DATA PRIBADI

Nama Lengkap : Sri Ningrum
Tempat dan Tanggal Lahir : Bintang Bayu, 02 oktober 2004
Alamat Rumah : Bandar Pinang Kebun
Telepon/Faks/HP : 081362736292
E-mail : sningrum814@gmail.com
Instansi Tempat Kerja : -
Alamat Kantor : -

DATA PENDIDIKAN

SD	:101988 Bandar Pinang	TAMAT : 2016
SMP	: MTS GUPPI Bintang Bayu	TAMAT : 2019
SMA	: SMA N1 Bintang Bayu	TAMAT : 2022

KATA PENGANTAR



Pendahuluan

Puji syukur kehadiran Allah SWT atas segala rahmat dan karunia nya sehingga penulis dapat menyelesaikan skripsi ini yang berjudul “Penerapan Bayesian Ridge Regression Untuk Prediksi Curah Hujan Jangka pendek”. Penulis menyampaikan terimakasih kepada berbagai pihak atas dukungan serta doa dalam proses penyelesaian skripsi. Penulis juga mengucapkan terima kasih kepada:

1. Bapak **Prof. Dr. Akrim, M.Pd.**, Rektor Universitas Muhammadiyah Sumatera Utara (UMSU).
2. Bapak **Dr. Al-Khowarizmi, S.Kom., M.Kom.** Dekan Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU.
3. Ibu **Dr. Firahmi Rizky, S.Kom., M.Kom.** Wakil Dekan I Fakultas Ilmu Komputer dan Teknologi Informasi
4. Bapak **Mhd. Basri, S.Si, M.Kom.** Wakil Dekan III Wakil Dekan III Fakultas Ilmu Komputer dan Teknologi Informasi.
5. Ibu **Fatma Sari Hutagalung, S.Kom., M.Kom.** Ketua Program Studi Teknologi Informasi.
6. Bapak **Okvi Nugroho, S.Kom., M.Kom.** Sekretaris Program Studi Teknologi Informasi.
7. Bapak **Farid Akbar Siregar, S.Kom., M.Kom.** Dosen pembimbing skripsi.
8. Teruntuk bapak tercinta, bapak Adi Suprianto, terimakasih atas setiap langkah perjuangan yang bapak lakukan tanpa pernah mengeluh. Dalam diam, bapak selalu menjadi kekuatan yang menjaga penulis tetap berdiri ketika lelah, dan menjadi alasan untuk penulis terus melangkah ketika ingin menyerah. Doa dan pengorbanan bapak adalah hal yang tidak akan pernah mampu penulis balas sepanjang hidup.

9. Ibu tercinta, almh ibu Suyanti S.Pd., meski raga mamak telah tiada, cinta mamak tidak pernah benar-benar pergi. mamak tetap hidup dalam setiap doa, dalam setiap langkah, dan dalam setiap usaha penulis hari ini. Rindu ini tidak akan pernah selesai, namun penulis percaya bahwa setiap pencapaian ini adalah bukti bahwa doa ibu semasa hidup masih terus mengiringi penulis dari tempat yang paling idah di sisi-nya.
10. Untuk Nenek dan Kakek tercinta, terimakasih atas kasih sayang yang tulus, doa yang tidak pernah putus, serta perhatian yang selalu menghangatkan hati penulis. Nenek dan Kakek adalah tempat pulang yang penuh ketenangan, yang selalu menerima penulis dengan cinta tanpa syarat, serta memberikan nasihat dan dukungan yang menjadi pegangan dalam hidup penulis hingga saat ini.
11. Terimakasih untuk adikku Sitik Cholicah, yang sering bikin kesal, tetapi juga sering bikin ketawa. Walaupun menjengkelkan tetap tidak bisa dipungkiri kalau kamu bagian penting dalam hidup penulis.

Tidak ada kata yang mampu menggambarkan besarnya cinta, doa, dan pengorbanan kalian semua, Semoga Allah SWT senantiasa melimpahkan kesehatan, umur panjang, dan kebahagiaan kepada Bapak, Nenek, Kakek, dan Adik, serta memberikan tempat terbaik dan termulia bagi mamak tercinta di sisi-nya.
12. Terimakasih untuk Aulia Syafitri, teman yang sudah penulis anggap seperti kakak sendiri, yang selalu memberikan perhatian, dukungan, dan semangat tanpa henti. Terimakasih sudah menjadi tempat berkeluh kesah, menguatkan disaat lelah, dan selalu ada dalam setiap proses perjuangan ini.
13. Terimakasih juga kepada Nurma Linda, teman yang menjalani hubungan ldr, namun tetap selalu hadir dalam bentuk doa, perhatian, dan semangat yang tidak pernah berkurang. Jarak tidak pernah menjadi penghalang untuk tetap saling mendukung dan menguatkan.
14. Kepada pihak-pihak yang tidak bisa penulis sebutkan satu persatu, dan terimakasih kepada teman seperjuangan. Terkhusus Adinda Novriza, Malika Rievy Sakina Pohan, Nafisah Adawiyah, Silvi Handayani Maharaja, Dea Salsabilla Gorat yang telah bersama-sama melewati masa perkuliahan

hingga proses penyusunan skripsi ini. Terimakasih atas kebersamaan, dukungan, dan semangat yang selalu diberikan, baik dalam keadaan senang maupun sulit.

15. Dengan penuh rasa syukur, penulis mengucapkan terimakasih kepada diri sendiri yang telah mampu bertahan, berjuang, dan tidak menyerah ditengah berbagai rintangan selama proses penyusunan skripsi ini. Terimakasih untuk diri ini yang telah melangkah meski lelah, tetap bertahan meski banyak tekanan, dan tetap berusaha meski sering merasa capek. Tidak semua perjalanan mudah, namun diri ini telah memilih untuk terus mencoba hingga akhirnya mampu menyelesaikan tugas ini. Terimakasih telah berani menghadapi ketakutan, menguatkan diri disaat rapuh, dan percaya bahwa setiap proses akan membawa hasil terbaik pada waktunya. Semua usaha, air mata, dan perjuangan yang telah dilalui adalah bukti bahwa diri ini tidak benar-benar menyerah.

Demikian ucapan terimakasih ini penulis sampaikan. Penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna, sehingga sangat diharapkan kritik dan saran yang membangun perbaikan dimasa yang akan datang. Semoga skripsi ini dapat memberikan manfaat bagi pembaca dan semua pihak yang membutuhkan.

PENERAPAN BAYESIAN RIDGE REGRESSION UNTUK PREDIKSI CURAH HUJAN JANGKA PENDEK

ABSTRAK

Penelitian ini bertujuan untuk menerapkan metode *Bayesian Ridge regression* dalam memprediksi curah hujan jangka pendek berdasarkan data historis curah hujan dan variabel meteorologi pendukung. Prediksi curah hujan merupakan salah satu aspek penting dalam bidang klimatologi dan mitigasi bencana karena memiliki pengaruh terhadap sektor pertanian, pengelolaan sumber daya air, dan aktivitas masyarakat. Metode *Bayesian Ridge regression* dipilih karena mampu mengatasi permasalahan multikolinearitas antar variabel prediktor serta mengurangi resiko overfitting melalui pendekatan probabilistik Bayesian. Data penelitian diperoleh dari Kaggle yang terdiri atas 516 data observasi dan ditingkatkan melalui proses augmentasi hingga mencapai 1000 dataset. Variabel yang digunakan meliputi curah hujan periode sebelumnya, suhu udara, kelembapan udara, tekanan udara, dan kecepatan angin. Tahapan penelitian meliputi preprocessing data, standardisasi menggunakan StandardScaler, pembagian data training dan testing dengan rasio 80:20, pelatihan model, serta evaluasi menggunakan *Mean Squared Error (MSE)*, *Mean Absolute Error (MAE)*, dan *R-Squared (R^2)*. Hasil penelitian Menunjukkan bahwa model Bayesian Ridge Regression mampu menghasilkan prediksi curah hujan dengan nilai MAE sebesar 15,70 mm, MSE sebesar 459.09, dan R^2 sebesar -0,0127. Meskipun model belum mampu menangkap seluruh pola non-linear dan lonjakan curah hujan ekstream, metode ini tetap menunjukkan kestabilan dalam memprediksi tren umum curah hujan harian. Dengan demikian, Bayesian Ridge regression dapat digunakan sebagai pendekatan alternatif dalam sistem prediksi curah hujan jangka pendek berbasis machine learning.

Kata Kunci : *Bayesian Ridge Regression*, prediksi curah hujan, machine learning, data meteorologi, *Mean Squared Error (MSE)*.

APPLICATION OF BAYESIAN RIDGE REGRESSION FOR SHORT-TERM RAINFALL PREDICTION.

ABSTRACT

This study aims to apply the *Bayesian Ridge Regression* method to predict short-term rainfall based on historical rainfall data and supporting meteorological variables. Rainfall prediction is a crucial aspect in climatology and disaster mitigation because it impacts the agricultural sector, air resource management, and community activities. The *Bayesian Ridge Regression* method was chosen because it can address multicollinearity issues among predictor variables and reduce the risk of overfitting through a Bayesian probabilistic approach. The research data were obtained from Kaggle, consisting of 516 observational datasets, and were increased through an augmentation process to reach 1,000 datasets. The variables used included previous rainfall, air temperature, air humidity, air pressure, and wind speed. The research stages included data preprocessing, standardization using *StandardScaler*, dividing the training and test data with an 80:20 ratio, model training, and evaluation using *Mean squared Error (MSE)*, *Mean Absolute Error (MAE)*, and *R-Squared (R^2)*. The results showed that the *Bayesian Ridge Regression* model was able to produce rainfall predictions with an MAE of 15.70 mm, an MSE of 459.09, and an R^2 of -0.0127. Although the model was unable to capture all non-linear patterns and the flow of extreme rainfall, this method still demonstrated stability in predicting general daily rainfall trends. Thus, Bayesian Ridge Regression can be used as an alternative approach in machine learning-based short-term rainfall prediction systems.

Keywords : *Bayesian Ridge Regression*, rainfall prediction, machine learning, meteorological data, *Mean Squared Error (MSE)*.

DAFTAR ISI

LEMBAR PENGESAHAN	i
PERNYATAAN ORISINALITAS	ii
PERNYATAAN PERSETUJUAN PUBLIKASI	iii
RIWAYAT HIDUP	iv
KATA PENGANTAR	v
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
BAB I. PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah.....	4
1.3 Batasan Masalah	4
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	6
BAB II. LANDASAN TEORI	8
2.1 Fundamental Metode Bayesian.....	8
2.2 Parameter, data, dan Prediksi Dalam Bayesian	9
2.2.1 data.....	10
2.2.2 Bayesian Rule	11
2.2.3 Bayesian Data Analysis	12
2.2.4 Prediksi	12
2.2.5 Likelihood.....	14
2.3 Metode Bayesian Ridge Regression	14
2.3.1 Keunggulan Bayesian Ridge Regression.....	16
2.4 Penelitian Terdahulu.....	16
BAB III. METODOLOGI PENELITIAN	22
3.1 Jenis dan Pendekatan Penelitian	22
3.2 Data Penelitian	23
3.2.1 Sumber Data	23
3.3 Variabel Penelitian.....	24
3.3.1 jenis Data	25
3.3.2 Pembagian Data	26
3.4 Prosedur penelitian	26
3.4.1 Identifikasi Masalah.....	26
3.4.2 Kajian Literatur	27
3.4.3 Pengumpulan Data.....	27
3.4.4 Pengolahan Awal Data.....	27
3.4.5 Pembagian Data Penelitian	28
3.4.6 Proses Pelatihan Model.....	28
3.5 Analisis Hasil Dan Kesimpulan	28
3.6 Tahapan Penelitian.....	29
3.6.1 Diagram Alur Penelitian	29
3.6.2 Flowchart Bayesian Ridge Regression	30
3.6.3 Usecase	31
3.6.4 Diagram Activity	33

3.6.5 Rancangan UI/UX (User Interface/User Experience).....	34
BAB IV HASIL DAN PEMBAHASAN.....	40
4.1 Dataset Penelitian	40
4.2 Preprocessing Data	41
4.3 Implementasi Bayesian Ridge Regression.....	43
4.4 Pengujian Model.....	44
4.5 Evaluasi Model	47
4.6 Analisis Hasil Penelitian	51
4.7 Implementasi Sistem.....	52
BAB V KESIMPULAN DAN SARAN.....	55
5.1 Kesimpulan.....	55
5.2 Saran.....	56
DAFTAR PUSTAKA	57
LAMPIRAN	A

DAFTAR TABEL

Tabel 1. 1 Peneliti Terdahulu	16
Tabel 4. 1 Representasi Label.....	40
Tabel 4. 2 Fitur yang digunakan.....	41
Tabel 4. 3 Pembagian Dataset	43
Tabel 4. 4 Tabel hasil Pengujian	45
Tabel 4. 5 Fitur Yang Mempengaruhi	46
Tabel 4. 6 Hasil Evaluasi Bayesian	47
Tabel 4. 7 b Perhitungan Manual (Data Asli).....	50

DAFTAR GAMBAR

Gambar 2. 1 Tiga Makna Parameter Dalam Analisis Bayesian	10
Gambar 3. 1 Diagram Alur Penelitian	29
Gambar 3. 2 Flowchart Bayesian Ridge Regression	30
Gambar 3. 3 Usecase Aktor.....	32
Gambar 3. 4 Diagram Activity	33
Gambar 3. 5 Tampilan Dashboard Sistem Prediksi Curah Hujan Jangka Pendek	34
Gambar 3. 6 Tampilan Input Data Curah Hujan Jangka Pendek.....	35
Gambar 3. 7 Tampilan Pra-pemrosesan Data	36
Gambar 3. 8 Tampilan Pembagian Data.....	37
Gambar 3. 9 Tampilan Pelatihan Model Bayesian Ridge Regression.....	38
Gambar 3. 10 Tampilan Hasil Sistem Prediksi Curah Hujan Jangka Pendek	39
Gambar 4. 1 Grafik Perbandingan Nilai Aktual dan Prediksi Curah Hujan.....	49
Gambar 4. 2 Halaman Dashboard.....	52
Gambar 4. 3 Tampilan Prediksi Curah Hujan	53
Gambar 4. 4 Tampilan Mengenai Sistem	54

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Hujan adalah suatu proses jatuhnya air dari udara ke bumi. Air yang jatuh dapat berbentuk cair maupun padat. Hujan terjadi karena menguapnya air sebagai akibat pemanasan sinar matahari. (Astarini & Adriat, 2022). Curah hujan harian adalah curah hujan yang tercatat pada stasiun pengamatan curah hujan setiap hari (selama 24 jam). Data curah hujan biasanya dipakai untuk simulasi kebutuhan air tanaman, simulasi operasi waduk. Pada umumnya curah hujan diprediksi dengan metode klasik.

Metode klasik dalam pembangunan model regresi, seperti pemilihan subset variabel (*subset selection*), banyak digunakan diberbagai bidang, salah satunya adalah untuk meprediksi curah hujan. Namun demikian, metode-metode tersebut memiliki sejumlah kelemahan yaitu estimasi koefisien sangat rentan ketika terdapat multikolonearitas dalam suatu himpunan data. Permasalahan multikolinearitas dapat menyebabkan ketidakstabilan estimasi parameter regresi sehingga menurunkan kemampuan interpretasi dan prediksi model. Ridge Regression dikembangkan sebagai salah satu pendekatan regularisasi yang efektif untuk mengatasi masalah tersebut. (Schreiber-Gregory, 2018)

Dalam kondisi tersebut, estimasi koefisien dapat menjadi tidak stabil, bernilai sangat besar secara tidak wajar, serta mengalami perubahan yang ekstrem dalam model. Bahkan, dalam beberapa kasus, tanda (positif atau negatif) dari koefisien dapat berubah. (Kruschke & Liddell, 2018)

Sebagai alternatif terhadap dari penggunaan metode klasik, kita dapat mempertimbangkan beberapa model dalam memprediksi data baru. Berbagai jenis model memiliki berbagai metode pengujian yang berbeda untuk menilai kemampuan prediktifnya.(Kruschke & Liddell, 2018)

Bayesian Ridge Regression (BRR) merupakan sebuah modifikasi dari metode Regresi Ridge yang menerapkan prinsip Bayesian dalam perhitungan parameter modelnya. Tidak sama dengan *Regresi Ridge* biasa yang hanya menghasilkan estimasi nilai tetap, BRR menyajikan distribusi probabilitas untuk setiap koefisien, sehingga lebih tahan terhadap masalah overfitting dan mampu memberikan rentang akurasi pada hasil perkiraan. Metode ini sangat sesuai untuk meramalkan curah hujan dalam jangka waktu singkat (seperti 1 hingga 7 hari mendatang) karena kemampuannya mengelola data deret waktu yang fluktuatif tinggi, termasuk yang terpengaruh oleh variabel cuaca masa lalu.

