

**PENERAPAN METODE *SUPPORT VECTOR REGRESSION*  
(SVR) PADA PREDIKSI CURAH HUJAN DI JAWA TIMUR**

**SKRIPSI**

**OLEH  
DELLA SAFITRI  
NIM. 19610030**



**PROGRAM STUDI MATEMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2024**

**PENERAPAN METODE *SUPPORT VECTOR REGRESSION*  
(SVR) PADA PREDIKSI CURAH HUJAN DI JAWA TIMUR**

**SKRIPSI**

**Diajukan Kepada  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri Maulana Malik Ibrahim Malang  
untuk Memenuhi Salah Satu Persyaratan dalam  
Memperoleh Gelar Sarjana Matematika (S.Mat)**

**Oleh  
Della Safitri  
NIM. 19610030**

**PROGRAM STUDI MATEMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2024**

**PENERAPAN METODE *SUPPORT VECTOR REGRESSION*  
(SVR) PADA PREDIKSI CURAH HUJAN DI JAWA TIMUR**

**SKRIPSI**

**Oleh  
Della Safitri  
NIM. 19610030**

**Telah Diperiksa dan Disetujui untuk Diuji  
Malang, 10 Desember 2024**

**Dosen Pembimbing I**



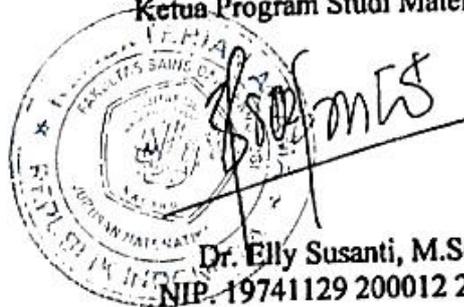
**Dr. Fachrur Rozi, M.Si  
NIP. 19800527 200801 1 012**

**Dosen Pembimbing II**



**Mohammad Nafie Jauhari, M.Si  
NIPPPK. 19870218 202321 1 018**

**Mengetahui,  
Ketua Program Studi Matematika**



**Dr. Elly Susanti, M.Sc.  
NIP. 19741129 200012 2 005**

# PENERAPAN METODE *SUPPORT VECTOR REGRESSION* (SVR) PADA PREDIKSI CURAH HUJAN DI JAWA TIMUR

## SKRIPSI

Oleh:  
**Della Safitri**  
NIM. 19610030

Telah Dipertahankan di Depan Dewan Penguji Skripsi  
Dan Dinyatakan Diterima sebagai Salah Satu Persyaratan  
Untuk Memperoleh Gelar Sarjana Matematika (S.Mat)  
Tanggal 19 Desember 2024

Ketua Penguji : Abdul Aziz, M.Si.  
Anggota Penguji I : Ria Dhea Layla Nur Karisma, M.Si.  
Anggota Penguji II : Dr. Fachrur Rozi, M.Si.  
Anggota Penguji III : Mohammad Nafie Jauhari, M.Si.



Mengetahui,  
Ketua Program Studi Matematika



  
Dr. Elty Susanti, M.Sc.  
NIP. 19741129 200012 2 005s

## PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Della Safitri  
NIM : 19610030  
Program Studi : Matematika  
Fakultas : Sains dan Teknologi  
Judul Skripsi : Penerapan Metode Support Vector Regression (SVR)  
pada Prediksi Curah Hujan di Jawa Timur.

Menyatakan dengan sebenar-benarnya bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambilan data, tulisan, atau pemikiran orang lain yang saya akui sebagai hasil tulisan saya sendiri, kecuali dengan menyertakan sumber cuplikan pada daftar pustaka. Apabila di kemudian hari terbukti atau dapat dibuktikan skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 19 Desember 2024

Yang membuat pernyataan,



Della Safitri  
NIM. 19610030

## **MOTO**

Allah tidak memberikan apa yang kita inginkan melainkan Allah memberikan apa yang kita butuhkan.  
(A-Baqarah: 216)

## **PERSEMBAHAN**

Skripsi ini penulis dipersembahkan kepada Bapak Sholi dan Ibu Siti Khoiriyah tercinta yang senantiasa memberikan doa, dukungan, nasihat, serta motivasi terbaik untuk keberhasilan dan kesuksesan penulis. Kepada adik tersayang Muhammad Hilal Al-Farizi, yang telah memberikan semangat terhadap penulis. Serta kepada kakek Suwaji dan almarhumah nenek Ngatminah yang selalu memberikan doa, semangat, motivasi, dan nasihat yang berharga. Kepada keluarga dan kerabat terdekat, baik dari pihak bapak atau ibu, yang telah membantu dan memberikan doa, semangat, motivasi, dan nasihat terbaik untuk keberhasilan dan kesuksesan penulis. Serta sahabat-sahabat penulis yang telah memberikan doa, bantuan, motivasi, dan semangat dalam menyelesaikan skripsi ini.

## KATA PENGANTAR

Segala puji bagi Allah SWT yang telah melimpahkan rahmat, taufik, serta hidayah-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Penerapan Metode *Support Vector Regression* (SVR) pada Prediksi Curah Hujan di Jawa Timur” dengan baik. Sholawat serta salam selalu tercurahkan kepada Nabi Muhammad Saw yang telah membawa kita dari zaman kegelapan menuju zaman yang terang yakni agama Islam, serta yang akan memberikan syafaatnya di akhirat kelak.

Dalam penyusunan skripsi ini penulis tak lupa menerima bantuan, bimbingan, arahan, dan doa dari berbagai pihak. Oleh karena itu penulis mengucapkan terima kasih sebesar-besarnya kepada:

1. Prof. Dr. H. M. Zainuddin, M.A., selaku rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Prof. Dr. Sri Harini, M.Si., selaku dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Dr. Elly Susanti, S.Pd., M.Sc., selaku ketua Program Studi Matematika, Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Dr. Fachrur Rozi, M.Si., selaku wali dosen dan pembimbing I yang telah memberikan pengetahuan, arahan, nasihat dan solusi kepada penulis.
5. Mohammad Nafie Jauhari, M.Si., selaku dosen pembimbing II yang telah memberikan pengetahuan, arahan, nasihat dan solusi kepada penulis.
6. Abdul Aziz, M.Si., selaku ketua penguji dalam ujian skripsi yang telah memberikan saran yang bermanfaat bagi penulis.
7. Ria Dhea Layla Nur Karisma, M.Si., selaku dosen penguji I dalam ujian skripsi yang telah memberikan saran yang bermanfaat untuk penulis.
8. Seluruh Dosen Program Studi Matematika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim yang telah memberikan banyak ilmunya selama perkuliahan.
9. Bapak Sholi dan Ibu Siti Khoiriyah selaku orang tua penulis dan segenap keluarga yang senantiasa memberikan doa terbaik serta memberikan

dukungan, motivasi, semangat serta materiil sehingga penulis dapat menyelesaikan perkuliahan dan skripsi ini.

10. Seluruh mahasiswa angkatan 2019 dan 2020 teman seperjuangan yang telah memberikan bantuan dan semangat kepada penulis.
11. Teman-teman terbaik penulis yang selalu menemani, membantu, mendukung, dan mau direpotkan selama perkuliahan hingga terselesainya skripsi ini.
12. Semua pihak yang tidak dapat disebutkan satu-persatu, yang telah membantu penulis dalam menyelesaikan skripsi ini.

Semoga Allah SWT melimpahkan rahmat dan karunia-Nya kepada kita semua. Penulis berharap skripsi ini dapat bermanfaat bagi penulis dan semua pihak yang membacanya. *Aamiin Ya Rabbal 'Alamiin*

Malang, 19 Desember 2024

Penulis

## DAFTAR ISI

<b>HALAMAN JUDUL</b> .....	<b>i</b>
<b>HALAMAN PENGAJUAN</b> .....	<b>ii</b>
<b>HALAMAN PERSETUJUAN</b> .....	<b>iii</b>
<b>HALAMAN PENGESAHAN</b> .....	<b>iv</b>
<b>PERNYATAAN KEASLIAN TULISAN</b> .....	<b>v</b>
<b>MOTO</b> .....	<b>vi</b>
<b>PERSEMBAHAN</b> .....	<b>vii</b>
<b>KATA PENGANTAR</b> .....	<b>viii</b>
<b>DAFTAR ISI</b> .....	<b>x</b>
<b>DAFTAR GAMBAR</b> .....	<b>xii</b>
<b>DAFTAR TABEL</b> .....	<b>xiii</b>
<b>DAFTAR LAMPIRAN</b> .....	<b>xiv</b>
<b>DAFTAR SIMBOL</b> .....	<b>xv</b>
<b>ABSTRAK</b> .....	<b>xvi</b>
<b>ABSTRACT</b> .....	<b>xvii</b>
<b>مستخلص البحث</b> .....	<b>xviii</b>
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah .....	6
1.3 Tujuan Penelitian .....	7
1.4 Manfaat Penelitian.....	7
1.5 Batasan Masalah .....	8
<b>BAB II KAJIAN TEORI</b> .....	<b>9</b>
2.1 <i>Data Preprocessing</i> .....	9
2.2 <i>Support Vector Regression</i> .....	12
2.2.1 <i>Linear Regression</i> .....	13
2.2.2 <i>Nonlinear Regression</i> .....	16
2.3 Fungsi Kernel .....	16
2.4 <i>Grid Search Optimization</i> .....	18
2.5 Evaluasi Hasil Prediksi .....	20
2.6 Tingkat Akurasi <i>R-squared</i> .....	21
2.7 Curah Hujan.....	22
2.8 Kajian Integrasi Prediksi dengan AI-quran .....	22
2.9 Kajian Prediksi Curah Hujan dengan Teori Pendukung.....	24
<b>BAB III METODE PENELITIAN</b> .....	<b>26</b>
3.1 Jenis Penelitian .....	26
3.2 Data dan Sumber Data .....	26
3.3 Teknik Pengumpulan Data .....	27
3.4 Instrumen Penelitian .....	27
3.5 Teknik Analisis Data .....	27
3.6 Diagram Alur Penelitian .....	29
<b>BAB IV PEMBAHASAN</b> .....	<b>30</b>
4.1 Analisis Deskriptif.....	30
4.2 <i>Data Preprocessing</i> .....	33
4.2.1 <i>Data Cleaning</i> .....	33
4.2.2 Mendeteksi Pencilan ( <i>Outlier</i> ).....	33

4.3	<i>Resampling Data</i> .....	34
4.4	Normalisasi <i>Min-Max Scaling</i> .....	35
4.5	Proses Pembangunan Model <i>Support Vector Regression</i> Algoritma <i>Grid Search</i> .....	36
4.5.1	Proses Pemodelan Curah Hujan tanpa Faktor yang Mempengaruhi .....	36
4.5.2	Proses Pemodelan Curah Hujan dengan Memperhitungkan Faktor Meteorologi .....	38
4.6	Pemodelan <i>Support Vector Regression</i> .....	39
4.7	Analisis Evaluasi Metode <i>Support Vector Regression</i> .....	44
4.8	Analisis Tingkat Akurasi Metode <i>Support Vector Regression</i> .....	46
4.9	Kajian Penelitian dalam Perspektif Islam .....	46
<b>BAB V</b>	<b>PENUTUP</b> .....	<b>49</b>
5.1	Kesimpulan .....	49
5.2	Saran .....	49
<b>DAFTAR PUSTAKA</b>	.....	<b>51</b>
<b>LAMPIRAN</b>	.....	<b>55</b>
<b>RIWAYAT HIDUP</b>	.....	<b>68</b>

## DAFTAR GAMBAR

Gambar 3.1	Diagram Alir Penelitian .....	29
Gambar 4.1	<i>Boxplot</i> Curah Hujan Per Minggu Jawa Timur dari Tahun 2017-2023.....	32
Gambar 4.2	<i>Boxplot</i> Mendeteksi <i>Outliers</i> .....	33
Gambar 4.3	a) Data dalam Harian, b) Data dalam Mingguan atau Setelah Proses <i>Downsampling</i> .....	35
Gambar 4.4	<i>Plot</i> Prediksi Curah Hujan tanpa Faktor Meteorologi.....	41
Gambar 4.5	<i>Plot</i> Prediksi Curah Hujan tanpa Faktor Meteorologi.....	43

## DAFTAR TABEL

Tabel 1.1	Jumlah Curah Hujan Per Tahun Jawa Timur .....	4
Tabel 3.1	Variabel Penelitian .....	27
Tabel 4.1	Analisis Deskriptif .....	30
Tabel 4.2	Hasil <i>Interquartile Range</i> .....	34
Tabel 4.3	Pembagian <i>Data Training</i> dan <i>Data Testing</i> pada Pemodelan SVR tanpa Faktor yang Mempengaruhi.....	37
Tabel 4.4	Pembagian <i>Data Training</i> dan <i>Data Testing</i> pada Pemodelan SVR dengan Mempertimbangkan Faktor Meteorologi.....	38
Tabel 4.5	Nilai <i>Weight</i> .....	42

## DAFTAR LAMPIRAN

Lampiran 1.	Data Curah Hujan Harian Jawa Timur pada Tahun 2017-2023 .....	55
Lampiran 2.	<i>Data Cleaning</i> dari <i>Data Missing</i> .....	56
Lampiran 3.	Data Curah Hujan Setelah Proses <i>Downsampling</i> .....	58
Lampiran 4.	Normalisasi Data <i>Min-max Scaling</i> .....	59
Lampiran 5.	Proses Perhitungan RMSE dan MAE Prediksi Curah Hujan tanpa Faktor yang Mempengaruhi dan Data Denormalisasi....	59
Lampiran 6.	Proses Perhitungan RMSE dan MAE Prediksi Curah Hujan dengan Memperhitungkan Faktor Meteorologi .....	60
Lampiran 7.	<i>Boxplot</i> Curah Hujan Harian dalam Akurasi Bulanan .....	62
Lampiran 8.	<i>Boxplot</i> Mendeteksi Adanya <i>Outliers</i> .....	62
Lampiran 9.	<i>Scatterplot</i> Data Curah Hujan Harian Jawa Timur.....	63
Lampiran 10.	<i>Scatterplot</i> Data Curah Hujan Mingguan Jawa Timur .....	63
Lampiran 11.	<i>Syntax</i> Model SVR pada Prediksi Curah Hujan Menggunakan <i>R</i> .....	64

## DAFTAR SIMBOL

$IQR$	: <i>Interquartile Range</i>
$Q_1$	: Kuartil pertama
$Q_3$	: Kuartil ketiga
$X_{scaled}$	: Nilai atribut hasil normalisasi
$X$	: Nilai asli atribut
$X_{min}$	: Nilai minimum dari atribut
$X_{max}$	: Nilai maksimum dari atribut
$w$	: <i>Weight</i> atau vektor pembobot
$x$	: Variabel input atau fitur dataset
$x_i$	: Vector fitur dari sampel ke-i
$y$	: Variabel output atau target yang diprediksi
$y_i$	: Nilai target sampel ke-i
$b$	: Bias atau skalar
$C$	: Kompleksitas
$L$	: <i>Primal lagrangian</i>
$W$	: <i>Dual lagrangian</i>
$\alpha, \beta$	: <i>Lagrange multiplier</i>
$\varepsilon$	: <i>Margin error</i>
$\xi_i$	: <i>Slack variable</i>
$l$	: Banyaknya data
$\alpha_i^*$	: <i>Lagrange multiplier</i>
$\beta^*$	: <i>Lagrange multiplier</i>
$\phi$	: Notasi transformasi atau pemetaan
$K(x_i, x_j)$	: Fungsi kernel
$\gamma$	: Parameter <i>gamma</i> fungsi kernel RBF
$\hat{y}$	: Nilai prediksi
$n$	: Jumlah sampel dataset

## ABSTRAK

Safitri, Della. 2024. **Penerapan Metode *Support Vector Regression* (SVR) pada Prediksi Curah Hujan di Jawa Timur**. Skripsi. Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Dr. Fachrur Rozi, M.Si. (II) Mohammad Nafie Jauhari, M.Si.

**Kata Kunci:** Prediksi, Curah Hujan, *Support Vector Regression*, SVR, *Grid Search*

Prediksi masa depan adalah menganalisis tren masa lalu dengan mengasumsikan bahwa tren masa depan akan terus mirip dengan tren historis. Penelitian ini bertujuan untuk mengevaluasi metode *Support Vector Regression* (SVR) dengan algoritma *grid search optimization* dalam memprediksi curah hujan di Jawa Timur. Dua pendekatan utama yang digunakan pada penelitian ini yaitu: prediksi curah hujan tanpa mempertimbangkan faktor meteorologi dan prediksi curah hujan dengan memperhitungkan faktor meteorologi. Hasil penelitian menunjukkan bahwa model SVR tanpa faktor meteorologi, menggunakan kernel *Radial Basis Function* (RBF) dengan parameter  $C = 1$ ,  $\gamma = 0,0625$  dan  $\varepsilon = 0$  menghasilkan tingkat akurasi yang akurat yaitu  $R^2$  sebesar 99,99%. Sementara model SVR dengan memperhitungkan faktor meteorologi, juga menggunakan kernel *Radial Basis Function* (RBF) dengan parameter terbaik diperoleh  $C = 1$ ,  $\gamma = 0,125$  dan  $\varepsilon = 0,5$  memiliki performa yang sedang dengan nilai  $R^2 = 54,56\%$ . Secara keseluruhan, model SVR tanpa faktor meteorologi lebih unggul dalam memprediksi curah hujan di Jawa Timur pada tahun 2017-2023, menjadikannya pilihan yang lebih efektif dalam penerapan model SVR. Peningkatan keakuratan prediksi curah hujan dapat memberikan manfaat signifikan dalam pengambilan keputusan terkait sumber daya air dan mitigasi dampak perubahan iklim di wilayah tersebut.

## ABSTRACT

Safitri, Della. 2024. **The Implementation of the Support Vector Regression (SVR) Method on the Rainfall Prediction in East Java**. Thesis. Mathematics Department, Faculty of Science and Technology, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Advisor: (I) Dr. Fachrur Rozi, M.Si. (II) Mohammad Nafie Jauhari, M.Si.

**Keywords:** Prediction, Rainfall, Support Vector Regression, SVR, Grid Search

Future prediction analyzes past trends by assuming that the future trends will follow historical trends. The research aims to evaluate the Support Vector Regression (SVR) method using the grid search optimization algorithm to predict rainfall in East Java. It employed two main approaches: rainfall prediction with and without considering meteorology factor. It employed two main approaches: rainfall prediction with and without considering the meteorology factor. The research result shows that the SVR model, without considering the meteorology factor, using the Radial Basis Function (RBF) kernel with the parameter  $C = 1$ ,  $\gamma = 0,0625$  and  $\varepsilon = 0$  creates an accurate accuracy level with  $R^2$  value of 99,98%. Meanwhile, the SVR model by considering meteorology factor, also using the Radial Basis Function (RBF) kernel but with the best parameter  $C = 1$ ,  $\gamma = 0,125$  dan  $\varepsilon = 0,5$  has a quite good performance with  $R^2$  value of 54,56%.. Generally, the SVR model, without considering the meteorology factor, is better in predicting the rainfall in East Java in 2017-2023. Therefore, it becomes a more effective choice in the SVR model implementation. The rainfall prediction accuracy can be significantly beneficial in making decisions related to water resources and mitigating climate changes in the area.

## مستخلص البحث

سافطري، ديلا. ٢٠٢٤. تطبيق طريقة انحدار المتجه الداعم (SVR) للتنبؤ بمطول الأمطار في جاوى الشرقية. البحث الجامعي. قسم الرياضيات، كلية العلوم والتكنولوجيا بجامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانج. المشرف الأول: د. فخر الرازي، الماجستير. المشرف الثاني: محمد نافع جوهرى، الماجستير.

**الكلمات الرئيسية:** تنبؤ، هطول أمطار، انحدار متجه الداعم، SVR، البحث الشبكي.

التنبؤ المستقبلي هو تحليل الاتجاهات السابقة من خلال افتراض أن الاتجاهات المستقبلية ستظل مشابهة للاتجاهات السابقة. هدف هذا البحث إلى تقييم طريقة انحدار المتجه الداعم (SVR) باستخدام خوارزمية تحسين البحث الشبكي في التنبؤ بمطول الأمطار في جاوى الشرقية. المدخلان الرئيسيان المستخدمان في هذا البحث هما: التنبؤ بمطول الأمطار دون مراعاة عوامل الأرصاد الجوية والتنبؤ بمطول الأمطار مع مراعاة عوامل الأرصاد الجوية. أظهرت النتائج أن نموذج SVR بدون عوامل الأرصاد الجوية، باستخدام نواة دالة الأساس الشعاعي (RBF) مع المعلمات  $C = 1$  و  $\gamma = 0,0625$  و  $\epsilon = 0$  ينتج مستوى جيدا من الدقة، أي بـ  $R^2 = 99,99\%$ . وفي الوقت نفسه، فإن نموذج SVR، مع مراعاة عوامل الأرصاد الجوية، استخدم أيضا نواة دالة الأساس الشعاعي (RBF) ولكن مع أفضل المعلمات التي تم الحصول عليها  $C = 1$  و  $\gamma = 0,125$  و  $\epsilon = 0,5$  لديه أداء جيد جدا مع قيمة  $R^2 = 54,56$ . بشكل عام، يتفوق نموذج SVR بدون عوامل الأرصاد الجوية في التنبؤ بمطول الأمطار في جاوى الشرقية في الفترة ٢٠١٧-٢٠٢٣، مما يجعله خيارا أكثر فاعلية في تطبيق نموذج SVR. يمكن أن يوفر تحسين دقة تنبؤات هطول الأمطار فوائد كبيرة في صنع القرار المتعلق بالموارد المائية والتخفيف من تأثير تغير المناخ في المنطقة.

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Di era informasi saat ini, masyarakat dan organisasi semakin mengandalkan prediksi untuk membuat keputusan yang lebih tepat dan efisien. Pengembangan model prediksi dalam ilmu statistik menjadi fokus penelitian yang terus berkembang. Peneliti harus berusaha meningkatkan kinerja model untuk mengatasi masalah-masalah yang muncul, seperti mengatasi *overfitting* dan menggali metode baru untuk mengekstraksi informasi yang lebih akurat. Ciri khas dari prediksi adalah melibatkan komponen waktu yang memerlukan ekstrapolasi data deret waktu (*time series*). Prediksi peristiwa masa depan dengan menganalisis tren masa lalu yaitu mengasumsikan bahwa tren masa depan akan terus mirip dengan tren historis (Yudistira dkk., 2023). Namun, tidak semua model prediksi akan memberikan hasil yang sama meskipun menggunakan kumpulan data yang sama. Oleh karena itu, sangat penting untuk memperhatikan dan menentukan model mana yang akan menghasilkan prediksi dengan paling baik (Yudistira dkk., 2023).