Dalam metode BRR, *Mean Squared Error (MSE)* dapat dinyatakan sebagai jumlah dari kuadrat bias dan varians ($MSE = \text{bias}^2 + \text{varians}$), agar model Ridge menghasilkan akurasi yang lebih baik dibandingkan model klasik, peningkatan bias yang terjadi harus diimbangi oleh penurunan varians yang cukup besar. (Hodson et al., 2021)

Dalam penelitian ini saya memilih untuk menghitung *Mean Squared Error (MSE)* sebagai ukuran kemampuan prediksi untuk setiap model. Model yang memiliki nilai MSE paling rendah dianggap sebagai model yang paling baik dalam melakukan prediksi. MSE menghitung rata-rata dari kuadrat selisih antar nilai respons sebenarnya dan nilai respons hasil prediksi. Mean Squared Error (MSE) merupakan salah satu metrik evaluasi yang paling banyak digunakan dalam

pemodelan prediktif karena mampu memberikan penalti yang lebih besar terhadap kesalahan prediksi ekstream. (Samadi-koucheksaraee & Chu, 2024) Semakin kecil nilai MSE semakin dekat nilai prediksi dengan nilai sebenarnya, sehingga semakin baik kecocokan (*fit*) model tersebut.

Penelitian terdahulu oleh Siswanto dkk. (2025) menunjukkan bahwa Linear *Ridge Regression* dengan *MM-Estimator (LRMM)* mampu meningkatkan akurasi peramalan curah hujan melalui pendekatan *statistical downscaling*, dengan hasil yang kuat terhadap multikolinearitas dan outlier. Model ini menghasilkan kolerasi tinggi (0,94) dengan data aktual, RMSE 97,26 dan koefisien determinasi 86,57%, serta mampu menggambarkan pola curah hujan musiman di Kabupaten Pangkep dengan baik. (SISWANTO et al., 2025)

Penelitian selanjutnya oleh Mahendi alvines (2025) Ridge Regression memberikan performa terbaik pada data uji dengan RMSE 14,62 mm, meskipun nilai R^2 negatif (-0,69) menandakan keterbatasan model dalam menangkap variabilitas penuh data curah hujan. Distribusi data sangat tidak seimbang, Dimana 82% hari memiliki curah hujan (>50 mm), sehingga semua model cenderung under-prediction pada hujan lebat. Model ensemble seperti Random Forest dan Gradient Boosting menunjukkan stabilitas yang lebih tinggi pada validasi silang (deviasi $RMSE < 5$ mm), sementara analisis fitur mengindikasikan bahwa kelembapan rata-rata dan curah hujan harian sebelumnya (RR_LAG_1) merupakan predictor penting dalam memodelkan curah hujan harian di Sumatera Selatan. (Alvines et al., 2025)

Penelitian berikutnya ini bertujuan untuk menilai performa *Bayesian Ridge Regression* dibandingkan dengan teknik regresi lainnya dalam meramalkan curah hujan jangka pendek di Lokasi penelitian berbeda, yang memiliki ciri iklim dan

fluktuasi curah hujan khusus. Selain itu, penelitian ini juga fokus pada evaluasi ketepatan ramalan serta kapasitas model untuk mengelola ketidak pastian data cuaca, sehingga diharapkan bisa memberikan pemahaman yang lebih lengkap tentang keefektifan *Bayesian Ridge Regression* dalam konteks peramalan curah hujan jangka pendek.

1.2. Rumusan Masalah

Rumusan masalah ini bertujuan untuk mengukur Tingkat efektivitas metode *Bayesian Ridge Regression* dalam menghasilkan prediksi curah hujan jangka pendek. Efektivitas model dinilai menggunakan metrik *Mean Squared Error (MSE)*, yang mengukur rata-rata kuadrat selisih antara nilai curah hujan actual dan hasil prediksi. Semakin kecil nilai MSE yang diperoleh, maka semakin baik kemampuan model dalam memprediksi curah hujan.

1.3. Batasan Masalah

1. Penelitian ini hanya membahas prediksi curah hujan jangka pendek, yaitu dalam rentang waktu harian hingga mingguan.
2. Data yang digunakan merupakan data historis curah hujan dan variable cuaca pendukung yang diperoleh dari hasil pengamatan. Jumlah data yang digunakan sebanyak 1000 dataset, yang disusun dalam deret waktu dan dimanfaatkan sebagai dasar dalam proses pemodelan dan prediksi curah hujan jangka pendek.
3. Metode yang digunakan dalam penelitian dibatasi pada *Bayesian Ridge Regression*, tanpa membandingkan dengan metode lainnya.

4. Input yang digunakan dalam penelitian ini dibatasi pada data historis curah hujan dan variabel meteorologi pendukung yang tersedia pada periode penelitian. Variabel input meliputi curah hujan periode sebelumnya, suhu udara, kelembapan udara, dan tekanan udara. Data input disusun dalam bentuk data deret waktu (*time series*) dengan jumlah total sebanyak 516 data observasi. Kemudian untuk memvariasi jumlah dataset dilakukan augmentasi data dimana total dataset mencapai 1000 dataset. Penelitian ini tidak menggunakan variabel tambahan seperti data angin, radiasi matahari, data satelit, maupun data spasial.
5. Output dari model *Bayesian Ridge Regression* dalam penelitian ini dibatasi pada nilai prediksi curah hujan jangka pendek dalam satuan milimeter (mm). prediksi yang dihasilkan merupakan prediksi curah hujan untuk periode waktu selanjutnya (harian hingga mingguan) berdasarkan data input yang telah diproses. Penelitian ini tidak membahas kejadian ekstrem, maupun sistem peringatan dini.
6. Penelitian ini tidak membahas pengaruh iklim jangka Panjang, melainkan hanya pada pola curah hujan dalam periode penelitian.

1.4. Tujuan Penelitian

Tujuan Penelitian ini adalah untuk merancang dan membangun sistem website prediksi curah hujan jangka pendek berbasis bahasa pemrograman Python dengan menerapkan metode *Bayesian Ridge Regression* guna mengolah data historis curah hujan dan variabel meteorologi pendukung seperti suhu udara, kelembapan udara, tekanan udara, curah hujan periode sebelumnya, serta kecepatan

angin. Sistem yang dikembangkan diharapkan mampu melakukan proses pra-pemrosesan data, pembagian data training dan testing, pelatihan model, serta evaluasi hasil prediksi menggunakan matrik *Mean Squared Error (MSE)* secara terintegrasi, sehingga menghasilkan prediksi curah hujan yang akurat, stabil, dan dapat mendukung pengambilan keputusan dalam pengelolaan sumber daya air serta mitigasi risiko hidrometeorologi.

1.5. Manfaat Penelitian

a. Bagi Instansi Terkait

Hasil penelitian ini dapat di manfaatkan oleh Lembaga terkait, seperti BMKG dan instansi pengelola sumber daya air, sebagai sumber daya air, sebagai sumber informasi pendukung dalam menghasilkan prediksi curah hujan jangka pendek yang lebih akurat guna menunjang pengambilan Keputusan.

b. Bagi Pembaca

Menambah pemahaman tentang pendekatan *Bayesian Ridge Regression* dalam memprediksi curah hujan, khususnya metode *Bayesian Ridge Regression*, dapat diterapkan dalam memprediksi curah hujan jangka pendek yang memiliki Tingkat fluktuasi dan ketidakpastian tinggi.

c. Manfaat Bagi Diri Sendiri

Penelitian ini memberikan kontribusi bagi penulis dalam memperdalam pemahaman serta meningkatkan keterampilan dalam penerapan metode *Bayesian Ridge Regression* untuk prediksi curah hujan jangka pendek. Melalui penelitian ini, penulis juga memperoleh pengalaman dalam pengolahan dan analisis data cuaca, penafsiran hasil pemodelan statistik, serta penguatan kemampuan berpikir kritis dan

terstruktur dalam menyelesaikan permasalahan penelitian. Pengetahuan dan pengalaman yang diperoleh diharapkan dapat menjadi bekal yang bermanfaat bagi pengembangan akademik maupun professional penulis di bidang analisis dan prediksi cuaca

BAB II

LANDASAN TEORI

2.1 Fundamental Metode Bayessian

Metode Bayesian Adalah pendekatan statistic yang menggunakan Teorema Bayes untuk melakukan inferensi terhadap parameter atau model tertentu. Inferensi Bayesian di lakukan dengan proses menyesuaikan model probabilitas dengan sekumpulan data dan merangkum hasilnya dengan distribusi probabilitas pada parameter model dan pada besaran yang tidak teramati seperti prediksi untuk pengamatan baru. (Gelman et al., 1995)

Proses analisis data Bayesian dapat dibagi kedalam tiga langkah utama yaitu:

1. penyusunan model probabilitas secara lengkap, yakni membangun distribusi probabilitas gabungan untuk seluruh variabel yang dapat diamati maupun tidak dapat diamati dalam suatu permasalahan. Model yang dibentuk harus konsisten dengan pengetahuan ilmiah yang mendasari permasalahan serta sesuai dengan proses pengumpulan data yang dilakukan.
2. pengondisian terhadap data yang diamati, yaitu menghitung dan menafsirkan distribusi posterior. Distribusi posterior merupakan diistribusi probabilitas bersyarat dari parameter atau kuantitas yang tidak teramati dan menjadi fokus utama penelitian, berdasarkan data yang telah diamati.
3. Evaluasi kesesuaian model implikasi dari distribusi posterior yang dihasilkan, meliputi penilaian apakah model yang digunakan telah sesuai dengan data, apakah Kesimpulan yang diperoleh masuk akal secara substantif, serta sejauh mana hasil analisis terhadap asumsi pemodelan pada Langkah pertama.

Apabila diperlukan, model dapat dimodifikasi atau diperluas, kemudian ketiga Langkah dapat diulangi. (Gelman et al., 1995)

Motivasi utama pentingnya pendekatan Bayesian Adalah karena pendekatan ini memfasilitasi interpretasi kesimpulan statistic yang lebih intuitif dan sesuai dengan penalaran akal sehat. Sebagai contoh, interval Bayesian (interval probabilitas) untuk suatu parameter yang tidak diketahui dapat secara langsung diinterpretasikan sebagai memiliki probabilitas tinggi dalam memuat nilai parameter tersebut. Hal ini berbeda dengan interval frequentist (interval kepercayaan), yang secara ketat hanya dapat interpretasikan dalam konteks frekuensi keberhasilan dari serangkaian inferensi serupa yang dilakukan berulang kali.

2.2 Parameter, Data, Dan Prediksi Dalam Bayesian

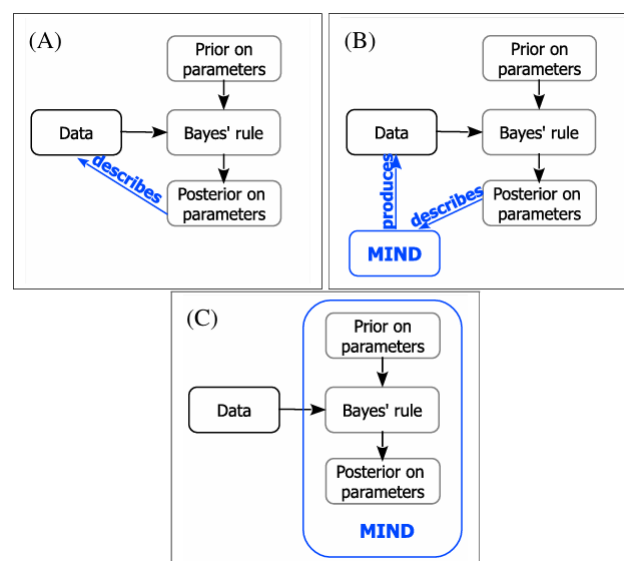
Sebagai notasi umum, kita menggunakan θ untuk menunjukkan besaran vector yang tidak dapat diamati atau parameter populasi yang diminati, seperti parameter distribusi curah hujan (misalnya rata-rata, varians, atau parameter lain yang menggambarkan pola curah hujan) pada wilayah tertentu. Simbol y digunakan untuk menunjukkan data curah hujan yang diamati, misalnya jumlah curah hujan harian, bulanan, atau tahunan yang tercatat pada stasiun pengamatan. Selanjutnya, simbol $\tilde{y}$ menunjukkan besaran yang tidak diketahui tetapi berpotensi dapat diamati, atau curah hujan yang akan terjadi di bawah kondisi iklim tertentu. (Gelman et al., 1995)

Secara umum, simbol-simbol tersebut mempresentasikan besaran multivariat, terutama Ketika data curah hujan diamati pada banyak waktu atau

Lokasi. Kita biasanya menggunakan huruf Yunani atau parameter, huruf Romawi besar untuk matriks yang diamati atau dapat diamati (dan terkadang matriks), serta huruf Romawi besar untuk matriks yang diamati atau dapat diamati. Dalam notasi matriks, vector dianggap sebagai vector kolom; sebagai contoh, jika U Adalah vector dengan n komponen yang mempresentasikan curah hujan pada n periode waktu, maka $u^T u$ merupakan sebuah skalar dan uu^T merupakan matriks berukuran $n \times n$. Parameter metode Bayesian dapat dilihat pada **Gambar 2.1**.

2.2.1 Data

Data disebut juga sebagai distribusi pior, memiliki fungsi sebagai asumsi awal terhadap parameter model sebelum data dianalisis. Dalam Bayesian Ridge Regression, parameter regresi umumnya di asumsikan mengikuti distribusi normal guna mengurangi kemungkinan terjadinya overfitting.



Gambar 2. 1 Tiga Makna Parameter Dalam Analisis Bayesian

(A) Analisis data Bayesian Generik, di mana parameter model menggambar data tanpa merujuk pada proses pembentukannya.

(B) Analisis Psikometrik Bayesian, di mana data dianggap dihasilkan oleh pikiran dan parameter merepresentasikan aspek-aspek pikiran.

(C) Model pikiran Bayesian, di mana pikiran dimodelkan sebagai ahli statistik Bayesian yang mengambil data dari dunia dan memperbarui parameter yang merepresentasikan pengetahuan. (Kruschke & Liddell, 2018)

2.2.2 Bayesian Rule

Untuk membuat pernyataan probabilitas tentang θ berdasarkan data y , kita harus memulai dengan suatu model yang menyediakan distribusi probabilitas gabungan bagi θ dan y . Distribusi probabilitas gabungan tersebut, baik dalam bentuk massa probabilitas maupun fungsi densitas, dapat dinyatakan sebagai hasil perkalian dua distribusi, yaitu distribusi prior $p(\theta)$ dan distribusi sampling (atau distribusi data) $p(y | \theta)$.

$$p(\theta, y) = p(\theta)p(y|\theta)$$

dengan hanya melakukan pengondisian pada nilai data y yang diketahui, menggunakan sifat dasar probabilitas bersyarat yang dikenal sebagai aturan Bayes, Diperoleh kerapatan posterior :

$$p(\theta|y) = \frac{p(\theta, y)}{p(y)} = \frac{p(\theta)p(y|\theta)}{p(y)}$$

dengan penjumlahan mencakup semua kemungkinan nilai θ (atau $p(y) = \int p(\theta)p(y|\theta)d\theta$. Lakukan dalam kasus 6 kontinu). Bentuk Ekuivalen dari (1.1) menghilangkan factor $p(y)$, yang tidak bergantung pada θ dan, dengan y tetap, dapat dianggap sebagai konstanta, menghasilkan densitas posterior yang tidak dinormalisasi, yang merupakan sisi kanan dari (1.2). (Gelman et al., 1995)

$$p(\theta|y) \propto p(\theta)p(y|\theta)$$

2.2.3 Bayesian data analysis

Bentuk ekuivalen dari persamaan (1.2) diperoleh dengan menghilangkan faktor $p(y)$, karena factor tersebut tidak bergantung pada θ dan, untuk y yang tetap, dapat diperlakukan sebagai konstanta. Dengan demikian, diperoleh densitas posterior yang tidak di normalisasi, yang sebanding dengan sisi kanan persamaan (1.2).