Seiring dengan perkembangan teknologi, metode-metode prediksi juga mengalami kemajuan yang pesat dan menciptakan peluang baru untuk meningkatkan akurasi yang relevan. Beberapa metode yang digunakan untuk prediksi yaitu metode *deep learning* seperti algoritma *Recurrent Neural Network* (RNN) dan *Long Short Term Memory* (LSTM), metode runtun waktu (*time series*) seperti *moving averages* dan *exponential smoothing* serta metode *machine learning*

seperti regresi linear, *Support Vector Regression* (SVR), *random forest*, dan lain-lainnya. Pada penelitian ini memutuskan untuk menggunakan salah satu dari metode *machine learning* yaitu *Support Vector Regression* (SVR).

*Support Vector Regression* (SVR) merupakan perkembangan dari metode *Support Vector Machine* (SVM) yang awalnya mengatasi masalah klasifikasi kemudian berkembang untuk mengatasi masalah regresi (Chen dkk., 2004). *Support Vector Regression* (SVR) merupakan salah satu teknik *machine learning* dengan kategori *supervised learning* yang sangat baik dalam mengatasi masalah regresi, seperti prediksi di perusahaan digunakan untuk mengoptimalkan inventaris di masa depan (Trilaksono dkk., 2023). Konsep dasar dari *supervised learning* adalah melibatkan hubungan antara *input* dan *output* yang ingin diprediksi, dengan tujuan untuk mengembangkan model yang dapat menggeneralisasi pola dari *data training* dan mampu memprediksikannya. Perbedaan antara *Support Vector Regression* dan *Support Vector Machine* adalah pendekatan yang digunakan. SVM mengatasi masalah klasifikasi dengan berusaha mencari jarak atau *margin* yang besar sehingga dapat memisahkan sampel-sampel dari kelas berbeda. Sedangkan SVR berusaha untuk mencari *margin* yang dapat menampung sebanyak mungkin sampel di *margin* tersebut (Suparwito dkk., 2023).

*Support Vector Regression* memiliki keunggulan yaitu dapat mengatasi masalah *nonlinear* dengan trik kernel untuk mengatasi pola yang kompleks. Namun untuk memaksimalkan kinerja model ini, optimasi parameter menjadi langkah kritis. Algoritma *grid search* merupakan salah satu metode optimasi yang populer. *Grid search* akan melatih model dengan mencoba semua kombinasi dan mengevaluasi *hyperparameter* dengan *range* telah ditentukan pada *grid*, sehingga

peneliti dapat menemukan parameter optimal untuk meningkatkan akurasi prediksi sekaligus berpotensi mengurangi *underfitting* atau *overfitting* yang mungkin untuk terjadi. Prediksi sering dimanfaatkan dalam kehidupan sehari-hari, misalnya di bidang klimatologi yaitu untuk memprediksi curah hujan, di bidang geofisika untuk memprediksi terjadinya gempa bumi atau gunung meletus, di bidang ekonometrika untuk memprediksi tingkat inflasi dan harga saham, serta di beberapa bidang lainnya.

Air merupakan bagian dari sumber kehidupan di bumi, karena tanpa adanya air kehidupan tidak akan berlangsung. Manusia tidak terlepas dari air dalam menjalankan kehidupan sehari-hari, baik digunakan dalam Pembangkit Listrik Tenaga Air (PLTA), kegiatan industri, sektor ekonomi, budidaya pertanian dan keperluan lainnya. Tuhan Yang Maha Esa memberi air sebagai karunia bagi alam semesta. Hujan merupakan anugerah yang tidak terukur, karena memberikan manfaat bagi keberlangsungan hidup. Dalam Q.S An-Nahl ayat 10-11 dijelaskan bahwa hujan merupakan rahmat bagi makhluk-Nya:

هُوَ الَّذِي أَنْزَلَ مِنَ السَّمَاءِ مَاءً لَكُمْ مِنْهُ شَرَابٌ وَمِنْهُ شَجَرٌ فِيهِ تُسِيمُونَ ۝ ۱۰ يُنبِثُ لَكُمْ بِهِ الزَّرْعَ وَالزَّيْتُونَ  
وَالنَّخِيلَ وَالْأَعْنَابَ وَمِنْ كُلِّ الثَّمَرَاتِ إِنَّ فِي ذَلِكَ لَآيَةً لِّقَوْمٍ يَتَفَكَّرُونَ ۝ ۱۱

Artinya: “10. Dialah yang telah menurunkan air hujan dari langit untuk kalian, sebagaimana menjadi minuman dan sebagiannya (menyuburkan) tumbuh-tumbuhan, yang pada (tempat tumbuhnya) kalian menggembalakan ternak kalian. 11. Dia menumbuhkan bagi kalian dengan air hujan itu tanam-tanaman; zaitun, kurma, anggur, dan segala macam buah-buahan. Sesungguhnya pada yang demikian itu benar-benar ada tanda (kekuasaan Allah) bagi kaum yang memikirkan” (Al-Qur’an dan Terjemahannya, 2019).

Air tidak akan pernah habis karena air merupakan sumber daya alam yang dapat diperbaharui. Tetapi dalam penggunaannya, air tetap harus dibatasi dan dijaga agar terus berkelanjutan (Iswandi dan Indang, 2020). Hampir 72% permukaan bumi tertutup oleh air. Situs web Lembaga Survei Geologi Amerika Serikat (US

*Geological Survey*) menyatakan bahwa terdapat 1.386 juta  $km^3$  atau 330 juta  $mil^3$  air di bumi (Zulfikar, 2022). Diperkirakan luas bumi antara daratan dan air berbanding 2 : 3 (Al-Azizi, 2018). Di antara dinamika siklus hidrologi air, hujan merupakan sumber utama.

Hujan merupakan peristiwa turunnya butir-butir air atau es dari atmosfer ke permukaan bumi. Sedangkan curah hujan merupakan banyaknya air hujan yang jatuh pada suatu daerah dalam waktu tertentu. Indonesia berpotensi memiliki curah hujan yang tidak stabil, sebab terjadinya penguapan dari air ke udara besar mengakibatkan banyaknya curah hujan yang turun pada musim kemarau (Bayu, 2021). Curah hujan di Indonesia bervariasi berdasarkan ruang dan waktu tertentu yaitu harian, mingguan, bulanan, atau tahunan (Laia dan Setyawan, 2020). Menurut *Basic Planet*, Indonesia menempati peringkat kesembilan dari sepuluh negara dengan curah hujan tertinggi di dunia, yaitu dengan curah hujan sebesar 2.702  $mm$  setiap tahunnya (Dihni, 2022).

Curah hujan di Jawa Timur lebih sedikit dibandingkan dengan bagian barat Pulau Jawa, karena Jawa Timur beriklim tropis basah. Setiap tahunnya, curah hujan di Jawa Timur sebesar 1.900  $mm$  selama 100 hari di musim hujan. Merangkum data iklim BPS pada tahun 2017, yaitu jumlah curah hujan di Jawa Timur yang diambil dari stasiun BMKG Juanda 10 tahun terakhir adalah:

**Tabel 1.1** Jumlah Curah Hujan Per Tahun Jawa Timur

<b>Tahun</b>	<b>Jumlah Curah Hujan (<math>mm</math>)</b>
2011	1790
2012	1389
2013	2270
2014	1980
2015	2024

Pada tahun 2018, Jawa Timur berada di urutan ketiga sebagai provinsi dengan jumlah kejadian banjir terbesar setelah Aceh dan Jawa Tengah, yaitu sebanyak 1.422 kejadian banjir. Sedangkan pada tahun 2021, Jawa Timur berada di urutan keempat dengan jumlah kejadian banjir sebanyak 1.176 kasus (BPS, 2022). Laporan BNPB menyatakan bahwa pada tahun 2022, jumlah bencana alam di Indonesia mencapai 3.522 kasus, dengan banjir sebagai bencana yang sering terjadi, yaitu sebanyak 1.520 kasus dan diikuti bencana cuaca ekstrem sebanyak 1.057 kejadian (Mustajab, 2023).

Penelitian lain menggunakan metode SVR dalam memprediksi produktivitas kelapa sawit di Provinsi Riau dengan membandingkan tiga jenis kernel, yaitu linear, polinomial, dan RBF. Model SVR terbaik dalam penelitian tersebut diperoleh dengan menggunakan kernel RBF, yang menghasilkan koefisien determinasi ( $R^2$ ) sebesar 95% dan nilai *error* galat (MSE) sebesar 6% (Mustakim dkk., 2015). Penelitian yang meramalkan inflasi bulanan nasional dengan menggunakan metode *Support Vector Regression* ini memperoleh hasil RMSE sebesar 0,0011 dan  $R^2$  sebesar 0,9999 pada data validasinya, sedangkan akurasi pada model prediksinya memperoleh nilai RMSE sebesar 0,0477 diperoleh dari model data testing inflasi selama 3 bulan yaitu mei hingga juli 2017 (Khatimi dkk., 2017). Kemudian penelitian lain juga memprediksikan jumlah kasus terkonfirmasi COVID-19 di Indonesia dengan menggunakan metode *Support Vector Regression* dan algoritma *grid search*. Penelitian tersebut menyebutkan bahwa model dari kernel polinomial merupakan model terbaik dengan nilai MSE sebesar 161606,1, nilai MAPE sebesar 0,4946215%, dan nilai  $R^2$  sebesar 85,65011%, artinya mempunyai hasil dengan tingkat akurasi atau ketepatan prediksi yang sangat baik (Safira dkk., 2022).

Penelitian selanjutnya menggunakan metode *Support Vector Regression* untuk memprediksi jumlah positif COVID-19 di Jawa Timur, pada penelitian ini membandingkan model sebelum dan sesudah penggunaan algoritma *grid search optimization* serta membandingkan empat kernel trik. Hasil penelitian tersebut menunjukkan bahwa tingkat akurasi terkecil yaitu dengan menggunakan algoritma *grid search optimization* dengan kernel RBF (*Radial Basis Function*), dengan nilai MAPE sebesar 18,3 dan RMSE sebesar 47,1 (Suroyo, 2022).

Penelitian lain memprediksi model curah hujan di Kabupaten Grobogan dengan proses gaussian menggunakan metode pendekatan regresi. Hasil penelitian tersebut menunjukkan bahwa hasil simpangan baku yang diperoleh sebesar 67,36 (Mukid dan Sugito, 2013). Simpangan baku ini berpengaruh terhadap penyebaran data satu sama lain (Sigit, 2008). Kemudian penelitian lain mengimplementasi model CRISP-DM untuk memprediksi curah hujan yang berpotensi menyebabkan banjir menggunakan algoritma CART, yaitu metode *decision tree*. Penelitian tersebut menyatakan bahwa kinerja algoritma cukup baik dengan tingkat akurasi sebesar 89,4% (Hasanah dkk., 2021). Berdasarkan latar belakang dan beberapa hasil peneliti sebelumnya, peneliti memutuskan untuk memprediksikan curah hujan di Jawa Timur pada tahun 2017-2023 menggunakan metode *Support Vector Regression* (SVR) dengan memanfaatkan algoritma *grid search* untuk memperoleh *hyperparameter*.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, adapun rumusan masalah penelitian ini adalah:

1. Bagaimana tingkat akurasi metode *Support Vector Regression* (SVR) dengan algoritma *grid search optimization* dalam memprediksi curah hujan di Jawa Timur pada tahun 2017-2023, jika hanya menggunakan data curah hujan historis, dibandingkan dengan memprediksi curah hujan di Jawa Timur dengan memperhitungkan faktor meteorologi?
2. Model mana yang memiliki tingkat akurasi terbaik dalam memprediksi curah hujan di Jawa Timur pada tahun 2017-2023 menggunakan metode *Support Vector Regression* (SVR) dengan algoritma *grid search optimization*?