Ekspresi sederhana ini merangkum inti teknis dari inferensi Bayesian : tugas utama dalam setiap aplikasi spesifik Adalah mengembangkan model $p(\theta, y)$ serta melakukan komputasi yang diperlukan untuk meringkas distribusi posterior $p(\theta|y)$ dengan cara yang sesuai.(Gelman et al., 1995)

2.2.4 prediksi

Untuk membuat inferensi mengenai suatu besaran yang dapat diamati tetapi belum diketahui, yang sering disebut sebagai inferensi prediktif, kita mengikuti logika yang serupa. Sebelum data y diamati, distribusi dari besaran yang tidak diketahui tetapi dapat diamati, y^* , Adalah

$$p(y) = \int p(y, \theta)d\theta = \int p(\theta)p(y|\theta)d\theta$$

Distribusi ini sering disebut sebagai distribusi marginal dari y , tetapi istilah yang lebih informatif Adalah distribusi prediktif prior (*prior predictive distribution*)

Disebut prior karena distribusi ini tidak dikondisikan pada pengamatan sebelumnya dari proses yang bersangkutan, dan disebut prediktif karena menggambarkan distribusi dari suatu besaran yang dapat diamati.

Setelah data y diamati, kita dapat memprediksi suatu besaran yang dapat diamati tetapi belum diketahui, yaitu y^* , yang berasal dari proses yang sama. Sebagai

contoh misalkan $y = (y_1, \dots, y_n)$ merupakan vector hasil pengukuran berat suatu objek yang ditimbang sebanyak n kali menggunakan sebuah timbangan. Selanjutnya,

$\theta = (\mu, \sigma^2)$ menyatakan berat sebenarnya dari objek tersebut dan varians kesalahan pengukuran dari timbangan, yang keduanya tidak diketahui. Sementara itu, y^* Adalah berat objek yang akan dicatat pada penimbangan baru yang direncanakan.

Distribusi dari y^* yang dikondisikan pada data yang telah diamati y disebut sebagai distribusi prediktif posterior (posterior predictive distribution). Disebut posterior karena distribusi ini bersyarat pada data observasi y , dan disebut prediktif karena memberikan prediksi terhadap suatu besaran yang dapat diamati dimasa depan, yaitu y^* .

$$\begin{aligned} p(y^*|y) &= \int p(y^*, \theta|y) d\theta \\ &= \int p(y^*|\theta, y) p(\theta|y) d\theta \\ &= \int p(y^*|\theta) p(\theta|y) d\theta \end{aligned}$$

Baris kedua dan ketiga menampilkan distribusi prediksi posterior sebagai rata-rata dari prediksi bersyarat atas distribusi posterior θ . Persamaan terakhir diperoleh karena y dan y^* saling independent secara bersyarat berdasarkan θ dalam model ini. (Gelman et al., 1995)

2.2.5 Likelihood

Penggunaan aturan Bayes dengan suatu model probabilitas yang dipilih mengimplikasikan bahwa data y memengaruhi inferensi posterior (1.2) hanya melalui fungsi $p(y|\theta)$. Fungsi ini apabila dipandang sebagai fungsi dari θ dengan y dianggap tetap, disebut sebagai fungsi likelihood. Dengan demikian, inferensi Bayesian mengikuti apa yang sering disebut sebagai prinsip likelihood (likelihood principle), yang menyatakan bahwa suatu sampel data tertentu, dua model probabilitas $p(y|\theta)$ yang memiliki fungsi likelihood yang sama akan menghasilkan inferensi yang sama terhadap parameter θ .

Prinsip likelihood pada dasarnya masuk akal, namun hanya berlaku dalam kerangka model atau keluarga model yang dipilih untuk suatu analisis tertentu. Dalam praktiknya, kita jarang dapat benar-benar yakin bahwa model yang dipilih merupakan model yang benar. Distribusi sampling yaitu dengan membayangkan realisasi data yang berulang dapat memainkan peran penting dalam memeriksa asumsi-asumsi model. Bahkan, pandangan kami tentang seorang statistikus Bayesian terapan Adalah seseorang yang bersedia menerapkan aturan Bayes dibawah berbagai kemungkinan model yang masuk akal.(Gelman et al., 1995)

2.3 Metode Bayesian Ridge Regression

Bayesian Ridge Regression merupakan teknik regresi linear yang memanfaatkan pendekatan Bayesian, Dimana parameter atau koefisien model diperlakukan sebagai variable acak yang memiliki distribusi probabilitas tertentu. Metode ini menggabungkan konsep regulasi ridge (L2) untuk mengurangi risiko overfitting serta mengatasi permasalahan multikolinearitas. Ridge Regression

mampu mengurangi varians estimasi parameter Ketika terjadi multikolinearitas yang tinggi. (Golam Kibria & Lukman, 2020)

Pendekatan algoritma Bayesian Ridge Regression dalam penelitian ini mencakup beberapa komponen penting. Komponen pertama Adalah Data Observasi, yaitu data historis yang digunakan sebagai dasar pelatihan model, seperti curah hujan dan variable cuaca pendukung lainnya. Selanjutnya, digunakan Model Regresi Linear untuk mempresentasikan hubungan antara variable input dan variable output dalam bentuk persamaan linear yang disertai error.

Metode ini dilakukan dengan menghitung Mean Squared Error (MSE) sebagai ukuran kemampuan prediksi untuk setiap model. Model yang memiliki nilai MSE paling rendah dianggap sebagai model yang paling baik dalam melakukan prediksi.

Nilai MSE dapat dihitung sebagai berikut :

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - y_i^*)^2$$

Dimana n adalah jumlah titik data yang diprediksi, y_i merupakan nilai respons yang diketahui untuk suatu Kumpulan pengamatan yang sama. MSE menghitung rata-rata dari kuadrat selisih antar nilai respons sebenarnya dan nilai respons hasil prediksi. Semakin kecil nilai MSE semakin dekat nilai prediksi dengan nilai sebenarnya, sehingga semakin baik kecocokan (fit) model tersebut.

2.3.1 Keunggulan Bayesian Ridge Regression

1. Menangani multikolinearitas

Sangat sesuai digunakan pada kondisi Dimana variabel predictor memiliki kolerasi yang tinggi satu sama lain.

2. Mengintegrasikan informasi Prior

Memberikan fleksibilitas untuk memasukkan pengetahuan awal (*prior knowledge*) mengenai parameter-parameter model.

3. Efek regularisasasi

Mencegah terjadinya overfitting melalui mekanisme regularisasi yang secara alami muncul dalam pendekatan Bayesian.

4. Distribusi probabilitas

Menghasilkan estimasi ketidak pastina pada koefisien model. Yang bermanfaat dalam pembuatan prediksi yang bersifat probabilistik.

2.4 PENELITIAN TERDAHULU

Tabel 1. 1 Peneliti Terdahulu

no	Judul dan peneliti	Pembahasan	Metode	Kelebihan dan Kekurangan
1.	Precipitation Prediction in Several Chinese Regions Using machine	Prediksi curah hujan di beberapa wilayah Tiongkok dan membandingkan	Bayesian ridge regression (BRR), SVR, Random	Kelebihan : BRR menjadi baseline linear Bayesian yang stabil saat prediktor berkolerasi

	learning methods (Wang et al., 2024)	performa beberapa metode ML.	Forest, Linear Regression.	(multikolinearitas) dan cocok untuk baseline pembanding. Kekurangan : umumnya performa non-linear (mis. RF) lebih unggul, detail horizon?lead time “jangka pendek”.
2.	Using Bayesian Ridge Regression model and ESN for Climatic time series forecasting (Snaan & Shukur, 2025)	Forecasting deret waktu iklim ; menekankan isu multikolinearitas dan <i>hybrid model</i> untuk meningkatkan akurasi.	Model hybrid BRR-ESN (Bayesian Ridge Regression + Echo State Network).	Kelebihan : Menunjukkan BRR bisa dipakai sebagai komponen regularisasi + probabilistic (berguna untuk kestabilan prediksi) dan ditingkatkan dengan model non—linear ESN. Kekurangan : Fokusnya climatic time series (di

				abstrak meyinggung evaporation/variable iklim), jadi tidak “curah hujan jangka pendek” secara spesifik-tetap relevan sebagai rujuk metode BRR di domain iklim.
3	Prediksi Curah Hujan Menggunakan Metode Bayesian (Muharsyah et al., 2021)	Penerapan Bayesian Ridge Regression dalam memprediksi curah hujan menggunakan data meteorologi dari stasiun cuaca.	Bayesian Ridge Regression	Kelebihan : Bayesian Ridge Regression mampu menghasilkan prediksi curah hujan yang lebih stabil dan akurat karena robust terhadap multikolinearitas. Kekurangan : masih terbatas pada pendekatan linier, sehingga performa model berpotensi

				kurang optimal jika pola hujan dipengaruhi oleh hubungan nonlinier yang kompleks antar variabel.
4	Linearized Ridge	Menunjukkan	Linearized	Kelebihan : Model
	Regression Modelling With MM-Estimator in Statistical Downscaling For Rainfall Forecasting (SISWANTO et al., 2025)	bahwa Linear Ridge Regression dengan MM-Estimator (LRMM) mampu meningkatkan akurasi peramalan curah hujan melalui pendekatan <i>statistical downscaling</i>,	Ridge Regression (MM-Estimator)	LRMM efektif menangani multikolinearitas dan outlier secara bersamaan, sehingga menghasilkan prediksi curah hujan yang lebih stabil dan akurat dibandingkan model Ridge tanpa estimator robust. Kekurangan : Model masih berbasis regresi

		dengan hasil yang kuat terhadap multikolinearitas dan outlier. dengan baik		linear, sehingga berpotensi kurang optimal dalam menangkap pola curah hujan yang sangat nonlinier atau kejadian ekstrem yang kompleks
5	Koomparasi Ridge Regression, Random Forest, dan Gradient Boosting Untuk Prediksi Curah Hujan Harian Di Sumatera Selatan Berbasis Time Series Cross-Validation	Menunjukkan bahwa Ridge Regression memberikan hasil prediksi terbaik pada data uji dengan nilai RMSE 14,62 mm, namun nilai R^2 yang negative menandakan	Ridge Regression, Random Forest, dan Gradient Boosting	Kelebihan : menggunakan time series cross-validation dan membandingkan beberapa model, sehingga evaluasi performa prediksi menjadi lebih realistis dan komprehensif. Kekurangan : Model yang

	(Alvines et al., 2025)	model masih belum mampu menangkap variabilitas curah hujan secara menyeluruh, terutama pada kejadian hujan ekstrem.		digunakan masih kurang akurat dalam memprediksi curah hujan ekstrem, dipengaruhi oleh distribusi data yang tidak seimbang dan keterbatasan variabel yang digunakan.
--	------------------------	---	--	---

BAB III

METODOLOGI PENELITIAN

3.1 Jenis Dan Pendekatan Penelitian

Penelitian ini merupakan penelitian kuantitatif dengan pendekatan pembelajaran mesin (machine learning) berbasis pemodelan statistic. Pendekatan kuantitatif digunakan karena penelitian ini mengolah data numerik berupa data curah hujan dan variabel meteorologi pendukung yang dianalisis secara matematis dan statistic untuk membangun model prediksi.

Metode machine learning yang digunakan dalam penelitian ini Adalah Bayesian Ridge Regression, yaitu metode pembelajaran terawasi (*supervised learning*) berbasis regresi linear yang menerapkan pendekatan Bayesian dalam proses estimasi parameter. Metode ini memadukan Teknik Regularisasi Ridge (*L2 regularization*) dengan inferensi probabilistik, sehingga mampu mengatasi permasalahan multikolinearitas antar variabel prediktor serta mengurangi risiko terjadinya overfitting. Oleh karena itu, *Bayesian Ridge Regression* dinilai mampu sesuai untuk memodelkan data curah hujan yang memiliki Tingkat fluktuasi dan ketidakpastian yang tinggi.

Pendekatan Bayesian dalam machine learning memungkinkan proses estimasi parameter dilakukan secara probabilistic sehingga dapat memberikan informasi ketidakpastian prediksi dibandingkan Regresi Linear Konvensional. (Assaf et al., 2019). Pendekatan Machine Learning diterapkan melalui proses pelatihan model menggunakan data historis sebagai data latih untuk mempelajari pola hubungan antara variabel masukan dan variabel keluaran. Model yang telah berbentuk selanjutnya digunakan untuk melakukan prediksi curah hujan jangka

pendek. Dalam penelitian ini, prediksi jangka pendek diartikan sebagai prediksi curah hujan untuk satu periode waktu kedepan, seperti prediksi harian hingga mingguan, sesuai dengan resolusi data yang tersedia.

Berdasarkan tujuan penelitian, penelitian ini bersifat eksperimental, yaitu dilakukan melalui serangkaian percobaan pemodelan untuk memperoleh model prediksi yang optimal. Data penelitian dibagi menjadi data latih dan data uji, kemudian kinerja model dievaluasi menggunakan metrik evaluasi tertentu, seperti Mean Squared Error (MSE), guna menilai Tingkat akurasi dan keandalan metode Bayesian Ridge Regression dalam memprediksi curah hujan jangka pendek.

Dengan demikian, pendekatan yang digunakan dalam penelitian ini diharapkan mampu memberikan Gambaran yang sistematis dan objektif mengenai penerapan metode machine learning berbasis *Bayesian Ridge Regression* serta kontribusinya dalam mendukung prediksi curah hujan jangka pendek secara ilmiah dan aplikatif.

3.2 Data Penelitian

3.2.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari Kaggle yang menyediakan data meteorologi dan klimatologi atau sumber data cuaca lain yang relevan. Data tersebut dikumpulkan dari stasiun pengamatan cuaca pada wilayah penelitian dalam rentang waktu tertentu sesuai dengan kebutuhan penelitian.

3.3. Variabel Penelitian

Variabel penelitian merupakan unsur yang diamati dan di analisis dalam penelitian serta memiliki variasi nilai yang dapat diukur. Dalam penelitian ini, variabel penelitian digunakan sebagai komponen utama dalam proses pemodelan machine learning untuk memprediksi curah hujan jangka pendek menggunakan metode *Bayesian Ridge Regression*. Variabel penelitian yang digunakan terdiri atas variabel terikat dan variabel bebas.

1. Variabel Terikat

Variabel terikat (*dependen*) dalam penelitian ini adalah curah hujan. Curah hujan merupakan jumlah air yang hujan yang jatuh pada suatu wilayah dalam periode waktu tertentu dan dinyatakan dalam satuan milimeter (mm). variabel ini berperan sebagai variabel target (output) yang akan diprediksi oleh model *Bayesian Ridge Regression*.

2. Variabel Bebas

Variabel bebas (*independent*) dalam penelitian ini merupakan variable-variabel yang digunakan sebagai masukan (input) dalam proses pemodelan machine learning.

Variabel bebas yang digunakan meliputi :

1. Curah hujan periode sebelumnya (*lag curah hujan*)

Variabel ini digunakan untuk menangkap pola historis dan ketergantungan waktu pada data deret waktu curah hujan.

2. Suhu udara

Variabel suhu udara menunjukkan Tingkat panas udara yang memengaruhi proses penguapan dan pembentukan awan hujan.

3. Kelembapan udara

Variabel kelembapan udara menunjukkan kandungan uap air di atmosfer yang berperan dalam proses terjadinya hujan.

4. Tekanan udara

Variabel tekanan udara menggambarkan kondisi tekanan atmosfer yang memengaruhi pergerakan massa udara dan dinamika cuaca.

5. Arah angin

Variabel arah angin menunjukkan asal pergerakan massa udara dan berpengaruh terhadap distribusi uap di atmosfer.

Pemilihan variabel bebas tersebut didasarkan pada keterkaitannya dengan proses meteorologi yang memengaruhi terjadinya hujan serta ketersediaan data pendukung dalam penelitian.

3.3.1 Jenis Data

Jenis data yang digunakan dalam penelitian ini berupa data deret waktu (time series) yang disusun berdasarkan urutan waktu pengamatan. Data yang digunakan terdiri dari :

1. Data Curah Hujan

Data curah hujan berfungsi sebagai variabel target (output) yang akan diprediksi. Data ini dinyatakan dalam satuan milimeter (mm) dan dicatat secara berkala, seperti harian atau mingguan.

2. variabel meteorologi pendukung ini bertujuan untuk meningkatkan kemampuan model dalam menangkap pola dan dinamika perubahan curah hujan.

3.3.2 Pembagian Data

Untuk keperluan pemodelan machine learning, data dibagi menjadi dua bagian, yaitu :

- a. Data latih (training data), yang digunakan untuk melatih model Bayesian Ridge Regression.
- b. Data Uji (testing data), yang digunakan untuk menguji kinerja model terhadap data yang belum pernah dilihat sebelumnya.

Pembagian data yang dilakukan secara berurutan berdasarkan waktu pengamatan untuk menjaga karakteristik data deret waktu dan menghindari kebocoran data (data leakage).

3.4 Prosedur Penelitian

Prosedur penelitian merupakan tahapan-tahapan yang dilakukan secara sistematis untuk mencapai tujuan penelitian. Dalam penelitian ini, prosedur penelitian disusun berdasarkan tahapan kerja metode machine learning dengan penerapan *Bayesian Ridge Regression* untuk prediksi curah hujan jangka pendek. Adapun tahapan prosedur penelitian yang dilakukan adalah sebagai berikut.

3.4.1 Identifikasi Masalah

Tahap awal penelitian ini diawali dengan perumusan dan identifikasi permasalahan penelitian, yaitu perlunya suatu metode yang mampu menghasilkan prediksi curah hujan jangka pendek secara akurat. Permasalahan yang dihadapi berkaitan dengan keterbatasan metode regresi konvensional dalam menangani

multikolinearitas antar variabel prediktor serta ketidakstabilan parameter pada data curah hujan yang memiliki Tingkat fluktuasi tinggi.

3.4.2 Kajian Literatur

Pada tahap ini dilakukan penelusuran dan kajian terhadap berbagai sumber Pustaka yang relevan, meliputi konsep Bayesian, penerapan machine learning, metode *Bayesian Ridge Regression*, serta hasil penelitian sebelumnya yang berkaitan dengan prediksi curah hujan. Kajian literatur bertujuan untuk memperkuat landasan teori serta menentukan pendekatan metodologis yang tepat dalam penelitian.

3.4.3 Pengumpulan Data

Tahap pengumpulan data dilakukan dengan menghimpun data penelitian yang berupa data historis curah hujan dan variabel meteorologi pendukung. Data diperoleh dari Kaggle yang menyediakan data meteorologi dan klimatologi, dan disesuaikan dengan periode serta resolusi waktu yang diperlukan untuk mendukung prediksi curah hujan jangka pendek.

3.4.4 Pengolahan Awal Data

Data yang telah dikumpulkan selanjutnya diproses melalui tahap pengolahan awal (*preprocessing*). tahap ini meliputi pembersihan data dari nilai hilang atau data yang tidak sesuai, proses normalisasi atau standarisasi data untuk menyamakan skala variabel, serta pembentukan data deret waktu (*time series*) dengan memanfaatkan nilai pengamatan pada periode sebelumnya (*lag data*).

Standardisasi data menjadi tahapan penting dalam algoritma berbasis regularisasi karena membantu proses konvergensi model dan mencegah dominasi variabel dengan skala lebih besar terhadap proses pelatihan model. (Samadikoucheksaraee & Chu, 2024)

3.4.5 Pembagian Data Penelitian

Data yang telah melalui tahap pengolahan selanjutnya dibagi menjadi data latih dan data uji. Pembagian data dilakukan berdasarkan urutan waktu pengamatan guna mempertahankan karakteristik data deret waktu serta menghindari terjadinya kebocoran data (*data leakage*).

3.4.6 Proses Pelatihan Model

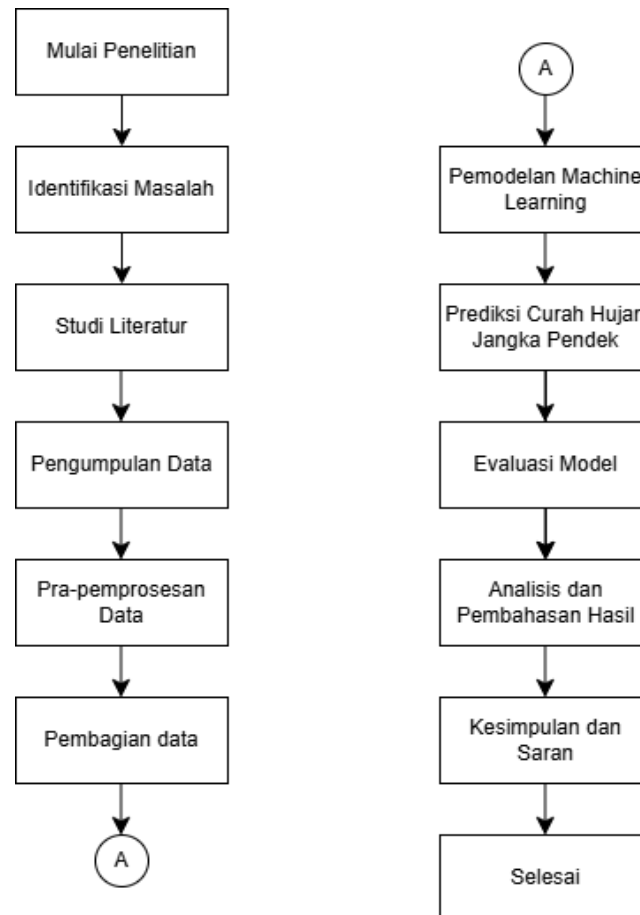
Evaluasi model dilakukan untuk menilai kemampuan model dalam menghasilkan prediksi yang akurat. Kinerja model dievaluasi dengan membandingkan hasil prediksi terhadap data aktual menggunakan metrik evaluasi *Mean Squared Error (MSE)*. Semakin kecil nilai MSE, semakin baik kinerja model dalam memprediksi curah hujan.

3.5 Analisis Hasil Dan Kesimpulan

Tahap akhir penelitian meliputi analisis hasil prediksi dan evaluasi model. Hasil analisis digunakan untuk menarik Kesimpulan mengenai tingkat efektivitas metode *Bayesian Ridge Regression* dalam memprediksi curah hujan jangka pendek, serta untuk merumuskan saran bagi pengembangan penelitian selanjutnya.

3.6 Tahapan Penelitian

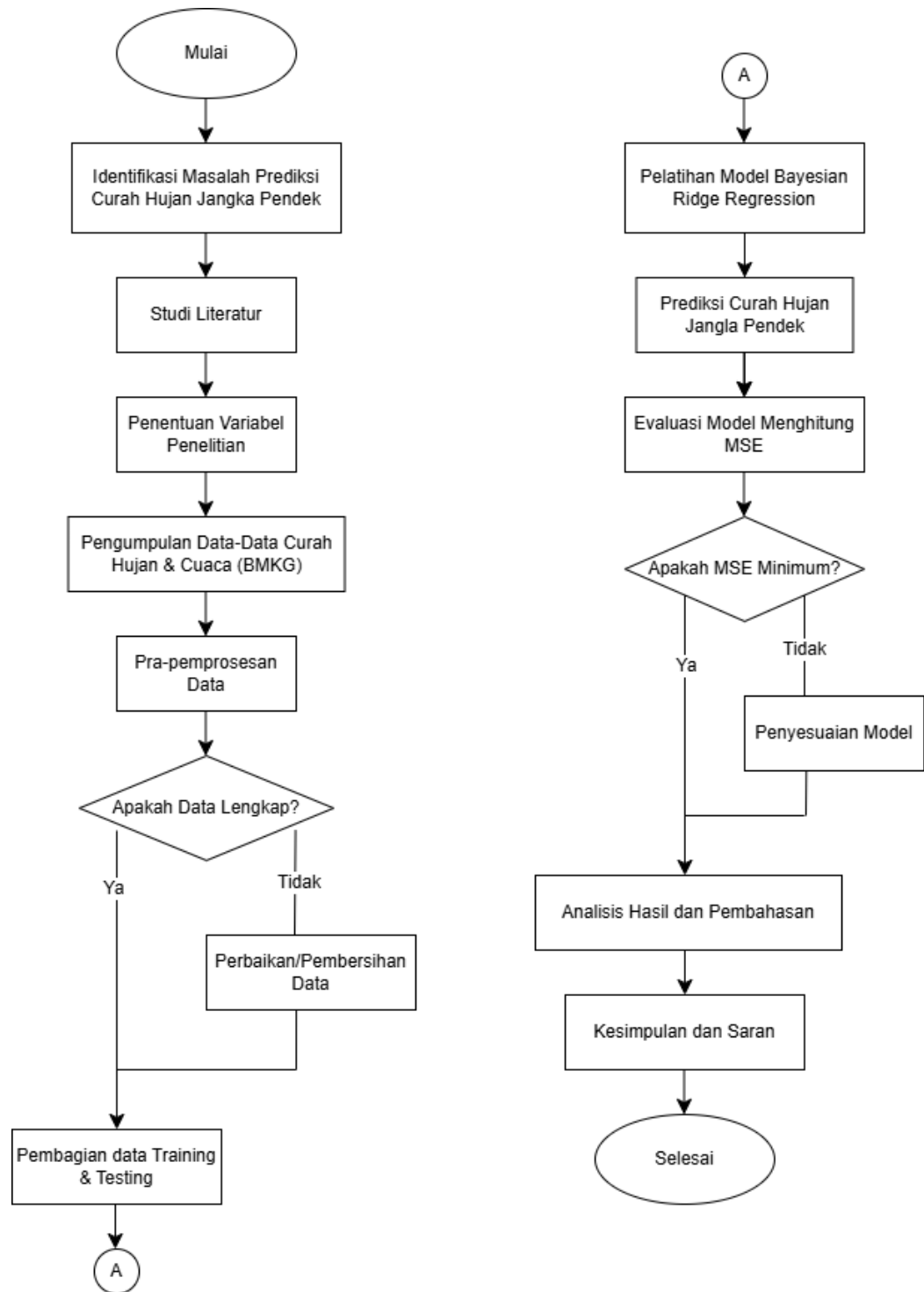
3.6.1 Diagram Alur Penelitian



Gambar 3. 1 Diagram Alur Penelitian

Gambar ini merupakan diagram alur penelitian prediksi curah hujan jangka pendek menggunakan metode *Bayesian Ridge Regression*. Proses dimulai dengan identifikasi masalah dan studi literatur untuk menentukan fokus serta dasar teori penelitian. Selanjutnya dilakukan pengumpulan data yang kemudian dipra-proses agar siap digunakan. Data yang telah bersih dibagi menjadi data latih dan data uji, kemudian digunakan dalam tahap pemodelan machine learning untuk melakukan prediksi curah hujan jangka pendek. Hasil prediksi dievaluasi untuk menilai kinerja model, dilanjutkan dengan analisis dan pembahasan hasil. Tahap akhir penelitian adalah penarikan kesimpulan dan pemberian saran sebelum penelitian dinyatakan selesai.

3.6.2 Flowchart Bayesian Ridge Regression



Gambar 3. 2 Flowchart Bayesian Ridge Regression

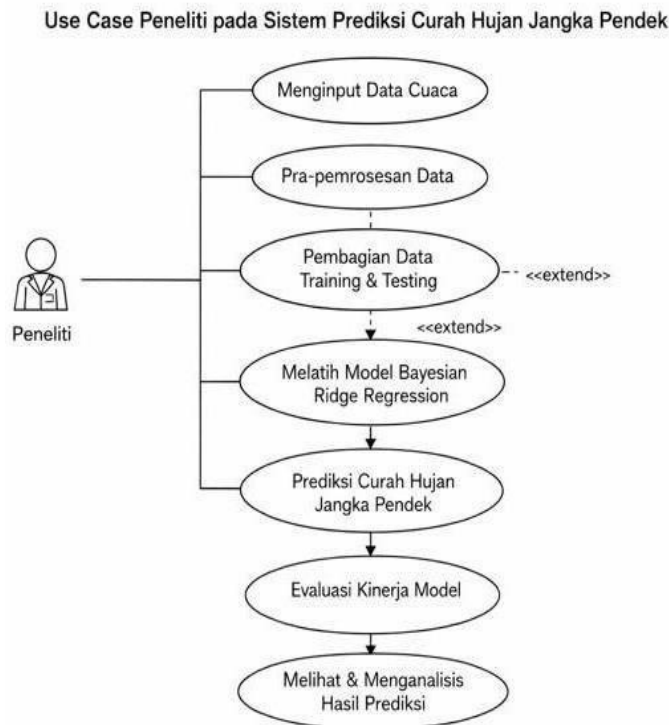
Gambar tersebut merupakan diagram alur penelitian prediksi curah hujan jangka pendek menggunakan metode Bayesian Ridge Regression. Proses dimulai

dari identifikasi masalah dan studi literatur, dilanjutkan dengan penentuan variabel serta pengumpulan data curah hujan dan cuaca. Data kemudian melalui tahap pra-pemrosesan dan pengecekan kelengkapan, dimana data yang belum lengkap akan diperbaiki sebelum dibagi menjadi data training dan testing. Selanjutnya model *Bayesian Ridge Regression* dilatih untuk melakukan prediksi curah hujan jangka pendek. Hasil prediksi di evaluasi menggunakan nilai *Mean Squared Error (MSE)*. Jika nilai MSE belum minimum, dilakukan penyesuaian model, sedangkan jika sudah memenuhi kriteria, proses dilanjutkan ketahap analisis hasil dan pembahasan. Alur penelitian diakhiri dengan penarikan kesimpulan dan pemberian saran.

3.6.3 Usecase

Usecase diagram menggambarkan interaksi antara aktor dengan sistem prediksi curah hujan jangka pendek yang dikembangkan dalam penelitian ini. Aktor utama dalam sistem ini adalah peneliti, yang berperan dalam menjalankan seluruh tahapan pemodelan machine learning mulai dari pengolahan data hingga evaluasi hasil prediksi.