### **1.3 Tujuan Penelitian**

Berdasarkan rumusan masalah di atas maka tujuan dari penelitian ini adalah:

1. Menentukan tingkat akurasi metode *Support Vector Regression* (SVR) dengan algoritma *grid search optimization* dalam memprediksi curah hujan di Jawa Timur pada tahun 2017-2023, jika hanya menggunakan data curah hujan historis saja dibandingkan dengan memprediksi curah hujan di Jawa Timur dengan memperhitungkan faktor meteorologi.
2. Mengetahui model yang memiliki tingkat akurasi terbaik dalam memprediksi curah hujan di Jawa Timur pada tahun 2017-2023 menggunakan metode *Support Vector Regression* (SVR) dengan algoritma *grid search optimization*.

### **1.4 Manfaat Penelitian**

Berdasarkan Penelitian ini, diharapkan dapat memberi manfaat:

1. Bagi Peneliti  
Meningkatkan informasi dan pengetahuan mengenai manfaat metode *Support Vector Regression* (SVR) dengan algoritma *grid search optimization* di

kehidupan sehari-hari terutama di bidang klimatologi yaitu untuk memprediksi curah hujan.

## 2. Bagi Pembaca

Sebagai referensi dan memberikan informasi kepada penelitian selanjutnya mengenai penerapan *Support Vector Regression* dengan algoritma *grid search optimization* untuk prediksi di bidang klimatologi yaitu curah hujan.

### 1.5 Batasan Masalah

Berdasarkan rumusan masalah dan tujuan penelitian, maka batasan masalah yang akan diberikan sebagai berikut:

1. Variabel yang digunakan dalam penelitian prediksi curah hujan ini terdapat lima variabel yaitu: kelembaban udara rata-rata, suhu udara rata-rata, kecepatan angin rata-rata, curah hujan dan lamanya penyinaran matahari.
2. Penelitian ini menggunakan kernel *trick* berupa *Radial Basis Function* (RBF).
3. *Range* parameter yang digunakan pada penelitian ini yaitu *cost* ( $C$ ), *gamma* ( $\gamma$ ) dan *error* ( $\epsilon$ ) berdasarkan beberapa rujukan dari penelitian terdahulu.
4. Alternatif pembagian *data training* dan *data testing* adalah: 65%: 35%, 70%: 30%, 75%: 25%, 80%: 20%, 85%: 15%, 90%: 10%, dan 95%: 5% serta pembagian data dilakukan secara acak.

## BAB II

### KAJIAN TEORI

#### 2.1 *Data Preprocessing*

Proses *preprocessing* dilakukan untuk mengonversi data agar memiliki format yang sama. Proses ini meliputi pemilahan, pengayaan, penyaringan, dan pembersihan data yang menghasilkan *preprocessed data*. *Data preprocessing* bertujuan untuk menyelesaikan masalah yang terjadi pada data mentah, seperti data yang hilang, *noise*, dimensi terlalu tinggi dan data yang tidakimbang (Benhar dkk., 2020). Umumnya *data preprocessing* terbagi menjadi empat teknik yaitu :

##### 1. *Data Cleaning*

Proses *data cleaning* terdiri dari teknik untuk menemukan dan memperbaiki data yang tidak lengkap (*missing*) dan mendeteksi *outliers*. *Data cleaning* akan mempengaruhi hasil informasi dari teknik *data mining*, karena data yang diolah akan berkurang jumlah dan kompleksitasnya (Yunardi dan Dina, 2022). Apabila tetap dipaksakan untuk diolah maka *output* yang dihasilkan akan sangat berbeda dibandingkan setelah proses *data cleaning* (Santoso, 2018). Berikut ini adalah proses *data cleaning* dalam mengatasi *data missing* dan *outliers*:

##### a. Mengatasi *Data Missing*

Pada umumnya data atau *database* yang diperoleh dari suatu perusahaan memiliki isi yang tidak sempurna seperti data yang hilang atau data yang tidak valid. Contohnya, pada kasus curah hujan *data missing* terjadi ketika proses pengumpulan/pengukuran data tidak terukur, sehingga data tercatat dengan kode tertentu misalnya `8888`. Selain itu, ketika tidak data artinya

tidak dilakukan pengukuran, tercatat dengan kode `999`. Penanganan *data missing* dapat dilakukan dengan beberapa cara seperti menghapus *data missing*, mengisi *data missing* dengan nilai *mean*, *modus* atau *median*, interpolasi dan lain-lainnya.

b. Mendeteksi dan Menangani *Outliers*

Mendeteksi *outliers* pada sebuah data adalah mengidentifikasi data yang sangat ekstrim yang berada di luar batas rentang normal. Keberadaan *outliers* akan mengganggu keseluruhan analisis data, karena dapat menyebabkan biasanya kesimpulan yang diambil (Santoso, 2018). Keberadaan *outlier* juga dapat meningkatkan nilai *error* dalam analisis data (Werdiningsih dkk., 2022). Beberapa metode untuk mendeteksi *outlier* antara lain visualisasi data dengan *scatterplot* atau *boxplot*, teknik statistik seperti *z-score* atau *Interquartile Range* (Sulianta, 2023). Penanganan *outlier* dapat dilakukan dengan beberapa cara seperti menghapus *outlier*, transformasi data, menggunakan teknik *binning* atau *discretization*, perhitungan *Interquartile Range* (IQR) dan sebagainya.

Berikut cara mendeteksi *outlier* dengan mengimplementasikan perhitungan *Interquartile Range* (IQR) yaitu langkah pertama mencari nilai median, langkah kedua mencari nilai kuartil pertama ( $Q_1$ ) dan nilai kuartil ketiga ( $Q_3$ ), langkah ketiga adalah mencari jarak interkuartil atau selisih jarak antara  $Q_3$  dan  $Q_1$  atau  $(Q_3 - Q_1)$ , langkah terakhir adalah mencari batas rentang atas dan batas rentang bawah dengan rumus:

$$\text{Batas rentang atas} : Q_3 + (IQR \times 1,5) \quad (2.1)$$

$$\text{Batas rentang bawah} : Q_1 - (IQR \times 1,5) \quad (2.2)$$

## 2. *Data Reduction*

Reduksi data merupakan teknik menampilkan data ke dalam dimensi yang lebih sedikit dari data aslinya, namun tetap menghasilkan analisis yang sama. Cara mereduksi data yaitu dengan cara mengeliminasi fitur atau atribut yang memiliki pengaruh sangat sedikit bahkan tidak berpengaruh terhadap analisis data.

## 3. *Data Transformation*

Transformasi data teknik mengubah adalah nilai dan jenis data agar sesuai dengan format yang dibutuhkan dalam analisis. Beberapa teknik dalam transformasi data seperti *normalization*, *smoothing*, dan *encoding*. Penelitian ini menggunakan transformasi data dengan teknik normalisasi yaitu menggunakan metode *min-max scaling*. Metode ini bekerja dengan *scaling* data atau menyesuaikan data ke dalam rentang tertentu, yaitu rentang nilai minimum (0) hingga nilai maksimum (1), sehingga nilai data menjadi seragam (Yudistira dkk., 2023). Normalisasi data multivariat dengan *min-max scaling* dapat dilakukan dengan mengikuti rumus sebagai berikut:

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (2.3)$$

dengan:

$X_{scaled}$  : Nilai atribut hasil normalisasi

$X$  : Nilai asli atribut

$X_{min}$  : Nilai minimum dari atribut

$X_{max}$  : Nilai maksimum dari atribut

Denormalisasi data adalah kebalikan dari normalisasi, yaitu pengembalian data yang telah dinormalisasi dalam rentang nol sampai satu agar kembali ke skala aslinya. Denormalisasi data berguna untuk meningkatkan kinerja *query* atau mengabaikan *anomaly storage* yang signifikan (Wadi, 2023). Adapun denormalisasi dilakukan dengan menggunakan rumus:

$$X_i = X_{scaled}(max - min) + min \quad (2.4)$$

dengan:

$X$  : Nilai asli atribut

$X_{scaled}$ : Nilai atribut hasil normalisasi

$Max$  : Nilai maksimum  $X$

$Min$  : Nilai minimum  $X$

#### 4. *Data Integration*

*Data integration* adalah teknik untuk menjaga integritas data sehingga dataset menjadi koheren dan terintegrasi dengan baik. *Data integration* membantu mengatasi ketidakseragaman data dan menyatukan data dari berbagai sistem atau sumber, sehingga dapat digunakan secara bersama-sama.

#### 2.2 *Support Vector Regression*

*Support Vector Regression* merupakan teknik *machine learning* yang berfokus pada masalah regresi, sedangkan SVM lebih fokus terhadap masalah klasifikasi. Prinsip dasar dari SVR adalah mencari fungsi regresi optimal dengan melalui pencarian *hyperplane* atau fungsi pemisah terbaik yang dapat memisahkan beberapa titik data (*support vector*) dari nilai sebenarnya dengan mempertimbangkan toleransi kesalahan yang telah ditentukan sebelumnya (Riyanarto dkk., 2018). SVR menggunakan fungsi  $\epsilon - insensitive loss$ , di mana

fungsi ini dapat mengontrol parameter yang setara dengan parameter *margin* yang mampu memisahkan *hyperplane*. Selain itu, dengan mempertimbangkan fungsi  $\varepsilon - insensitive loss$ , SVR dapat memastikan ketersebaran variabel ganda, seperti halnya ada klasifikasi SVM. Variabel *Slank*  $\xi$  digunakan mengukur nilai kesalahan pada titik *training* (Chen dkk., 2004).

### 2.2.1 Linear Regression

Regresi linear menentukan *hyperplane* pada *data training* yaitu dengan cara meminimalkan kesalahan. Kesalahan tersebut ditandai dengan *residual*, yaitu selisih antara *output* model dengan nilai sebenarnya dari *data training*. Sehingga tujuan dari regresi linear adalah untuk meminimalkan *residual* tersebut (Luz, 2009). Misalkan diberikan sekumpulan *data training* dengan banyaknya data  $l$  adalah:  $(\mathbf{x}_i, y_i)$  dengan  $i = 1, 2, \dots, l$ , di mana  $\mathbf{x}_i \in \mathbb{R}^n$  adalah vektor *input* dan  $y_i \in \mathbb{R}$  adalah *output* skalar. Fungsi regresi linear dinyatakan sebagai (Chen dkk., 2004):

$$f(x) = \langle w \cdot x \rangle + b \quad (2.5)$$

$w$  adalah vektor bobot,  $\langle \cdot \rangle$  adalah *dot product* antara vektor bobot dan vektor fitur  $x$ , dan  $b$  adalah konstanta bias. Sedangkan vektor bobot  $w$  diperoleh:

$$w = \sum_{i=1}^l (\alpha_i - \alpha_i^*) \mathbf{x}_i \quad (2.6)$$

Pada kasus persamaan (2.6) tujuannya adalah mencari nilai  $w$  terkecil. Hal tersebut dapat dilakukan dengan meminimalkan *norm*  $w$ , yang dinyatakan sebagai  $\|w\|^2 = \langle w \cdot w \rangle$ , di mana  $\|w\|^2$  adalah reguralisasi yang dapat mengontrol kapasitas fungsi, sehingga memperoleh generalisasi yang lebih baik. Adapun penyelesaian masalah optimasi tersebut adalah meminimalkan nilai bobot  $w$ :

$$\frac{1}{2} \|w\|^2 \quad (2.7)$$

dengan syarat:

$$\langle w \cdot x_i \rangle + b - y_i \leq \varepsilon$$

$$y_i - \langle w \cdot x_i \rangle - b \leq \varepsilon$$

Secara umum asumsi pada persamaan (2.7) tidak berlaku pada kasus di dunia nyata.