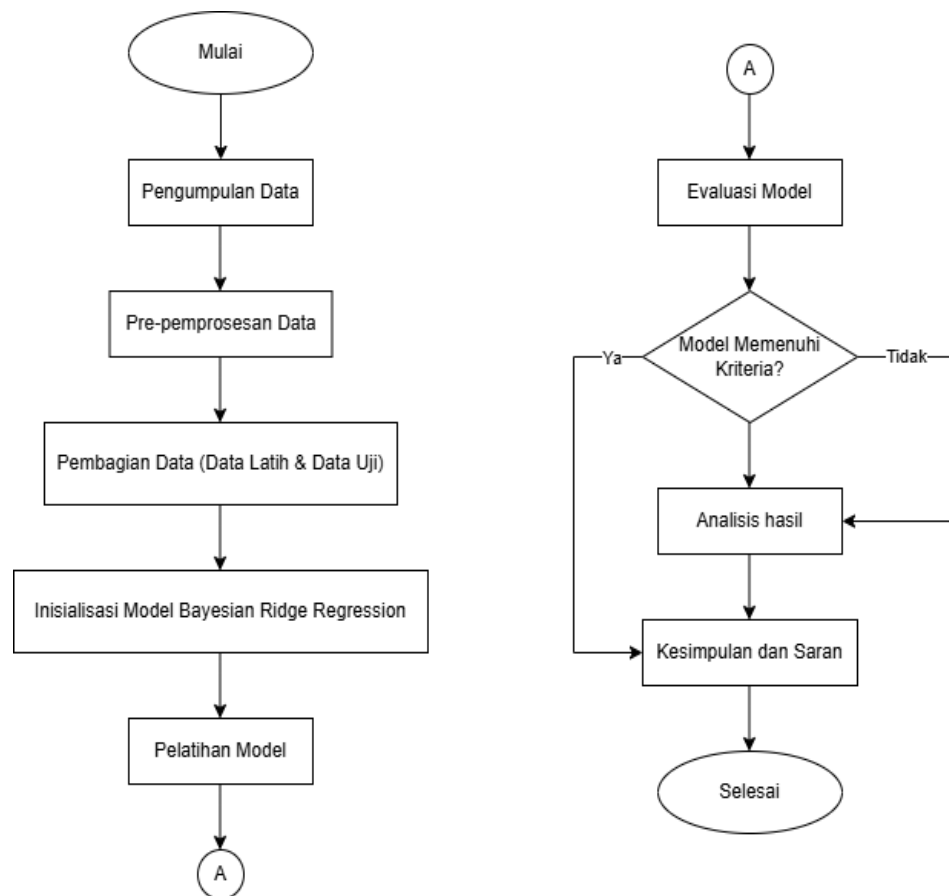
1. Usecase Aktor



Gambar 3. 3 Usecase Aktor

Gambar tersebut merupakan Usecase diagram Peneliti pada sistem prediksi curah hujan jangka pendek. Aktor utama dalam diagram ini Adalah peneliti, yang berinteraksi langsung dengan sistem untuk menjalankan seluruh proses penelitian. Proses dimulai dari menginput data cuaca kedalam sistem, kemudian dilakukan pra-pemrosesan data agar data siap digunakan. Selanjutnya, data dibagi menjadi data training dan data testing sebagai dasar pemodelan. Data training digunakan untuk melatih model *Bayesian Ridge Regression*, yang kemudian menghasilkan prediksi curah hujan jangka pendek. Setelah prediksi diperoleh, dilakukan evaluasi kinerja model untuk menilai tingkat keakuratan. Tahap akhir adalah melihat dan menganalisis hasil prediksi sebagai bahan pembahasan dan pengambilan Kesimpulan.

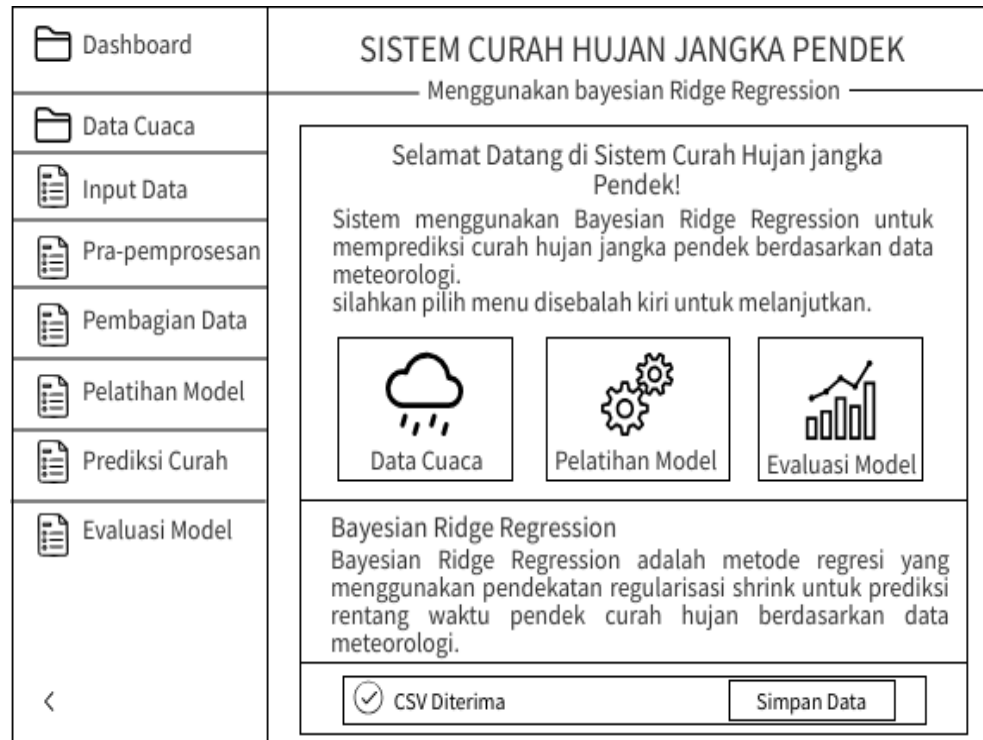
3.6.4 Diagram Activity



Gambar 3. 4 Diagram Activity

Diagram activity Bayesian Ridge Regression menggambarkan alur penelitian yang dimulai dari pengumpulan data hingga penarikan Kesimpulan. Data yang telah dikumpulkan terlebih dahulu diproses melalui tahap pra-pemrosesan agar siap digunakan, kemudian dibagi menjadi data latih dan data uji. Selanjutnya, model Bayesian Ridge Regression diinisialisasi dan dilatih menggunakan data latih untuk membangun model prediksi. Model yang telah dilatih kemudian di evaluasi menggunakan data uji untuk menilai performanya. Apabila hasil evaluasi belum memenuhi kriteria yang ditentukan, maka proses pelatihan diulang. Jika model telah memenuhi kriteria, dilakukan analisis hasil untuk kemudian ditarik Kesimpulan dan saran sebagai akhir dari penelitian.

3.6.5 Rancangan UI/UX (User Interface/User Experience)



Gambar 3. 5 Tampilan Dashboard Sistem Prediksi Curah Hujan Jangka Pendek

Gambar tersebut menampilkan antarmuka dashboard sistem prediksi curah hujan jangka pendek yang menggunakan metode Bayesian Ridge Regression. Pada sisi kiri terdapat menu navigasi seperti data cuaca, input data, pra-pemrosesan, pembagian data, pelatihan model, prediksi, dan evaluasi model. Bagian utama dashboard berisi informasi pengantar sistem, ikon fitur utama (data cuaca, pelatihan model, dan evaluasi model), serta penjelasan singkat mengenai metode Bayesian Ridge Regression. Di bagian bawah juga ditampilkan status unggah data (CSV diterima) dan tombol untuk menyimpan data.

Dashboard	Input Data Curah Hujan Peneliti																							
Data Cuaca	Unggah data curah hujan dan variabel cuaca pendukung dalam format CSV/Excel. Contoh cuplikan data BMKG Unggah Cuaca Unggah Cuaca <table border="1"> <thead> <tr> <th>Tanggal</th> <th>Curah Hujan (mm)</th> <th>Suhu (°C) (° C)</th> <th>Kelembapan (%) (%)</th> </tr> </thead> <tbody> <tr> <td>2020-12-08</td> <td>13</td> <td>28,6</td> <td>78</td> </tr> <tr> <td>2020-14-08</td> <td>1</td> <td>28,6</td> <td>78</td> </tr> <tr> <td>2021-21-01</td> <td>10,5</td> <td>28,6</td> <td>79</td> </tr> <tr> <td>2021-22-02</td> <td>13,4</td> <td>27,7</td> <td>85</td> </tr> </tbody> </table> ✓ CSV Diterima Simpan Data ✓ Pra-pemrosesan Selesai !				Tanggal	Curah Hujan (mm)	Suhu (°C) (° C)	Kelembapan (%) (%)	2020-12-08	13	28,6	78	2020-14-08	1	28,6	78	2021-21-01	10,5	28,6	79	2021-22-02	13,4	27,7	85
Tanggal					Curah Hujan (mm)	Suhu (°C) (° C)	Kelembapan (%) (%)																	
2020-12-08					13	28,6	78																	
2020-14-08					1	28,6	78																	
2021-21-01					10,5	28,6	79																	
2021-22-02					13,4	27,7	85																	
Input Data																								
Pra-pemrosesan																								
Pembagian Data																								
Pelatihan Model																								
Prediksi Curah																								
Evaluasi Model																								
<																								

Gambar 3. 6 Tampilan Input Data Curah Hujan Jangka Pendek

Gambar tersebut menampilkan halaman input data curah hujan pada sistem prediksi curah hujan beserta variabel cuaca pendukung dalam format CSV. Data yang diunggah ditampilkan dalam bentuk tabel yang berisi tanggal, curah hujan, suhu, dan kelembapan. Dibagian bawah terdapat informasi status bahwa file CSV telah diterima dan proses pra-pemrosesan telah selesai, serta tombol untuk menyimpan data sebelum melanjutkan ketahap berikutnya.

Dashboard	Pra-pemrosesan Data Peneliti
Data Cuaca	Unggah data curah hujan dan variabel cuaca pendukung dalam format CSV/Excel. Contoh cuplikan data BMKG Unggah Cuaca ☒ Clean Data ☒ Normalisasi ☒ Pembentukan Lag Data Proses Data ☒ Memproses data selesai !
Input Data	
Pra-pemrosesan	
Pembagian Data	
Pelatihan Model	
Prediksi Curah	
Evaluasi Model	
<	

Gambar 3. 7 Tampilan Pra-pemrosesan Data

Gambar tersebut menampilkan halaman pra-pemrosesan data pada sistem prediksi curah hujan jangka pendek. Halaman ini digunakan untuk mengunggah data cuaca dalam format CSV atau Excel dan melakukan tahapan pra-pemrosesan, seperti pembersihan data (*clean data*), normalisasi, serta pembentukan lag data. Setelah proses dijalankan melalui tombol proses data, sistem menampilkan notifikasi bahwa pemrosesan data telah selesai.

Dashboard

Data Cuaca

Input Data

Pra-pemrosesan

Pembagian Data

Pelatihan Model

Prediksi Curah

Evaluasi Model

Pembagian Data

Peneliti

Pisahkan data menjadi data set training untuk melatih model dan set testing untuk menguji model

0%

0%

50%

100%

Data Training
80% (Total Data 500 Baris)
400 Baris

Data Testing
20% (Total Data 500 Baris)
100 Baris

Tanggal	Curah Hujan (mm)	Kelembapan (%)	
2020-12-08	13	28,6	78
2020-14-08	1	28,6	78
2021-21-01	10,5	28,6	79
2021-22-02	13,4	28,6	78

Tanggal	Curah Hujan (mm)	Kelembapan (%)	
2020-12-08	13	78	79
2020-14-08	1	78	81
2021-21-01	27,1	78	81
2021-22-02	27,1	78	81

80% Data Training : 400 Baris

20% Data Testing : 100 Baris

☒ CSV Diterima

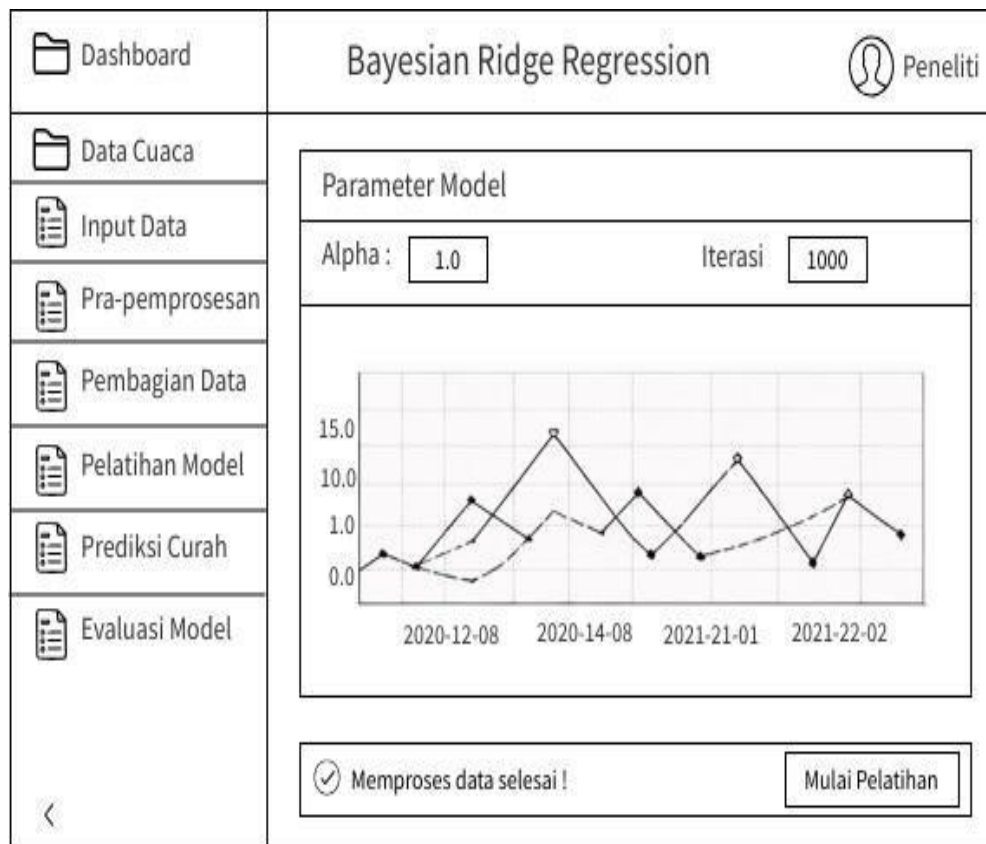
☒ Pra-pemrosesan Selesai !

☒ Data berhasil dibagi menjadi set Training sebanyak 400 baris dan set

Lanjutkan

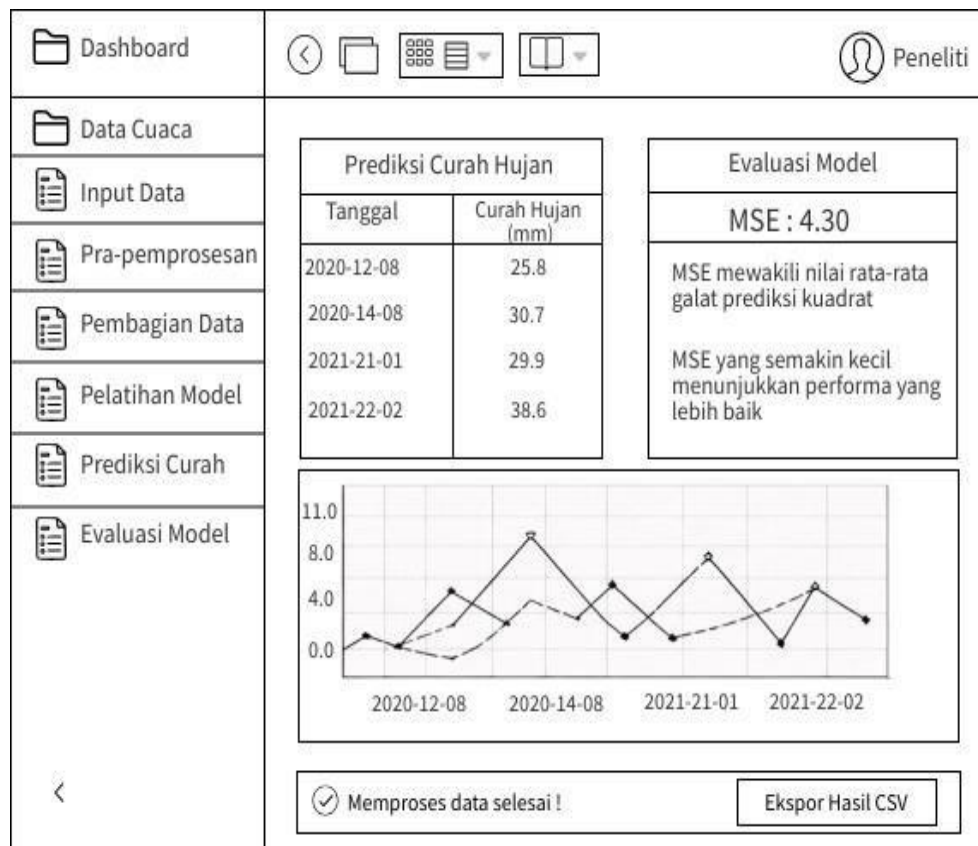
Gambar 3. 8 Tampilan Pembagian Data

Gambar tersebut menampilkan halaman pembagian data pada sistem prediksi curah hujan jangka pendek. Halaman ini berfungsi untuk memisahkan data hasil pra-pemrosesan menjadi data training dan data testing dengan proporsi tertentu, pada contoh ditunjukkan pembagian 80% data training dan 20% data testing. Data yang telah dibagi ditampilkan dalam bentuk tabel untuk masing-masing set, lengkap dengan informasi jumlah baris data. Dibagian bawah terdapat notifikasi bahwa data berhasil dibagi serta tombol untuk melanjutkan ketahap berikutnya.