Kemudian dikenalkan variabel slank  $\xi_i, \xi_i^*$  yang digunakan untuk mencari fungsi regresi yang optimal dengan persamaan minimum yang dinyatakan sebagai berikut:

$$\phi(w, \xi, \xi^*) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) \quad (2.8)$$

dengan syarat:

$$\langle w \cdot x_i \rangle + b - y_i \leq \varepsilon + \xi_i$$

$$y_i - \langle w \cdot x_i \rangle - b \leq \varepsilon + \xi_i^*$$

$$\xi_i, \xi_i^* \geq 0$$

konstanta  $C > 0$  menentukan *trade-off* antara *flatness* dari  $f$  dan jumlah hingga penyimpangan yang lebih besar dari  $\varepsilon$  yang ditoleransi.

Menggeneralisasi hasil yang diperoleh untuk memperkirakan atau memprediksi analisis yang kuat diperlukan fungsi  $\varepsilon - insensitive loss$  (Vapnik, 2000). Pembangunan model SVR dengan memanfaatkan persamaan  $\varepsilon - insensitive loss$  yaitu sebagai berikut:

$$L(y, f(x, \alpha)) = L(|y - f(x, \alpha)|_\varepsilon), \quad (2.9)$$

dengan

$$|y - f(x, \alpha)|_\varepsilon = \begin{cases} 0, & \text{jika } |y - f(x, \alpha)| \leq \varepsilon \\ |y - f(x, \alpha)| - \varepsilon, & \text{lainnya.} \end{cases}$$

Fungsi ini menggambarkan model  $\varepsilon - insensitive loss$  di mana kerugian sama dengan nol ( $0 loss$ ) jika terdapat perbedaan antara nilai prediksi dan nilai observasi. Sedangkan persamaan linear  $\varepsilon - insensitive loss$ , dengan solusi dari (2.8) yang diberikan oleh *saddle point* dari *lagrangian* maka menjadi:

$$\begin{aligned}
 L(w, b, \xi_i, \xi_i^*, \alpha, \alpha^*, \beta, \beta^*) \\
 &= \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) - \sum_{i=1}^l (\beta_i \xi_i + \beta_i^* \xi_i^*) \\
 &\quad - \sum_{i=1}^l \alpha_i (\varepsilon + \xi_i - y_i + \langle w \cdot x_i \rangle + b)
 \end{aligned} \tag{2.10}$$

$\alpha_i, \alpha_i^*, \beta_i, \beta_i^*$  adalah fungsi *lagrange multipliers*:

$$\alpha_i^{(*)}, \beta_i^{(*)} \geq 0$$

untuk memperoleh hasil solusi yang maksimal, perlu dilakukan proses turunan  $L$  terhadap variabel primal  $(w, b, \xi_i, \xi_i^*)$  sehingga diperoleh persamaan:

$$w = \sum_{i=1}^l (a_i - a_i^*) x_i \tag{2.11}$$

dengan menggabungkan persamaan (2.5) dan persamaan (2.11), maka diperoleh persamaan SVR linear sebagai berikut:

$$f(x) = \sum_{i=1}^l (a_i - a_i^*) \langle x \cdot x_i \rangle + b \tag{2.12}$$

Variabel  $w$  dideskripsikan sebagai kondisi linear dari vektor training  $x_i$ . Sedangkan nilai  $b$  diperoleh dengan memanfaatkan kondisi KKT (*Karush Kuhn Tucker*) yaitu:

$$b = y_i - \langle w \cdot x_i \rangle - \varepsilon \text{ untuk } 0 < \alpha_i < C \tag{2.13}$$

$$b = y_i - \langle w \cdot x_i \rangle + \varepsilon \text{ untuk } 0 < \alpha_i^* < C \tag{2.14}$$

### 2.2.2 *Nonlinear Regression*

Model regresi *nonlinear* digunakan dalam menyelesaikan data yang tidak dapat diselesaikan menggunakan model regresi linear. Namun dengan cara yang sama seperti pendekatan SVM *nonlinear*, yaitu pemetaan yang memetakan data ke dalam fitur berdimensi tinggi. Solusi SVR *nonlinear* dapat dipecahkan menggunakan persamaan  $\varepsilon - insensitive loss$  yaitu (Chen dkk., 2004):

$$W(\alpha) = -\frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)K(\mathbf{x}_i, \mathbf{x}_j) - \varepsilon \sum_{i=1}^l (\alpha_i + \alpha_i^*) + \sum_{i=1}^l (\alpha_i + \alpha_i^*) \quad (2.15)$$

dengan

$$\sum_{i=1}^l (\alpha_i + \alpha_i^*) \quad \text{dan} \quad 0 < \alpha_i^* \leq C$$

menggunakan persamaan (2.5) dan persamaan (2.15) dapat ditulis sebagai:

$$w = \sum_{i=1}^l (\alpha_i - \alpha_i^*) \phi(\mathbf{x}_i) \quad (2.16)$$

sehingga diperoleh persamaan SVR *nonlinear*:

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*)K(\mathbf{x}, \mathbf{x}_i) + b \quad (2.17)$$

### 2.3 *Fungsi Kernel*

Sebagian besar beberapa data bersifat *nonlinear*, sedangkan *Support Vector Regression* (SVR) didasarkan pada mesin pembelajaran linear dengan cara mengoptimalkan *margin*. Sehingga diperlukan beberapa teknik pemetaan *nonlinear* agar sebagian besar kumpulan data dapat ditangani menggunakan teknik SVR.

Maka diperlukan fungsi kernel yang efektif untuk mengatasi hal tersebut. Persamaan pemetaan *nonlinear* dari data dalam ruang *input* ke dalam ruang fitur dinyatakan dalam bentuk:

$$f(x) = \sum_{i=1}^l w_i \phi_i(x) + b \quad (2.18)$$

$\phi: X \rightarrow F$  adalah sebuah pemetaan *nonlinear* dari ruang input  $X$  menuju ruang fitur. Fungsi hipotesis dapat dinyatakan sebagai kombinasi linear dari titik-titik *training*, sehingga keputusan dapat dievaluasi hanya dengan menggunakan produk antara *point testing* dan *point training*. Sehingga dapat dinyatakan dalam persamaan:

$$f(x) = \sum_{i=1}^l a_i y_i \langle \phi(x_i) \cdot \phi(x) \rangle + b \quad (2.19)$$

Fungsi  $K$  adalah kernel, sehingga untuk semua  $x, z \in X$  maka berlaku

$$K(x, z) = \langle \phi(x) \cdot \phi(z) \rangle \quad (2.20)$$

$\phi$  adalah pemetaan dari  $X$  ke ruang fitur  $F$ . Vektor fitur tidak dapat dinyatakan secara eksplisit, karena jumlah operasi yang diperlukan untuk menghitung hasil kali dalam belum tentu sebanding dengan jumlah fitur. Penggunaan kernel dapat memungkinkan untuk pemetaan data secara implisit ke dalam ruang fitur. Fungsi kernel yang sering digunakan yaitu (Riyanarto dkk., 2023):

#### 1. Kernel Linear

Kernel linear digunakan untuk *dataset* yang linear yaitu cukup membutuhkan *hyperplane* lurus. Jarak dua vektor adalah jumlah perkalian dari setiap pasangan nilai *input*. Kernel linear dihitung dengan persamaan (2.21)

$$K(x_i, x_j) = \sum x_i \cdot x_j \quad (2.21)$$

## 2. Kernel Polinomial

Kernel polinomial adalah bentuk yang lebih umum dari kernel linear. Kernel polinomial memiliki bentuk *hyperplane* yang melengkung atau *nonlinear* sehingga dapat memisahkan data dengan bentuk yang lebih kompleks. Kernel polinomial dihitung dengan persamaan (2.22)

$$K(\mathbf{x}_i, \mathbf{x}_j) = 1 + \sum (\mathbf{x}_i \cdot \mathbf{x}_j)^m \quad (2.22)$$

$m$  adalah derajat polinomial yang berarti  $m = 1$  berbentuk mirip dengan *hyperplane* linear. Hal ini menunjukkan bahwa derajat tersebut perlu ditentukan secara manual ketika proses *training*.

## 3. Kernel *Radial Basis Function* (RBF)

Kernel *Radial Basis function* adalah fungsi kernel yang populer, karena sering digunakan untuk klasifikasi SVM. RBF dapat memetakan *input space* dalam ruang dengan dimensi tak terhingga. RBF dapat dihitung dengan persamaan (2.23)

$$K(\mathbf{x}_i, \mathbf{x}_j) = e \left( -\gamma \cdot \sum (\mathbf{x}_i - \mathbf{x}_j)^2 \right) \quad (2.23)$$

Nilai gamma yang lebih tinggi akan sangat cocok dengan *data training* yang dapat menyebabkan *overfitting*.

## 4. Kernel Sigmoid (tangen Hiperbolik)

$$K(\mathbf{x}_i, \mathbf{x}_j) = \tanh (\alpha \mathbf{x}_i \cdot \mathbf{x}_j + \beta) \quad (2.24)$$

### 2.4 *Grid Search Optimization*

Kelebihan dari algoritma SVM dan SVR adalah dapat mencapai optimum global dan dapat menangani masalah pada dimensi tinggi. Salah satu metode untuk memperoleh *hyperparameter* tuning adalah *grid search*. Metode ini melakukan pelatihan model dengan mencoba semua kombinasi dan mengevaluasi

*hyperparameter* dengan *range* yang telah ditentukan pada *grid*. Proses ini akan memakan waktu dan komputasi yang besar jika jumlah *hyperparameter* yang akan diuji coba banyak (Putra dkk., 2023). Namun pengaturan parameter dan kernel dapat mempengaruhi kinerja proses pembelajaran dan kinerja umum SVR. Pengujian validasi menggunakan jenis metode tertentu, seperti *cross-fold cross validation* dan *bootstrapping*, dengan kombinasi parameter yang berbeda dapat menyebabkan estimasi tingkat kesalahan pada data tertentu (Muflikhah dkk., 2021).

Adapun parameter SVR dengan *Radial Basis Function* (RBF) kernel yang digunakan meliputi *gamma* ( $\gamma$ ) dan *cost* ( $C$ ) yang membutuhkan nilai optimum. Parameter *cost* ( $C$ ) dari formulasi prediksi SVR untuk mengontrol penalti data latih yang salah diklasifikasikan mengakibatkan kompleksitas fungsi. Oleh karena itu, nilai  $C$  yang tinggi dapat memaksa SVR untuk membuat fungsi prediksi yang cukup kompleks untuk salah mengklasifikasikan data pelatih sesedikit mungkin, sedangkan parameter *cost* yang rendah dapat mengarah ke arah prediksi yang lebih sederhana (Gandhi, 2018). Selain itu, parameter *gamma* menunjukkan seberapa jauh pengaruh data latih tunggal dicapai dengan nilai rendah dan tinggi yang menunjukkan jauh atau dekat. Parameter *gamma* dapat dianggap sebagai kebalikan dari radius pengaruh sampel yang dipilih oleh model sebagai *support vector*.

Hasil akurasi maksimum dalam membagi data menjadi *data training* dan *data testing* memerlukan teknik *cross validation*. Salah satu jenis pengujian *cross validation* adalah *k-fold cross validation*, yang berfungsi untuk menilai kinerja sebuah metode algoritma dengan membagi sampel data secara acak dan mengelompokkan data tersebut ke dalam nilai  $K$  subset sesuai dengan nilai  $K$  pada

*k-fold* (Arisandi dan Hakim, 2022). Penggunaan *k-fold cross validation* dapat mengurangi waktu komputasi yang mungkin terjadi karena adanya proses iterasi dengan tetap menjaga keakuratan estimasi model (Mustika dkk., 2021).

## 2.5 Evaluasi Hasil Prediksi

Beberapa metode yang digunakan untuk mengevaluasi hasil prediksi (Ramadhan dkk., 2023):

### 1. *Mean Absolute Error* (MAE)

*Mean Absolute Error* (MAE) merupakan rata-rata selisih mutlak sebenarnya dengan nilai prediksi. MAE digunakan untuk mengukur keakuratan suatu model dalam melakukan prediksi. Dengan kata lain, MAE menghitung beberapa rata-rata kesalahan absolut dalam prediksi. Semakin kecil nilai MAE, maka semakin baik kualitas model tersebut. Secara matematis persamaan MAE adalah sebagai berikut:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y - \hat{y}| \quad (2.25)$$

dengan:

- $n$  : Jumlah data
- $y$  : Nilai asli atribut
- $\hat{y}$  : Nilai prediksi

### 2. *Root Mean Square Error* (RMSE)

*Root Mean Square Error* (RMSE) merupakan salah satu cara untuk mengevaluasi model dengan mengukur kesalahan (*error*) model pada prediksi data. Keakuratan model ditandai dengan adanya nilai RMSE yang kecil. Semakin kecil nilai RMSE atau yang paling mendekati 0 maka metode tersebut

dinyatakan memiliki model yang baik dan akurat. Secara matematis persamaan dari RMSE adalah:

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(\hat{y}_i - y_i)^2}{n}} \quad (2.26)$$

## 2.6 Tingkat Akurasi *R-squared*

Besarnya nilai *R-squared* atau koefisien determinasi terletak antara 0 dan 1. Jika nilai *R-squared* sama dengan satu, berarti garis regresi yang terbentuk cocok secara sempurna dengan nilai-nilai prediksi yang diperoleh. Salah satu manfaat *R-squared* adalah untuk mengukur besar proporsi atau persentase dari hasil prediksi yang diterapkan oleh model *Support Vector Regression*. Adapun nilai *R-squared* adalah  $0 \leq R^2 \leq 1$ :

$R^2 = 0$  artinya tidak ada hubungan antara variabel dependen dan variabel independent, atau model regresi yang terbentuk tidak tepat untuk meramalkan variabel Y.

$R^2 = 1$  artinya garis regresi yang terbentuk dapat meramalkan Y secara sempurna. Semakin dekat nilai  $R^2$  ke nilai 1, maka semakin cocok (tepat) garis regresi yang terbentuk untuk memprediksi (Siagian & Sugiarto, 2000). Formula *R-squared* pada dasarnya yaitu:

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (2.27)$$

dengan:

- $y_i$  : Nilai aktual
- $\hat{y}_i$  : Nilai prediksi
- $\bar{y}$  : Rata-rata nilai aktual

## 2.7 Curah Hujan

Hujan adalah air dalam bentuk cair atau padat yang jatuh dari atmosfer sampai ke permukaan bumi. Hujan terjadi apabila butir-butir uap air yang terkumpul dalam bentuk awan bersatu menjadi lebih besar dan berat. Berubahnya uap air dari bentuk gas saat penguapan menjadi butir-butir air disebut kondensasi. Besarnya curah hujan dijabarkan sebagai banyaknya air yang mencapai permukaan tanah dalam waktu tertentu. Curah hujan dinyatakan dalam inci atau milimeter (1 inci = 25 mm). Banyaknya hujan yang jatuh ke bumi diketahui dengan mengukur besarnya curah hujan menggunakan penakar hujan yaitu *fluviometer* dan *ombrometer* (Bayu, 2021).

## 2.8 Kajian Integrasi Prediksi dengan Al-quran

Dalam bahasa arab prediksi artinya adalah التنبؤات. Sedangkan menurut KBBI prediksi adalah proses memperkirakan nilai di masa mendatang berdasarkan informasi dari masa sebelumnya dengan menggunakan metode atau algoritma. Secara implisit dalam Q.S Rum ayat 2-4 menjelaskan tentang prediksi Al-Qur'an terhadap kejadian di masa depan yaitu:

Artinya: “2) Bangsa romawi telah dikalahkan, 3) di negara yang terdekat dan mereka setelah kekalahannya itu akan menang, 4) dan beberapa tahun (lagi), Milik Allahlah urusan sebelum dan setelah (mereka menang). Pada hari itu (kemenangan bangsa Romawi) itu bergembiralah orang-orang mukmin” (Al-Qur'an dan Terjemahannya, 2019).

Tafsir dari ayat di atas menceritakan kekalahan bangsa Romawi (Ahli Kitab Nasrani) tepatnya Kerajaan Romawi Timur yang berpusat di Konstantinopel oleh bangsa Persia (Musyrik) yaitu Syiria. Beberapa tahun kemudian setelah mereka dikalahkan, maka bangsa Romawi akan mengalahkan bangsa Persia sebagai balasan atas kekalahan itu. Ayat 2 dalam Surat ar-Rum menjelaskan tentang hal-hal

gaib yang menunjukkan kemukjizatan al-Qur'an dan menerangkan bahwa ayat ini diturunkan saat bangsa Romawi dikalahkan oleh bangsa Persia. Namun kekalahan itu tidak akan lama diderita oleh bangsa Romawi. Setelah tujuh atau delapan tahun kekalahan bangsa Romawi dari bangsa Persia, peperangan antara kedua bangsa itu berperang kembali untuk kedua kalinya dan bangsa Romawi yang memenangkan peperangan tersebut (*Al-Qur'an dan Terjemahannya*, 2019).

Cara menguasai ilmu prediksi adalah dengan mengamati peristiwa yang terjadi saat ini untuk melihat perkembangan *trend* yang akan terjadi di masa depan. Sebagai muslim sepatutnya menyadari bahwa pentingnya ilmu prediksi atau peramalan terutama di dalam Al-Qur'an, karena al-Qur'an adalah kitab suci umat islam yang dipercaya berisi firman Allah SWT dan sebagai pedoman bagi umat manusia.

Kemudian Q.S An-Nahl ayat 8 menjelaskan alat transportasi di masa depan yang artinya:

Artinya: “(Dia telah menciptakan) kuda, bagal, dan keledai untuk kamu tunggangi dan (menjadi) perhiasan. Allah menciptakan apa yang tidak kamu ketahui” (*Al-Qur'an dan Terjemahannya*, 2019).

Allah SWT menyebutkan beberapa binatang ternak yaitu kuda, bagal dan keledai yang bermanfaat bagi manusia. Hewan ternak tersebut digunakan sebagai kendaraan, dijadikan sebagai hewan peliharaan, dan menjadi perhiasan. Pada zaman itu, manusia masih belum mengetahui tentang manfaat hewan ternak tersebut. Dari ayat tersebut dapat diketahui bahwa Allah SWT menciptakan untuk hamba-Nya apa yang diketahui dan tidak diketahui pada saat ini namun kelak akan kamu ketahui manfaat dan kegunaannya (*Al-Qur'an dan Terjemahannya*, 2019). Kemudian dari zaman ke zaman saat ini sudah ada alat transportasi modern seperti

sepeda motor, mobil, pesawat dan lain-lainnya yang semakin bagus bahkan saat ini alat transportasi juga sebagai perhiasan dan sebagai alat status sosial.

## 2.9 Kajian Prediksi Curah Hujan dengan Teori Pendukung

Kajian prediksi curah hujan pada penelitian ini memiliki beberapa tahapan. Tahapan pertama yang harus dilakukan adalah analisis deskriptif untuk mengetahui karakteristik data, seperti jumlah data, mean, standar deviasi, nilai minimum dan nilai maksimum. Selanjutnya *preprocessing data* meliputi *data cleaning* dari *data missing* dan *outliers*. Penanganan *missing value* dengan mengisi nilai yang hilang dengan nilai *mean* dari masing-masing variabel, sedangkan *outliers* ditangani dengan menggunakan metode *Interquartile Range (IQR)*. Tahapan selanjutnya yaitu *resampling data* dengan mengubah frekuensi data curah hujan menjadi data curah hujan mingguan. Data selanjutnya di normalisasi untuk menyamakan satuan dari variabel dengan menggunakan *min max scaling*.

Kemudian proses pembangunan model *Support Vector Regression (SVR)* meliputi: membagi data menjadi *data training* dan *data testing* dengan beberapa alternatif proporsi pembagian data, pemilihan atribut untuk variabel dependen dan variabel independen. Selanjutnya menetapkan *range* parameter dari model SVR, yaitu menggunakan kernel RBF dengan *range* parameter *cost (C)* dan *gamma ( $\gamma$ )* =  $2^{-5}$  hingga  $2^5$  serta *epsilon ( $\epsilon$ )* = *seq(0,1,0.1)*, yaitu dimulai dari 0 hingga 1 dengan langkah 0,1. Setelah itu, dengan menggunakan algoritma *grid search* akan diperoleh parameter terbaik berdasarkan hasil metrik evaluasi MAE dan RMSE terkecil. Sebelum menghitung metrik evaluasi kesalahan, data terlebih dahulu dinormalisasikan yaitu mengembalikan data ke bentuk awal setelah dinormalisasikan.

Mengulangi proses pembagian *data training* pada beberapa proporsi yang telah ditentukan, sehingga diperoleh nilai MAE dan RMSE yang terkecil sebagai acuan untuk memilih parameter optimal. Kemudian pemodelan SVR dengan mencari nilai matriks kernel K untuk memperoleh nilai *weight* dan *bias* serta visualisasi berupa *scatterplot* perbandingan antara data aktual dan hasil prediksi curah hujan Jawa Timur. Selanjutnya menghitung akurasi prediksi menggunakan *R-squared*. Terakhir adalah interpretasi hasil prediksi yang telah diperoleh.

## **BAB III**

### **METODE PENELITIAN**

#### **3.1 Jenis Penelitian**

Jenis penelitian yang digunakan dalam penelitian ini adalah pendekatan studi literatur dan penelitian kuantitatif. Pendekatan studi literatur terdiri dari proses mengupas, membandingkan, meringkas, dan mengumpulkan literatur atau pustaka (Restu dkk., 2021). Sedangkan penelitian kuantitatif mulai awal hingga pembentukan model penelitian sistematis dan terstruktur. Dengan penelitian kuantitatif berhubungan dengan angka disertai grafik atau gambar pada tahap kesimpulannya.