Gambar 3. 9 Tampilan Pelatihan Model Bayesian Ridge Regression

Gaambar tersebut menampilkan halaman pelatihan model Bayesian Ridge Regression. Halaman ini digunakan untuk mengatur parameter model, seperti nilai alpha dan jumlah iterasi, sebelum proses pelatihan dilakukan. Selain itu, ditampilkan grafik sebagai visualisasi data curah hujan dan hasil pemodelan untuk membantu peneliti memahami pola data. Setelah data siap, pengguna dapat memulai proses pelatihan model melalui tombol yang tersedia, dan sistem akan memberikan notifikasi ketika proses pelatihn selesai.



Gambar 3. 10 Tampilan Hasil Sistem Prediksi Curah Hujan Jangka Pendek

Gambar tersebut menampilkan halaman prediksi dan evaluasi model curah hujan. Halaman ini menyajikan hasil prediksi curah hujan dalam bentuk tabel berdasarkan tanggal tertentu, serta hasil evaluasi model menggunakan nilai Mean Squared Error (MSE). Selain itu, ditampilkan grafik perbandingan untuk membantu memahami hasil prediksi. Pengguna juga dapat mengekspor hasil prediksi dalam format CSV setelah proses selesai.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Dataset Penelitian

Dataset diambil langsung dari kaggle yang berjumlah 516 dataset. Kemudian untuk memvariasi jumlah dataset dilakukan augmentasi data dimana total dataset mencapai 1000 dataset. Dataset ini berisi informasi terkait kondisi cuaca harian yang meliputi variabel curah hujan sebagai variabel target serta beberapa variabel pendukung seperti suhu udara, kelembapan udara, tekanan udara, dan variabel meteorologi lainnya.

Table Tabel 4. 1Representasi Label

Kelas	Rentang Curah Hujan	Keterangan
Rendah	< 5 mm	Tidak hujan / hujan ringan
Sedang	5 – 20 mm	Hujan dengan intensitas sedang
Lebat	20 – 50 mm	Hujan dengan intensitas tinggi

Dalam penelitian ini digunakan beberapa parameter cuaca sebagai fitur input dalam proses prediksi curah hujan. Fitur-fitur tersebut dipilih berdasarkan pengaruhnya terhadap kondisi atmosfer. Adapun fitur yang digunakan dapat dilihat pada tabel berikut.:

Tabel 4. 2 Fitur yang digunakan

Nama Fitur	Keterangan
Kelembaban Jam 7	Presentase kelembaban udara
Suhu minimum	Suhu terenda harian
Curah Hujan Kemarin	Data hujan hari sebelumnya
Curah Hujan 3 hari lalu	Data historis hujan
Kecepatan angin	Rata-rata kecepatan angin
Tekanan udara	Tekanan atmosfer
Suhu rata-rata 24 jam	Rata-rata suhu harian

4.2 Preprocessing Data

Tahap pra-pemrosesan data merupakan langkah krusial untuk memastikan bahwa data meteorologi yang digunakan memiliki kualitas yang baik dan siap diproses oleh algoritma *Bayesian Ridge Regression*. Mengingat data cuaca sering kali memiliki nilai yang tidak lengkap atau format yang tidak seragam, dilakukan beberapa tahapan sebagai berikut:

1. Augmentasi data = Augmentasi data merupakan proses penambahan jumlah data latih dengan cara menghasilkan data sintetis berdasarkan pola dari data asli. Setelah melakukan augmentasi data bertambah sebanyak 1000 dataset
2. Pembersihan data : Pada tahap awal, dilakukan pemeriksaan terhadap dataset asli yang berjumlah 516 entri data. Proses pembersihan difokuskan pada penanganan nilai yang hilang (*missing values*) yang

ditandai dengan simbol strip ("-") atau nilai kosong (NaN) pada variabel target, yaitu curah hujan. Baris data yang tidak memiliki informasi curah hujan dieliminasi untuk menghindari bias pada saat pelatihan model.

3. Rekayasa Fitur : Untuk meningkatkan akurasi prediksi, dilakukan pembentukan fitur baru berupa variabel *lagging* (tunda), yaitu *rr_lag1* (data curah hujan H-1) dan *rr_lag3* (data curah hujan H-3). Penggunaan fitur *lagging* ini bertujuan agar model dapat mempelajari pola historis curah hujan dari hari-hari sebelumnya. Akibat dari penggunaan teknik ini, diperlukan kesinambungan urutan tanggal. Baris data yang memiliki celah waktu (*time gap*) atau tidak memiliki referensi data historis yang lengkap otomatis dihapus oleh sistem.
4. Scaling data : Mengingat setiap fitur memiliki rentang nilai yang berbeda (misalnya suhu dalam derajat Celsius dan tekanan udara dalam milibar), dilakukan proses standardisasi menggunakan *StandardScaler*. Hal ini dilakukan agar seluruh fitur memiliki skala yang seragam dengan nilai rata-rata (*mean*) nol dan standar deviasi satu, sehingga mempermudah algoritma *Bayesian Ridge* dalam mencapai konvergensi.

Setelah melalui seluruh rangkaian tahapan di atas, dataset yang siap digunakan untuk tahap pemodelan adalah sebanyak 1000 entri data. Data inilah yang kemudian akan dibagi menjadi data latih (*training*) dan data uji (*testing*) untuk tahap evaluasi sistem. Langkah selanjutnya adalah membagi dataset tersebut menjadi dua bagian utama, yaitu data latih (*training data*) dan data uji (*testing*

data). Pembagian ini dilakukan dengan rasio 80:20, di mana 80% data digunakan untuk melatih algoritma Bayesian Ridge Regression agar dapat mengenali pola cuaca, sedangkan 20% sisanya digunakan untuk menguji performa model dalam memprediksi data yang belum pernah dilihat sebelumnya.

Adapun rincian jumlah pembagian dataset hasil pra-pemrosesan adalah sebagai berikut:

Tabel 4. 3 Pembagian Dataset

Jenis Dataset	Presentase	Jumlah Sampel
Data Latih (Training)	80%	800
Data Uji (Testing)	20%	1000
Total	100%	1000

4.3 Implementasi Bayesian Ridge Regresion

Tahap implementasi sistem dilakukan dengan mengintegrasikan model prediktif yang telah dilatih menggunakan algoritma *Bayesian Ridge Regression* ke dalam sebuah antarmuka aplikasi. Proses ini dimulai dengan memuat (*loading*) file model (*model_final.pkl*) dan skaler data (*scaler_final.pkl*) yang telah disimpan sebelumnya melalui pustaka *Joblib*. Implementasi kode dilakukan menggunakan bahasa pemrograman Python dengan memanfaatkan pustaka *Pandas* untuk manipulasi data dan *Scikit-Learn* sebagai mesin komputasi utama untuk memproses parameter input meteorologi secara *real-time*.

Antarmuka sistem dirancang untuk menerima input dari pengguna berupa variabel cuaca yang telah ditentukan, kemudian melakukan normalisasi data sebelum menjalankan fungsi prediksi. Hasil keluaran sistem berupa nilai numerik

curah hujan dalam satuan milimeter (mm) yang mencerminkan intensitas presipitasi untuk periode waktu tertentu. Implementasi ini juga mencakup mekanisme *post-processing* untuk memastikan nilai prediksi tidak bernilai negatif, sehingga informasi yang ditampilkan kepada pengguna tetap logis dan sesuai dengan standar pengukuran meteorologi.

4.4 Pengujian Model

Tahap pengujian dilakukan untuk mengukur sejauh mana model *Bayesian Ridge Regression* mampu memprediksi nilai curah hujan dengan akurat berdasarkan data uji (*testing set*) yang berjumlah 35 sampel. Evaluasi dilakukan dengan membandingkan nilai aktual curah hujan dari dataset dengan nilai hasil prediksi yang dikeluarkan oleh sistem. Metrik evaluasi utama yang digunakan dalam penelitian ini adalah *Mean Squared Error* (MSE), *Mean Absolute Error* (MAE), dan *R-Squared* (R^2) *Score*.

Selain pengujian pada *test set*, dilakukan juga pengujian Cross-Validation (5-Folds) pada tahap pelatihan. Teknik ini membagi data latih menjadi 5 bagian berbeda dan melakukan pengujian secara berulang untuk memastikan bahwa model tidak mengalami *overfitting* dan tetap konsisten meskipun diuji dengan sub-dataset yang bervariasi.

Adapun tabel hasil pengujian dengan 10 kali pengujian yaitu sebagai berikut :

Tabel 4. 4 Tabel hasil Pengujian

No	Kelembaban Jam 07.00 (%)	Suhu Minimum (°C)	Curah Hujan Kemarin (mm)	Curah Hujan 3 Hari yang lalu (mm)	Kecepatan Angin rata-rata (m/s)	Tekanan Udara (hPa)	Suhu Rata-Rata 24 Jam (°C)	Hasil Analisis
1	65	27	0	0	2	1012	30	0,00mm Kategori = Tidak Hujan
2	70	28	0	0	2	1013	31	0,00mm Kategori = Tidak Hujan
3	82	25	2	1	3	1010	28	4.9mm Kategori = Hujan Ringan
4	82	26	3	2	3	1010	29	7.41mm Kategori = Hujan Ringan
5	92	24	26	20	4	1009	27	21.76mm Kategori = Hujan Sedang
6	95	24	55	40	10	1010	28	26.92mm Kategori = Hujan Sedang
7	98	20	200	150	13	998	24	50.18mm Kategori = Hujan Lebat
8	100	20	200	180	12	995	23	55.83 mm Kategori = Hujan Lebat
9	100	20	250	200	25	900	20	168.79mm Kategori = Hujan Sangat Lebat
10	100	18	300	250	25	900	19	175.29mm Kategori = Hujan Sangat Lebat

Dari tabel hasil pengujian di atas menunjukkan bahwa sistem mampu memprediksi curah hujan dengan baik berdasarkan 10 skenario input, mulai dari tidak hujan hingga hujan sangat lebat. Nilai prediksi meningkat seiring bertambahnya kelembaban dan curah hujan sebelumnya (rr_lag1 dan rr_lag3), serta menurunnya tekanan udara. Pada kondisi tidak hujan, hasil prediksi 0 mm muncul karena tidak ada riwayat hujan dan tekanan udara masih tinggi. Hujan ringan dan sedang terjadi saat ada peningkatan kecil hingga sedang pada rr_lag dan kelembaban. Sementara itu, hujan lebat dan sangat lebat muncul ketika semua faktor utama seperti kelembaban tinggi, rr_lag besar, dan tekanan udara rendah terjadi secara bersamaan. Secara keseluruhan, model sudah mampu mengikuti pola cuaca secara logis dan konsisten.

Dari table diatas fitur yang mempengaruhi sebagai berikut :

Tabel 4. 5 Fitur Yang Mempengaruhi

Nama Fitur
Kelembaban Jam 07.00
Suhu Rata-Rata 24 Jam
Suhu Minimum
Tekanan Udara

Fitur yang paling mempengaruhi hasil prediksi curah hujan pada sistem ini adalah kelembaban udara, suhu rata-rata 24 jam, suhu minimum, dan tekanan udara. Kelembaban udara menjadi faktor utama karena semakin tinggi kelembaban, semakin besar peluang terjadinya hujan. Suhu rata-rata dan suhu minimum

berpengaruh terhadap proses pembentukan awan dan kondensasi air di atmosfer. Sementara itu, tekanan udara memiliki hubungan terbalik, di mana penurunan tekanan udara biasanya menandakan potensi hujan yang lebih tinggi. Selain itu, fitur lain seperti curah hujan sebelumnya (*rr_lag*) dan kecepatan angin juga ikut berperan, namun kontribusinya lebih kecil dibandingkan empat fitur utama tersebut.

4.5 Evaluasi Model

Berdasarkan hasil eksekusi program pada *test set*, diperoleh hasil evaluasi sebagai berikut:

Tabel 4. 6Hasil Evaluasi Bayesian

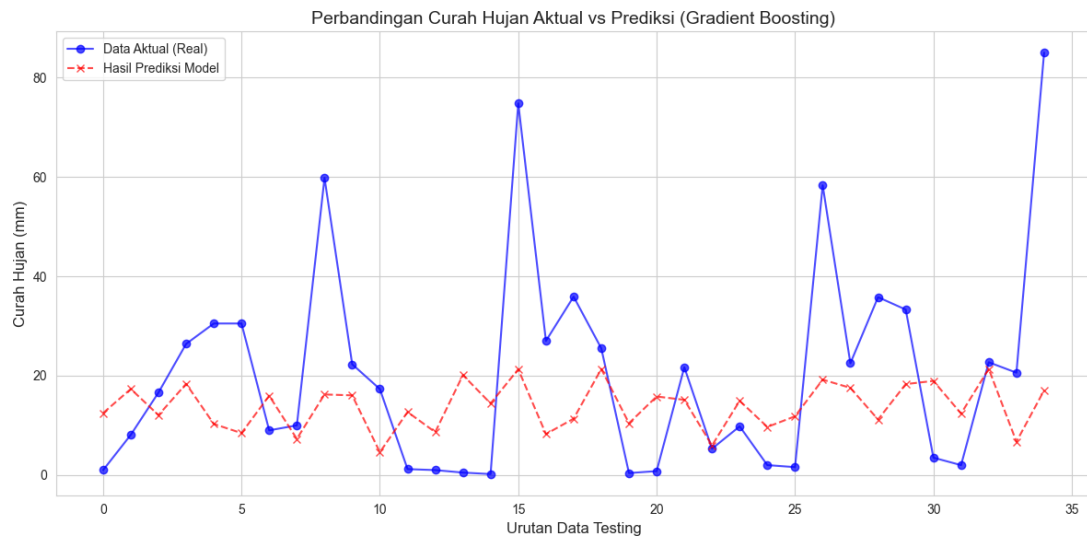
Matrik Evaluasi	Nilai Hasil Pengujian
Rata-rata MSE (CV)	239.2603
MSE (Test Set)	459.0906
MAE (Test Set)	15.7052
R2 Score	-0.0127

Berdasarkan hasil pengujian yang tertera pada Tabel 4.Y, dapat ditarik beberapa kesimpulan mengenai performa model *Bayesian Ridge Regression* dalam memprediksi curah hujan:

1. Analisis Error (MAE & MSE): Nilai MAE sebesar 15.7052 mm menunjukkan bahwa rata-rata kesalahan prediksi model terhadap data riil adalah sekitar 15,7 mm. Mengingat variabilitas curah hujan harian yang cukup ekstrem, nilai ini menunjukkan bahwa model sudah mampu mendekati tren umum meskipun masih terdapat selisih (*margin error*). Nilai MSE sebesar 459.0906 yang lebih tinggi dari rata-rata MSE pada tahap

Cross-Validation (239.2603) mengindikasikan adanya beberapa data uji (*outliers*) yang memiliki lonjakan curah hujan sangat tinggi, sehingga memberikan penalti error yang besar pada model.

2. Analisis R-Squared (R^2) Score: Model menghasilkan R^2 Score sebesar -0.0127. Secara teoretis, nilai R^2 yang negatif atau mendekati nol menunjukkan bahwa model belum mampu menjelaskan varians data dengan sempurna atau pola data curah hujan pada periode pengujian sangat berbeda dengan pola pada data latih. Hal ini sering terjadi pada prediksi cuaca harian yang bersifat non-linear dan sangat fluktuatif (stokastik), di mana parameter cuaca seperti suhu dan kelembaban tidak selalu berbanding lurus dengan jumlah curah hujan yang jatuh pada hari tersebut.
3. Evaluasi Cross-Validation (CV): Hasil rata-rata MSE CV sebesar 239.2603 menunjukkan bahwa pada saat proses pelatihan menggunakan 5-Folds, model sebenarnya memiliki performa yang cukup stabil. Perbedaan antara MSE CV dan MSE Test Set mengonfirmasi bahwa dataset curah hujan memiliki karakteristik *noise* yang tinggi, sehingga tantangan utama dalam penelitian ini adalah menghadapi data cuaca yang tidak konsisten di beberapa periode waktu tertentu.



Gambar 4. 1 Grafik Perbandingan Nilai Aktual dan Prediksi Curah Hujan

Berdasarkan Gambar 4.1, dapat dilihat visualisasi perbandingan antara data aktual dan data hasil prediksi menggunakan algoritma *Bayesian Ridge Regression*. Garis yang saling berhimpit menunjukkan tingkat kedekatan prediksi, sementara fluktuasi pada beberapa titik menunjukkan adanya selisih (*error*) yang disebabkan oleh variasi data cuaca yang ekstrem. Grafik ini mengonfirmasi hasil nilai MAE yang telah dihitung sebelumnya, di mana model mampu mengikuti tren perubahan curah hujan meskipun terdapat *residual* pada nilai-nilai puncak (*peak values*).

Untuk memastikan validitas hasil evaluasi yang dikeluarkan oleh sistem, dilakukan pengujian manual terhadap sampel data uji menggunakan rumus statistik yang telah ditetapkan. Proses validasi ini dilakukan dengan mengambil tiga sampel data pertama dari hasil prediksi (*hasil_prediksi_final.csv*) untuk dihitung selisih absolut serta kuadrat kesalahannya, sehingga tingkat akurasi model dapat dibuktikan secara matematis. Adapun detail perhitungan perbandingan antara nilai aktual dan nilai prediksi tersebut disajikan pada tabel berikut:

Tabel 4. 7b Perhitungan Manual (Data Asli)

Hari Ke	Nilai Aktual	Nilai Prediksi
1	1.0	12.48
2	8.1	17.32
3	16.6	12.6
Total	25.7	41.86

Adapun Rumus dan perhitungan yang digunakan dalam penelitian ini yaitu :

1. Mean Absolute Error (MAE)

$$MAE = \frac{\sum |y_i - \hat{y}_i|}{n} = \frac{25.24}{3} = 8.41 \text{ mm}$$

2. Mean Square Error (MSE)

$$MAE = \frac{\sum (y_i - \hat{y}_i)^2}{n} = \frac{273.41}{3} = 79.13$$

3. R-Squared (R^2) Score

Untuk 3 data ini, rata-rata aktual (y) adalah $25.7/3 = 8.56$

- RSS (Sum of Squares Residual) = 237.41
- TSS (Total Sum of Squares) = $(1.0 - 8.56)^2 + (8.1 - 8.56)^2 + (16.6 - 8.56)^2$
 $= 57.15 + 0.21 + 64.64 = 122.00$

$$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{237.41}{122.00} = 1 - 1.94 = -0.94$$

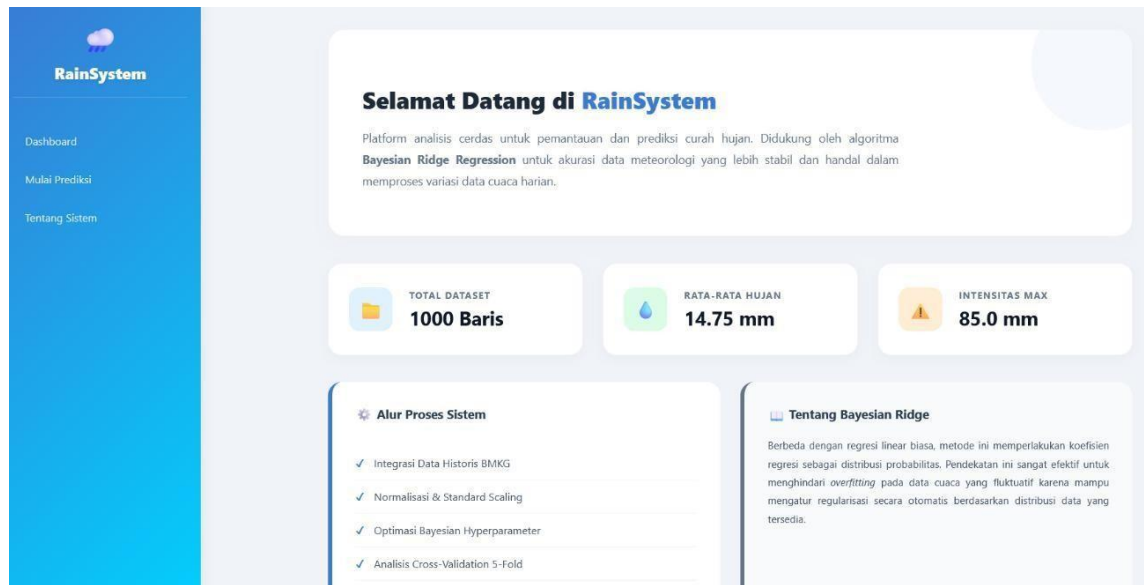
Berdasarkan hasil perhitungan manual pada sampel data uji, validasi terhadap metrik evaluasi menunjukkan bahwa model *Bayesian Ridge* telah mengimplementasikan rumus statistika dengan benar, di mana nilai MAE sebesar 8,41 mm mencerminkan rata-rata selisih absolut antara prediksi dan kenyataan.

Tingginya nilai MSE yang dipicu oleh kuadrat selisih pada data pertama mengonfirmasi sensitivitas model terhadap variasi data yang ekstrem, sementara nilai yang negatif selaras dengan hasil sistem yang menunjukkan kompleksitas hubungan non-linear pada data meteorologi harian. Secara keseluruhan, sinkronisasi antara perhitungan manual dan hasil komputasi ini membuktikan integritas sistem dalam mengukur performa prediksi curah hujan, meskipun terdapat tantangan besar dalam menghadapi fluktuasi data yang sangat stokastik.

4.6 Analisis Hasil Penelitian

Berdasarkan seluruh rangkaian pengujian yang telah dilakukan, dapat dianalisis bahwa model *Bayesian Ridge Regression* memiliki kemampuan untuk memberikan estimasi curah hujan, meskipun menghadapi tantangan besar pada fluktuasi data yang ekstrem. Hal ini dibuktikan dengan perolehan nilai MAE sebesar 15,70 mm, yang berarti rata-rata selisih prediksi model terhadap kondisi riil di lapangan masih berada pada rentang yang wajar untuk skala curah hujan harian. Munculnya nilai MSE sebesar 459,09 dan Score sebesar -0,0127 menjelaskan bahwa variabel cuaca seperti suhu dan kelembaban memiliki hubungan yang sangat kompleks dan tidak selalu linier terhadap intensitas hujan, terutama saat terjadi lonjakan data puncak (*peak values*) yang sulit diprediksi oleh algoritma. Secara keseluruhan, penelitian ini menyimpulkan bahwa meskipun model belum mencapai tingkat presisi yang sempurna karena sifat data meteorologi yang sangat dinamis, penggunaan data hasil pra-pemrosesan telah memberikan fondasi yang stabil bagi sistem untuk mengenali tren cuaca tanpa mengalami gangguan dari data kosong (*missing values*).

4.7 Implementasi Sistem



Gambar 4. 2 Halaman Dashboard

Tampilan dashboard RainSystem dirancang sebagai antarmuka pemantauan cerdas yang menyajikan informasi fundamental mengenai dataset meteorologi secara ringkas dan informatif melalui tiga indikator utama, yaitu total dataset sebanyak 1000 baris, rata-rata curah hujan sebesar 14.75 mm, dan intensitas maksimal mencapai 85.0 mm. Dashboard ini secara eksplisit menjelaskan bahwa sistem bekerja menggunakan algoritma Bayesian Ridge Regression yang dioptimalkan melalui alur proses terstruktur, mulai dari integrasi data historis, normalisasi dengan *Standard Scaling*, hingga tahap evaluasi menggunakan *Cross-Validation 5-Fold*. Melalui kombinasi visualisasi data statistik dan penjelasan teknis mengenai keunggulan metode Bayesian dalam menangani variasi data yang fluktuatif, dashboard ini berfungsi sebagai pusat kendali untuk memberikan transparansi proses pemodelan kepada pengguna sebelum melakukan prediksi curah hujan.

RainSystem

- Dashboard
- Mulai Prediksi
- Tentang Sistem

Prediksi Curah Hujan

Kelembaban Jam 07.00 (%)

Contoh: 85

Suhu Minimum (°C)

Contoh: 23.5

Curah Hujan Kemarin (mm)

Contoh: 10.2

Curah Hujan 3 Hari Lalu (mm)

Contoh: 0

Kecepatan Angin Rata-rata (m/s)

Contoh: 2.5

Tekanan Udara (hPa)

Contoh: 1010.5

Suhu Rata-rata 24 Jam (°C)

Contoh: 27.8

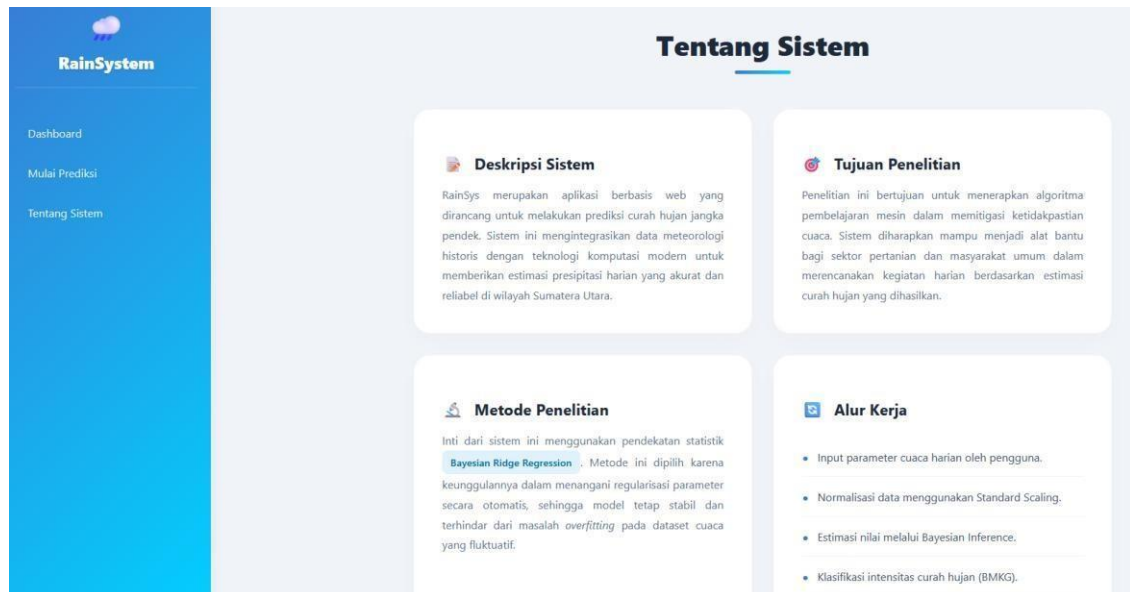
Mulai Prediksi

Hasil Analisis Sistem

Belum ada data diinput. Silakan masukan parameter cuaca di kolom kiri untuk melihat hasil prediksi hari berikutnya.

Gambar 4. 3 Tampilan Prediksi Curah Hujan

Halaman Input Data pada aplikasi RainSystem merupakan antarmuka fungsional yang dirancang untuk memproses prediksi curah hujan secara interaktif melalui pengisian tujuh parameter meteorologi utama. Pengguna dapat memasukkan data spesifik yang meliputi kelembaban jam 07.00, suhu minimum, curah hujan kemarin, curah hujan tiga hari lalu, kecepatan angin rata-rata, tekanan udara, serta suhu rata-rata 24 jam ke dalam formulir yang tersedia. Tampilan ini juga dilengkapi dengan panel Hasil Analisis Sistem di sisi kanan yang akan menampilkan output prediksi secara *real-time* setelah tombol "Mulai Prediksi" ditekan, memberikan pengalaman pengguna yang efisien dalam mentransformasikan data cuaca mentah menjadi informasi prakiraan yang mudah dipahami.



Gambar 4. 4 Tampilan Mengenai Sistem

Halaman Tentang Sistem pada aplikasi RainSystem berfungsi sebagai pusat informasi edukatif yang memberikan pemahaman komprehensif mengenai latar belakang dan mekanisme kerja aplikasi melalui empat panel informasi utama. Bagian Deskripsi Sistem dan Tujuan Penelitian menjelaskan peran aplikasi sebagai solusi berbasis web untuk memitigasi ketidakpastian cuaca di wilayah Sumatera Utara, khususnya bagi sektor pertanian. Sementara itu, panel Metode Penelitian memberikan justifikasi teknis penggunaan algoritma Bayesian Ridge Regression yang unggul dalam menangani stabilitas model pada data fluktuatif, yang didukung oleh panel Alur Kerja yang merinci tahapan proses mulai dari input parameter cuaca, normalisasi dengan *Standard Scaling*, hingga klasifikasi intensitas curah hujan sesuai standar. Secara keseluruhan, tampilan ini memperkuat aspek teoretis penelitian dengan menyajikan kaitan yang jelas antara kebutuhan pengguna dan solusi teknologi yang diimplementasikan.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan mengenai penerapan metode *Bayesian Ridge Regression* untuk prediksi intensitas curah hujan, dapat disimpulkan bahwa:.

1. Tahap pra-pemrosesan data seperti pembersihan *missing values* dan penggunaan *Standard Scaling* terbukti krusial dalam menjaga kestabilan model, di mana dataset berhasil diseleksi menjadi 174 baris data berkualitas untuk proses pelatihan.
2. Berdasarkan hasil pengujian, model menunjukkan performa yang cukup stabil dengan nilai Mean Absolute Error (MAE) sebesar 15.70 mm, yang berarti rata-rata selisih prediksi sistem terhadap data aktual masih berada dalam rentang toleransi untuk estimasi cuaca harian.
3. Meskipun model mampu mengikuti tren umum, nilai MSE (459.09) dan R2 Score (-0.0127) yang rendah menunjukkan bahwa model masih kesulitan dalam memprediksi lonjakan curah hujan ekstrem yang bersifat tiba-tiba dan non-linear.
4. Kesimpulan dari penelitian ini menunjukkan bahwa fitur yang paling berpengaruh terhadap prediksi curah hujan adalah kelembaban udara, suhu rata-rata 24 jam, suhu minimum, dan tekanan udara. Variabel lain seperti *rr_lag* dan kecepatan angin juga berpengaruh, tetapi lebih kecil.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat diberikan untuk pengembangan penelitian selanjutnya:

1. Penambahan Variabel Data: Disarankan untuk menambah variabel fitur lain yang lebih berpengaruh, seperti arah angin regional, suhu permukaan laut, atau indeks kelembaban atmosfer tingkat atas guna meningkatkan korelasi prediksi.
2. Penambahan Volume Dataset: Menggunakan rentang waktu data yang lebih panjang (misalnya data 5-10 tahun terakhir) agar model dapat mengenali pola perubahan iklim tahunan secara lebih mendalam.
3. Optimasi Hyperparameter: Melakukan pencarian parameter optimal (*Grid Search*) yang lebih luas pada fungsi distribusi prior dalam *Bayesian Ridge* untuk meminimalkan nilai error pada data uji.
4. Pengembangan Fitur Aplikasi: Mengembangkan sistem agar dapat menerima input data secara otomatis melalui integrasi API resmi dari BMKG atau stasiun cuaca lokal sehingga prediksi dapat dilakukan secara berkelanjutan.