#### **3.2 Data dan Sumber Data**

Data yang digunakan dalam penelitian ini merupakan data sekunder yaitu data yang diambil dari data website resmi BMKG Indonesia. Populasi dalam penelitian ini adalah curah hujan harian di seluruh Indonesia, sedangkan sampel yang digunakan dalam penelitian ini adalah data curah hujan harian di Jawa Timur selama enam tahun terakhir yaitu mulai 1 Januari 2017 hingga 31 Desember 2023 pada seluruh stasiun BMKG di Jawa Timur, sehingga diperoleh data sebanyak 2556 data. Hal tersebut untuk memperoleh informasi yang lengkap dan akurat mengenai kondisi curah hujan di Jawa Timur. Data yang digunakan bersumber dari website resmi BMKG meliputi variabel dependen ( $Y$ ) dan variabel independen ( $X$ ). Adapun beberapa variabel yang digunakan adalah sebagai berikut:

**Tabel 3.1** Variabel Penelitian

No	Variabel		Skala
1	Curah hujan	Y	Ordinal
2	Suhu udara rata-rata ( $^{\circ}C$ )	$X_1$	Rasio
3	Kelembaban udara rata-rata (%)	$X_2$	Nominal
4	Lamanya penyinaran matahari ( <i>jam</i> )	$X_3$	Nominal
5	Kecepatan angin rata-rata ( <i>knot</i> )	$X_4$	Rasio

### 3.3 Teknik Pengumpulan Data

Teknik pengumpulan data pada penelitian ini adalah teknik dokumentasi. Teknik dokumentasi yaitu data-data yang mudah diakses, ditinjau dengan mudah sehingga kasus yang diteliti menjadi baik (Anggito dan Setiawan, 2018).

### 3.4 Instrumen Penelitian

Instrumen Penelitian yang digunakan dalam penelitian ini berupa *software Rstudio* dan *Microsoft Excel*.

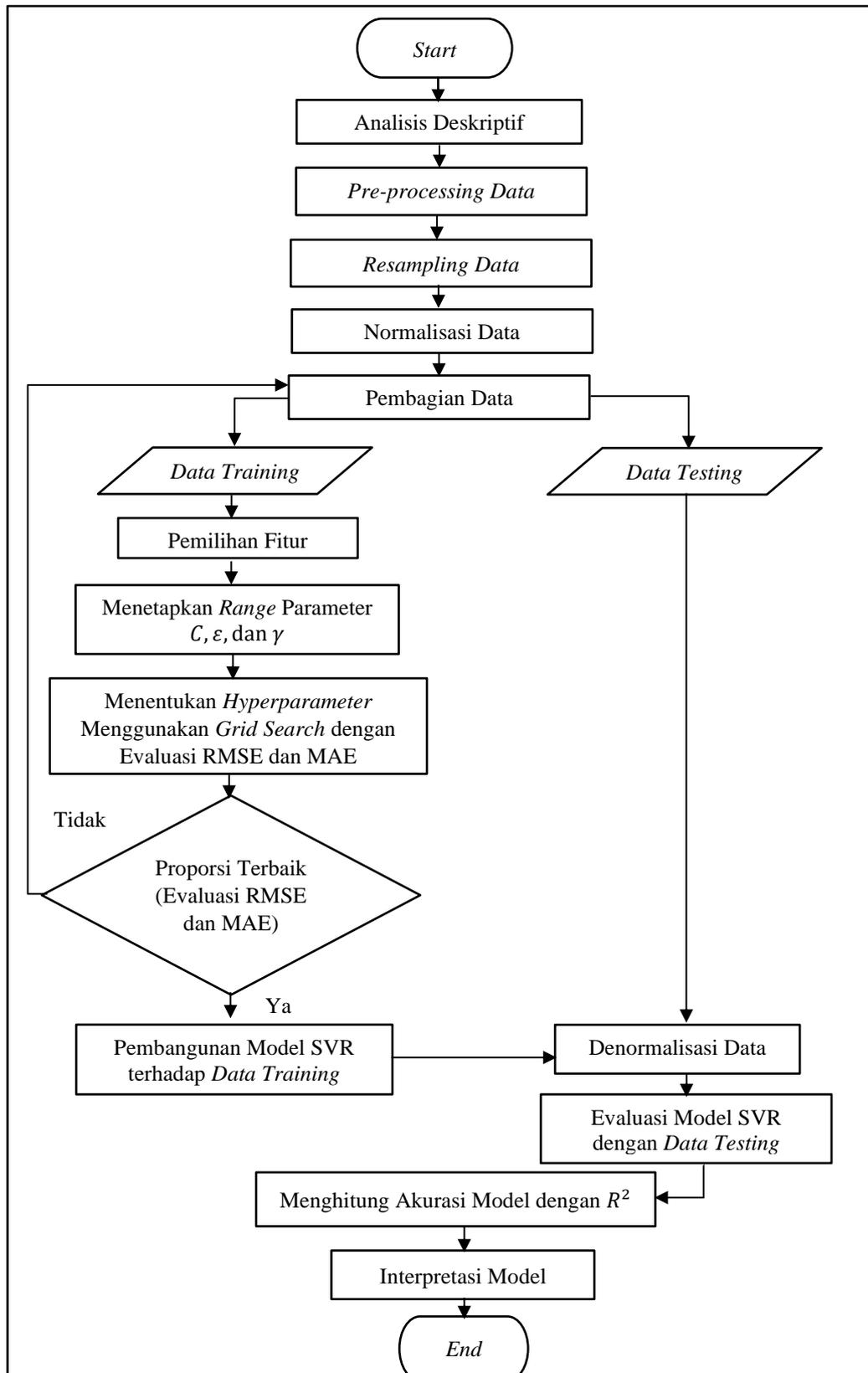
### 3.5 Teknik Analisis Data

Tahapan analisis data pada penelitian ini, yaitu:

1. Melakukan analisis deskriptif untuk mengetahui karakteristik umum data penelitian.
2. *Preprocessing* data terdiri dari penanganan data *outlier* dan *data missing* yaitu proses membersihkan data.
3. *Resampling data* yaitu mengubah frekuensi data curah hujan harian menjadi data curah hujan mingguan.
4. Melakukan normalisasi data untuk menyamakan satuan dari variabel.
5. Proses pembangunan model *Support Vector Regression (SVR)* pada *data training*.

- a. Membagi data menjadi *data training* dan *data testing* dengan beberapa alternatif proporsi pembagian data sesuai yang diberikan dalam batasan masalah.
  - b. Pemilihan fitur untuk variabel dependen dan variabel independen.
  - c. Menetapkan *range* parameter dari model SVR, yaitu kernel RBF dengan parameter *cost* ( $C$ ), *epsilon* ( $\epsilon$ ) dan *gamma* ( $\gamma$ ).
  - d. Penentuan parameter optimal menggunakan *grid search* berdasarkan metrik evaluasi RMSE dan MAE yang memiliki nilai kesalahan terkecil.
6. Mengulangi langkah (5) untuk setiap proporsi pembagian data.
  7. Pembangunan model *Support Vector Regression* (SVR).
  8. Denormalisasi data yaitu mengembalikan data ke bentuk awal setelah data dinormalisasi.
  9. Melakukan evaluasi pemodelan SVR yang telah dilatih menggunakan *data testing*.
  10. Menghitung keakuratan model SVR berdasarkan nilai *R-squared*.
  11. Interpretasi hasil keakuratan model.

### 3.6 Diagram Alur Penelitian



Gambar 3.1 Diagram Alir Penelitian

## BAB IV

### PEMBAHASAN

#### 4.1 Analisis Deskriptif

Variabel dependen ( $Y$ ) pada penelitian ini adalah curah hujan harian di Jawa Timur selama tujuh tahun yaitu dari tahun 2017 hingga tahun 2023. Sedangkan untuk variabel independen terdiri dari suhu udara rata-rata ( $X_1$ ), kelembaban udara rata-rata ( $X_2$ ), lamanya penyinaran matahari ( $X_3$ ), dan kecepatan angin rata-rata ( $X_4$ ). Analisis deskriptif merupakan ilmu statistika yang berguna untuk meringkas dan mendeskripsikan data sehingga data dapat mudah dipahami. Selain itu, analisis deskriptif juga memperjelas dalam menafsirkan data yang lebih efisien melalui perhitungan numerik, grafik atau tabel. Beberapa proses dari analisis deskriptif seperti tabulasi, *mean*, *median*, standar deviasi hingga analisis deret waktu atau *time series*. Berikut adalah hasil statistik deskriptif curah hujan harian Jawa Timur selama tujuh tahun:

**Tabel 4.1** Analisis Deskriptif

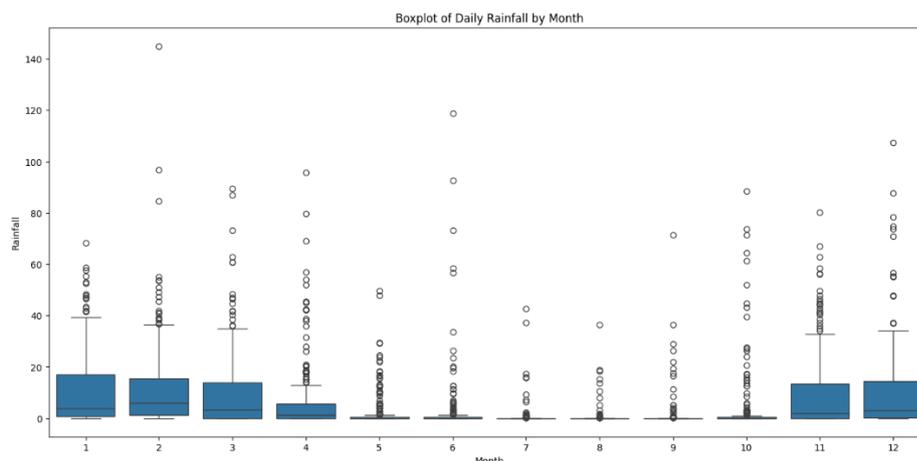
Variabel	$N$	$Mean$	Std.Deviasi	Minimum	Maksimum
$Y$	2304	6,48	13,99	0,00	145,00
$X_1$	2535	24,06	1,15	19,60	28,8
$X_2$	2534	78,33	7,48	6,00	96,00
$X_3$	2524	6,19	2,85	0,00	11,30
$X_4$	2540	1,82	0,76	0,00	6,00

Berdasarkan tabel 4.1 di atas jumlah data curah hujan tidak lengkap karena berjumlah 2304, sehingga variabel curah hujan memiliki *data missing* sebanyak 252 data. Kemudian nilai rata-rata *mean* dari variabel ( $Y$ ) yaitu curah hujan harian di Jawa Timur sebesar 6,48 mm dengan nilai standar deviasi atau ukuran sebuah sebaran data dari curah hujan sebesar 13,99 mm, artinya data curah hujan adalah

data yang bervariasi atau data lebih tersebar. Sedangkan nilai minimum curah hujan sebesar 0 dan nilai maksimum curah hujan sebesar 145 mm. Pada variabel  $X_1$  yaitu suhu udara rata-rata menunjukkan jumlah data sebanyak 2535 data dengan 21 data adalah *data missing*. Selain itu, variabel  $X_1$  memiliki nilai *mean* sebesar 24,06 °C dengan nilai standar deviasi sebesar 1,15°C, serta nilai minimum kecepatan angin rata-rata sebesar 19,6°C dengan nilai maksimumnya sebesar 28,8 °C.

Variabel ( $X_2$ ) yaitu kelembaban udara rata-rata menunjukkan jumlah data sebanyak 2534 data, artinya terdapat *data missing* sebanyak 22 data, serta memiliki nilai *mean* sebesar 78,33%, dengan nilai standar deviasi sebesar 7,48%. Sedangkan nilai minimum dan nilai maksimum dari kelembaban udara rata-rata secara berurutan adalah sebesar 6% dan 96%. Selanjutnya variabel lamanya penyinaran matahari ( $X_3$ ) memiliki jumlah data sebanyak 2524 data dengan *data missing* sebanyak 32 data, serta nilai *mean* sebesar 6,19 jam dengan nilai standar deviasi sebesar 2,85 jam. Sedangkan nilai minimum dan maksimum dari variabel lamanya penyinaran matahari secara berurutan adalah sebesar 0 jam dan 11,3 jam.