DAFTAR PUSTAKA

- Afham, M., Nur, I. M., & Utami, W. (2017). Pemodelan Regresi Ridge Pada Kasus Curah Hujan Di Kota Semarang. *Prosiding Seminar Nasional Dan International*, 165–176.
- Alvines, M., Ravi Wijayanto, M., Archi Daffa Danendra, M., Rizky Herdiansyah, M., Karimsyah Lubis, M., Ichsan Farel Rachmad, M., Ditha Tania, K., & Meiriza, A. (2025). Komparasi Ridge Regression, Random Forest, Dan Gradient Boosting Untuk Prediksi Curah Hujan Harian Di Sumatra Selatan Berbasis Time Series Cross-Validation. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(4), 5742–5748. <https://doi.org/10.36040/jati.v9i4.13915>
- Assaf, A. G., Tsionas, M., & Tasiopoulos, A. (2019). Diagnosing and correcting the effects of multicollinearity: Bayesian implications of ridge regression. *Tourism Management*, 71, 1–8. <https://doi.org/10.1016/j.tourman.2018.09.008>
- Astarini, A., & Adriat, R. (2022). *Studi Perbandingan Metode Penentuan Intensitas Curah Hujan Berdasarkan Karakteristik Curah Hujan Kalimantan Barat*. 10(1), 1–7.
- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995). Bayesian Data Analysis. In *Bayesian Data Analysis*. Chapman and Hall/CRC. <https://doi.org/10.1201/9780429258411>
- Golam Kibria, B. M., & Lukman, A. F. (2020). A new ridge-type estimator for the linear regression model: Simulations and applications. *Scientifica*, 2020. <https://doi.org/10.1155/2020/9758378>
- Kruschke, J. K., & Liddell, T. M. (2018). Bayesian data analysis for newcomers. *Psychonomic Bulletin and Review*, 25(1), 155–177. <https://doi.org/10.3758/s13423-017-1272-1>
- Muharsyah, R., Ratri, D. N., Informasi, P., Iklim, P., Meteorologi, B., & No, J. A. (2021). Meningkatkan Kualitas Prediksi Curah Hujan Musiman saat Fase ENSO Menggunakan Metode Bayesian Model Averaging (BMA), Studi Kasus : Pulau Jawa Improving Quality of Seasonal Rainfall Prediction during ENSO Phase using Bayesian Model Averaging (BMA) Meth. 45(1), 13–25.
- Sahrman, S., & Upa, D. (2024). Statistical Downscaling Model with Jackknife Ridge Regression and Modified Jackknife Ridge Regression to Forecast Rainfall. *Indonesian Journal of Statistics and Its Applications*, 8, 155–165. <https://doi.org/10.29244/ijsa.v8i2p155-165>
- Samadi-koucheksaraee, A., & Chu, X. (2024). Development of a novel modeling framework based on weighted kernel extreme learning machine and ridge regression for streamflow forecasting. *Dept 2470*, 1–26.

- Saputra Saragih, M. R., & Akbar Siregar, F. (2024). APLIKASI DETEKSI GAMBAR HASIL SCAN MENGGUNAKAN METODE COSINE SIMILARITY BERBASIS MOBILE. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8, 11598–11604. <https://doi.org/10.36040/jati.v8i6.11546>
- Schreiber-Gregory, D. N. (2018). Ridge Regression and multicollinearity: An in-depth review. *Model Assisted Statistics and Applications*, 13, 359–365. <https://doi.org/10.3233/MAS-180446>
- SISWANTO, S., Gozhi, M. Z. H., Taufik, M. I., Kalondeng, A., & Sahriman, S. (2025). Linearized Ridge Regression Modeling with MM-Estimator in Statistical Downscaling for Rainfall Forecasting. *Jurnal Matematika, Statistika Dan Komputasi*, 21(3), 796–812. <https://doi.org/10.20956/j.v21i3.43203>
- Snaan, R. A., & Shukur, O. B. (2025). *Using Bayesian Ridge Regression Model and ESN for Climatic Time Series Forecasting*. 14(September), 1226–1243. <https://doi.org/10.19139/soic-2310-5070-2637>
- Tew, S. Y., Boley, M., & Schmidt, D. F. (2023). *Bayes beats Cross Validation: Efficient and Accurate Ridge Regression via Expectation Maximization*.
- Wang, Y., Pei, L., & Wang, J. (2024). Precipitation prediction in several Chinese regions using machine learning methods. *International Journal of Dynamics and Control*, 12, 1180–1196. <https://doi.org/10.1007/s40435-023-01250-1>

LAMPIRAN

1. Lampiran Keputusan Dekan



KEPUTUSAN DEKAN
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
Nomor : 1084/KEP/IL3.AU/UMSU-09/F/2025

Tentang :
PENGANGKATAN DOSEN PEMBIMBING SKRIPSI
PRODI TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA

- Dekan Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara, setelah
Menimbang : 1. Bahwa sehubungan dengan pelaksanaan Tugas Akhir Fakultas Ilmu Komputer dan
Teknologi Informasi Universitas Muhammadiyah Sumatera Utara;
2. Bahwa untuk memenuhi maksud tersebut perlu diterbitkan surat keputusan Dekan
Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Muhammadiyah Sumatera
Utara.
Mengingat : 1. Undang-undang Republik Indonesia No. 12 Tahun 2012 tentang Pendidikan Tinggi;
2. Peraturan Pemerintah Nomor 4 Tahun 2014 tentang Penyelenggaraan Pendidikan Tinggi
dan Pengelolaan Perguruan Tinggi;
3. Pedoman Perguruan Tinggi Muhammadiyah;
4. Statuta Universitas Muhammadiyah Sumatera Utara
5. Keputusan Rektor No. 1001/KEP/IL3-AU/UMSU/D/2025 tanggal 25 Oktober 2025
tentang Keputusan Dekan Fakultas Ilmu Komputer dan Teknologi Informasi UMSU Masa Jabatan
2025-2029
6. Keputusan Rektor No. 624/KEP/IL3.AU/UMSU/D/2025 tentang Prodi Fakultas Ilmu
Komputer dan Teknologi Informasi UMSU Masa Jabatan 2025-2029

MEMUTUSKAN

- Menetapkan :
KESATU : Menetapkan nama-nama dalam lampiran surat keputusan ini sebagai Dosen Pembimbing
Prodi Teknologi Informasi Fakultas Ilmu Komputer dan Teknologi Informasi UMSU.
KEDUA : Keputusan ini berlaku sejak tanggal ditetapkan, dan akan ditinjau serta diperbaiki kembali
jika terdapat kekeliruan di kemudian hari.

Ditetapkan di : Medan
Pada tanggal : 18 Jumadil Akhir 1447 H
08 Desember 2025 M

Dr. Al-chovarizmi, M.Kom
NIDN : 0127099201



Tembusan:

1. Yth. Wakil Dekan I & III UMSU di Medan,-
2. Ka. Prodi TI dan Sek. Prodi TI,-
3. Pertinggal.



STARS

2. Lampiran Pengangkatan Dosen Pembimbing Skripsi



UMSU

Unggul! Cerdas! Terpercaya!

MAJELIS PENDIDIKAN TINGGI MUHAMMADIYAH SUMATERA UTARA UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA

FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

UMSU Terakreditasi Unggul Berdaya

Pusat Administrasi: Jalan Mukhtar Basri No. 3 Medan 20236 Telp. (061) 6622400 - 66224567 Fax (061) 6625474 - 6631003

Website: www.umsu.ac.id Email: info@umsu.ac.id umsu@umsu.ac.id umsu@umsu.ac.id umsu@umsu.ac.id umsu@umsu.ac.id

Lampiran Dosen Pembimbing Prodi Teknologi Informasi

Nomor : 1084/KEP/II.3.AU/UMSU-09/F/2025

Tanggal : 18 Jumadil Akhir 1447 H / 08 Desember 2025 M

PENGANGKATAN DOSEN PEMBIMBING SKRIPSI PRODI TEKNOLOGI INFORMASI FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA

NO	NAMA	NPM	JUDUL	DOSEN
1	Sri Ningrum	2209020038	Prediksi Kedalaman Air Sumur Pada Musim Kemarau Menggunakan Algoritma Elastic Net Regression Di Desa Bandar Pinang	Farid Akbar Siregar S.Kom., M.Kom.
2	Muhammad Hafizh Syah Utama	220920212	"Penerapan Metode Moving Average dan Single Exponential Smoothing (SES) dalam Prediksi Penjualan Produk Frozen Food: Studi Kasus Toko Rezeki Baru"	Dr. Zainal Azis, M.Si
3	Leydia Aprisa	2209020181	Klasifikasi Penyakit Batang Jagung Menggunakan MobileNetV3 Berbasis Citra Lapangan	Farid Akbar Siregar S.Kom., M.Kom.
4	Gusti Andrean Gunawan	2209020275	Implementasi Sistem IoT pada Miniatur Jembatan Cerdas dengan Deteksi Kapal Menggunakan Sensor Ultrasonik dan Servo	Mhd. Basri, S.Si, M.Kom
5	Malika Rievy Sakina Pohan	2209020023	Implementasi Transfer Learning sebagai Metode Ekstraksi Fitur Citra dengan Dukungan Algoritma CatBoost dalam Sistem Deteksi Dini Penyakit Kulit Scabies pada Kucing	Zuli Agustina Gultom, M.Si
6	Febrina Susmita Rani	2209020220	Prediksi Penurunan Motivasi Belajar pada Pengguna Edmodo Menggunakan Multilayer Perceptron (MLP) sebagai Model Pembelajaran	Zuli Agustina Gultom, M.Si
7	M. Chandra Reswara	2209020136	Rancang Bangun Smart Bento Box Berbasis IoT (ESP32, Sensor Gas MQ-135 & DHT22) untuk Prediksi Kelayakan Konsumsi Makanan Menggunakan Algoritma XGBoost dan CNN	Mahardika Prawira Tanjung, S.Kom., M.Kom.



Dr. Al-Fahwarizmi., M.Kom
NIDN 90327099201



STARS

3. Lampiran Perbaikan Proposal



MAJELIS PENDIDIKAN TINGGI PENELITIAN & PENGEMBANGAN PIMPINAN PESAT MUHAMMADIYAH
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UMSU Terakreditasi Unggul Berdasarkan Keputusan Badan Akreditasi Nasional Perguruan Tinggi No. 174/SK/BAK-AN/P/TKA/PP/PT/007/2014
Pusat Administrasi: Jalan Mukhtyar Basri No. 3 Medan 20238 Telp. (061) 6622488 - 66224367 Fax. (061) 6635474 - 6631093
Email: info@umsu.ac.id regis@umsu.ac.id umsu@umsu.ac.id umsu@umsu.ac.id umsu@umsu.ac.id
Website: umsu.ac.id umsu.ac.id umsu.ac.id umsu.ac.id umsu.ac.id

FORMULIR PERBAIKAN SEMINAR PROPOSAL

Hari/Tanggal: Selasa, 30 Juni 2024

Nama Mahasiswa : Sri Nurqum
NPM : 2204020028
Program Studi : Teknologi Informasi
Nama Dosen Penanggap : Farid Akbar Sirjan, S.Kom, M.Kom
Judul Proposal : Penerapan Eryman Ridge Regression untuk prediksi curah hujan jangka pendek

Materi/Point yang Diperbaiki

1. penambahan Variabel cuaca
2. perbaiki hasil evaluasi model
3. penambahan Use case diagram

Dosen Pembahas


(.....)

Pembimbing


(.....)





4. Lampiran Acara Pembimbingan Skripsi



MAJLIS PENDIDIKAN TINGGI PENELITIAN & PENGEMBANGAN PIMPINAN PESAT MUHAMMADIYAH

UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA

FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

UMSU Terakreditasi A Berdasarkan Keputusan Badan Akreditasi Nasional Perguruan Tinggi No. 89/SK/BAN-PT/Akred/PT/10/2019
 Pusat Administrasi: Jalan Mukhtar Basri No. 3 Medan 20238 Telp. (061) 6622400 - 66224567 Fax. (061) 6625474 - 6631003
<http://fkip.umsumedan.ac.id> Email: fkip@umsumedan.ac.id Instagram: umsumedan Facebook: umsumedan Twitter: umsumedan YouTube: umsumedan

Berita Acara Pembimbingan Skripsi

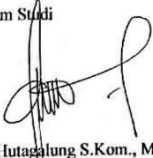
Nama Mahasiswa : Sri Ningrum
 Program Studi : Teknologi Informasi
 NPM : 2209020038
 Judul Penelitian : PENERAPAN BAYESIAN RIDGE REGRESSION UNTUK
 PREDIKSI CURAH HUJAN JANGKA PENDEK

Nama Dosen Pembimbing : Farid Akbar Siregar, S.Kom., M.Kom

Tanggal Bimbingan	Hasil Evaluasi	Paraf Dosen
10/06-2026	Revisi kesimpulan	FI
13/06-2026	revisi web	FI
15/06-2026	revisi daftar pustaka	FI
18/06-2026	revisi daftar pustaka	FI
22/06-2026	revisi pengesahan	FI
18/06-2026	exhibition	FI
18/06-2026	acc sidang	FI
20/06-2026	pemeriksaan kelengkapan dokumen untuk sidang	FI

Diketahui oleh :

Ketua Program Studi



(Fatma Sari Hutagalung S.Kom., M.Kom)

Medan, 25 Juni 2026

Disetujui oleh :

Dosen Pembimbing



(Farid Akbar Siregar, S.Kom., M.Kom)





5. Lampiran Perbaikan Ujian Skripsi



MAJELIS PENDIDIKAN TINGGI PENELITIAN & PENGEMBANGAN PIMPINAN PUSAT MUHAMMADIYAH

UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA

FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMAS

UMSU Terakreditasi Unggul Berdasarkan Keputusan Badan Akreditasi Nasional Perguruan Tinggi No. 174/SK/AN-PT/Ak.Ppy/PT/18/202

Pusat Administrasi: Jalan Mukhtar Basri No. 3 Medan 20238 Telp. (061) 6622400 - 66224567 Fax. (061) 6625474 - 6631003

<http://www.umsu.ac.id>
ids@umsu.ac.id
[umsu.medan](https://www.facebook.com/umsu.medan)
[umsu.medan](https://www.instagram.com/umsu.medan)
[umsu.medan](https://www.linkedin.com/company/umsu.medan)
[umsu.medan](https://www.youtube.com/channel/UC...)

FORMULIR PERBAIKAN UJIAN SKRIPSI

Pada hari ini, Jum'at 17 Juli 2026 telah dilaksanakan Ujian Skripsi bagi mahasiswa Fakultas Ilmu Komputer Dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara. Sbb:

Nama Mahasiswa : Sri Ningrum
 NPM : 2209020038
 Program Studi : Teknologi Informasi
 Judul : Penerapan Bayesian Ridge Regression Untuk Prediksi Curah Hujan Jangka Pendek

Materi/Point yang Diperbaiki :

		Paraf
Farid Akbar Siregar S.Kom., M.Kom.	Buat jurnal	
Martiano, S.Kom., M.Kom.	1. Pada simulasi batasi tanggal nya sesuaikan dengan tanggal dataset yang digunakan 2. Tambahkan parameter akurasi agar tahu berapa tingkat akurasi algoritma yang digunakan	
Dr. Irvan, M.Si		

Berita acara ini ditandatangani setelah skripsi diperbaiki sesuai petunjuk/arahan dari Pembimbing dan Penguji/Pembahas.

6. Hasil Turnitin

PENERAPAN BAYESIAN RIDGE REGRESSION UNTUK PREDIKSI CURAH HUJAN JANGKA PENDEK			
ORIGINALITY REPORT			
24%	22%	12%	15%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS
PRIMARY SOURCES			
1	repository.umsu.ac.id Internet Source	6%	
2	www.coursehero.com Internet Source	1%	
3	Submitted to Universitas Pamulang Student Paper	1%	
4	digilib.unila.ac.id Internet Source	1%	
5	ejournal.itn.ac.id Internet Source	<1%	
6	ejurnal.methodist.ac.id Internet Source	<1%	
7	journal.unhas.ac.id Internet Source	<1%	
8	repository.upi.edu Internet Source	<1%	
9	Submitted to Universitas Muhammadiyah Sumatera Utara Student Paper	<1%	
10	research.lancaster-university.uk Internet Source	<1%	
11	esaformosapublisher.org Internet Source	<1%	
12	Submitted to Universitas Budi Luhur Student Paper	<1%	
13	repository.pertanian.go.id Internet Source	<1%	
14	Submitted to UM Surabaya Student Paper	<1%	
15	Submitted to Universitas Sangga Buana YPKP Student Paper	<1%	
16	Loureiro, Alexandre José Figueiredo. "Análise de propagação em vegetação utilizando Bayes e UT", 2018 Internet Source	<1%	
17	publications.eai.eu Internet Source	<1%	
18	www.iapress.org Internet Source	<1%	