Terakhir adalah variabel kecepatan angin rata-rata ( $X_4$ ) yang memiliki jumlah data sebanyak 2540 data, artinya terdapat *data missing* sebanyak 16 data, serta nilai *mean* sebesar 1,82 m/s, dengan nilai standar deviasi sebesar 0,76 m/s. Sedangkan nilai minimum dan maksimum dari variabel kecepatan angin rata-rata secara berurutan adalah sebesar 0 m/s dan 6m/s. Sedangkan kondisi curah hujan harian di Jawa Timur jika diakumulasikan setiap bulan pada tahun 2017-2023 dapat dilihat pada gambar *boxplot* berikut:



**Gambar 4.1** *Boxplot* Curah Hujan Per Minggu Jawa Timur dari Tahun 2017-2023

Tabel analisis deskriptif di atas menunjukkan rata-rata akumulasi curah hujan Jawa Timur secara keseluruhan adalah sebesar  $6.48 \text{ mm}$ . Nilai rata-rata curah hujan tersebut digunakan sebagai acuan untuk membandingkan rata-rata curah hujan harian tiap bulan, untuk mengidentifikasi kondisi dan pola musiman curah hujan di Jawa Timur. Gambar 4.1 *boxplot* di atas menunjukkan bahwa pada bulan November hingga bulan April, nilai rata-rata curah hujan lebih tinggi dibandingkan nilai rata-rata akumulasi curah hujan secara keseluruhan. Hal tersebut menunjukkan bahwa terjadi musim hujan, dengan peningkatan curah hujan dari bulan November hingga bulan Februari dan penurunan curah hujan pada bulan Maret hingga April.

Pada bulan Mei, Juni dan Oktober memiliki nilai rata-rata yang mendekati nilai rata-rata akumulasi curah hujan secara keseluruhan, hal tersebut menunjukkan bahwa pada bulan tersebut terjadi pergantian musim hujan menuju musim kemarau atau sebaliknya yang dikenal dengan musim pancaroba. Sedangkan pada bulan Juli, Agustus, dan September memiliki nilai rata-rata dibawah nilai rata-rata akumulasi curah hujan secara keseluruhan, artinya pada tiga bulan tersebut sudah memasuki musim kemarau.

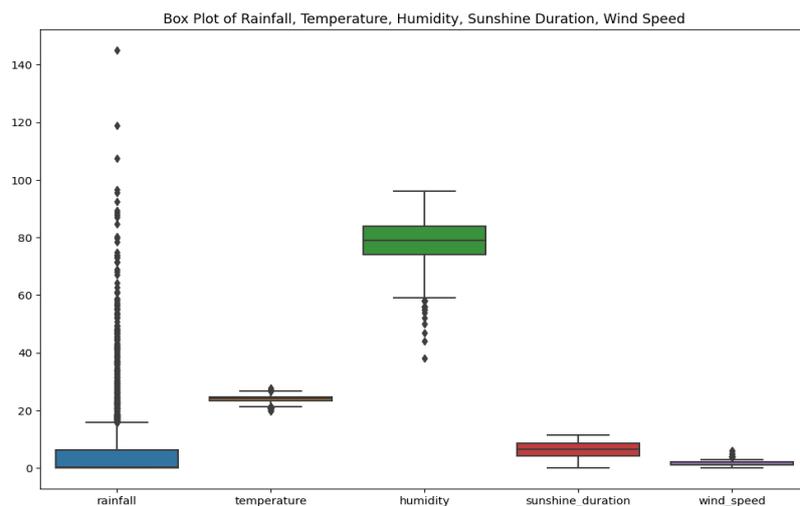
## 4.2 Data Preprocessing

### 4.2.1 Data Cleaning

Setelah dilakukan analisis deskriptif di atas, dapat diketahui bahwa jumlah *data missing* secara keseluruhan sebanyak 343 data, sehingga diperlukan penanganan untuk mendapatkan hasil yang optimal dan akurat. Pada penelitian ini memutuskan untuk menangani *data missing* dengan mengganti nilai yang hilang menggunakan nilai rata-rata dari masing-masing variabel.

### 4.2.2 Mendeteksi Pencilan (*Outlier*)

Salah satu cara untuk mendeteksi *outlier* adalah dengan menggunakan visualisasi *boxplot*. Berikut ini adalah *boxplot* dari data curah hujan Jawa Timur.



**Gambar 4.2** *Boxplot* Mendeteksi *Outliers*

Gambar 4.2 di atas dapat dilihat bahwa variabel curah hujan, suhu udara rata-rata, kelembaban udara rata-rata dan kecepatan angin rata-rata memiliki pencilan atau *outlier*. Sedangkan pada variabel lamanya penyinaran matahari tidak terdeteksi adanya *outlier*. Penanganan *outlier* pada penelitian ini menggunakan metode perhitungan *Interquartile Range* (IQR), yang dihitung menggunakan persamaan (2.1) dan (2.2) sehingga diperoleh:

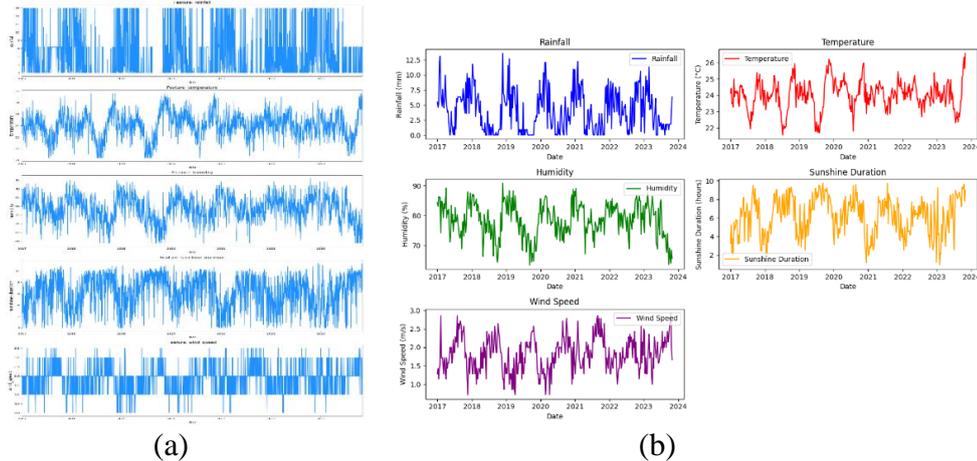
**Tabel 4.2** Hasil *Interquartile Range*

Parameter	Nilai
$Q_1$ (Kuartil 1)	0
$Q_3$ (Kuartil 3)	6,484
<i>Interquartile Range</i>	6,484
Batas Rentang Atas	16,21
Batas Rentang Bawah	-9,73

Nilai batas rentang atas dan batas rentang bawah curah hujan tersebut, digunakan untuk menangani nilai-nilai ekstrim curah hujan yang dapat mempengaruhi hasil analisis. Apabila terdapat data curah hujan yang melebihi batas rentang atas, nilai tersebut akan diganti dengan batas atas. Sebaliknya, jika data curah hujan yang kurang dari batas rentang bawah, nilai tersebut akan diganti dengan batas bawah.

### 4.3 *Resampling Data*

Hal pertama yang perlu diperhatikan dalam menganalisis data runtun waktu adalah mengatur dan memastikan format tanggal dan urutan tanggal sudah benar. Dalam kasus *machine learning*, jumlah data yang besar sangat berpengaruh terhadap keakuratan model. Mengatasi hal tersebut dapat dilakukan dengan mengubah ukuran atau frekuensi data, yaitu dengan meningkatkan frekuensi data (*upsampling*) atau menurunkan frekuensi sampel data (*downsampling*). Pada penelitian ini menggunakan metode *downsampling* yaitu menurunkan frekuensi sample data curah hujan harian menjadi data curah hujan mingguan. *Downsampling* bertujuan untuk mengurangi fluktuasi harian yang tinggi, sehingga data lebih mudah untuk diinterpretasikan. Setelah dilakukan *downsampling*, jumlah data berubah menjadi sebanyak 366 data. Berikut adalah perbedaan plot antara data harian dan data setelah dilakukan proses *downsampling*:



**Gambar 4.3** a) Data dalam Harian, b) Data dalam Mingguan atau Setelah Proses *Downsampling*

#### 4.4 Normalisasi *Min-Max Scaling*

Teknik normalisasi data digunakan untuk memastikan bahwa variabel-variabel memiliki nilai dalam rentang yang sama, yaitu tidak terlalu besar maupun terlalu kecil, sehingga mempermudah proses analisis data. Tujuan dilakukan normalisasi data diantaranya dapat meningkatkan integritas data yang konsisten, menghindari ketidakteraturan data dan tujuan lainnya.

Normalisasi *min-max scaling* bekerja dengan menyesuaikan data dalam rentang nilai minimum hingga maksimum, yaitu dalam rentang 0 hingga 1. Berikut adalah contoh normalisasi *min-max scaling* pada atribut kelembaban udara rata-rata dan suhu udara rata-rata menggunakan persamaan (2.3):

Normalisasi *min-max scaling* atribut kelembaban udara rata-rata:

$$X_{scaled} = \frac{84,333183 - 63,428571}{90,857143 - 63,428571} = 0,762147$$

Normalisasi *min-max scaling* atribut suhu udara rata-rata:

$$X_{scaled} = \frac{24,136805 - 21,614286}{26,600000 - 21,614286} = 0,505949$$

Hasil normalisasi *min-max scaling* pada semua atribut selanjutnya dibantu dengan *software Rstudio* dapat dilihat pada lampiran 4.

## 4.5 Proses Pembangunan Model *Support Vector Regression* Algoritma *Grid Search*

### 4.5.1 Proses Pemodelan Curah Hujan tanpa Faktor yang Mempengaruhi

Proses pemodelan *Support Vector Regression* (SVR) dibangun dengan menggunakan bantuan *software R* yang membutuhkan *package 'e1071'* yang menyediakan fungsi *Support Vector Machine* (SVM). Selanjutnya data dibagi menjadi dua yaitu *data training* untuk pembangunan model dan *data testing* untuk mengukur tingkat akurasi dan kinerja model. Pembagian *data training* juga dapat mengoptimalkan kinerja dari *Support Vector Regression* (SVR) dalam menentukan parameter untuk pembangunan model prediksi terhadap *data testing*. Pembagian data dilakukan secara acak dengan menggunakan fungsi *set.seed (123)* dengan tujuan untuk memastikan bahwa proses pengacakan bersifat deterministik, artinya meskipun pembagian data secara acak hasilnya akan sama setiap kali kode tersebut dijalankan.

Langkah selanjutnya adalah pemilihan fitur untuk variabel dependen menggunakan curah hujan mingguan dan variabel independen menggunakan tanggal. Setelah itu, mengoptimasi parameter pada model *Support Vector Regression* (SVR) dengan menggunakan algoritma *grid search*. Pada penelitian ini menggunakan kernel *Radial Basis Function* (RBF) dengan *range* parameter *Cost* ( $C$ ) =  $2^{-5}$  hingga  $2^5$  dan *gamma* ( $\gamma$ ) =  $2^{-5}$  hingga  $2^5$ . Agar memperoleh *hyperparameter* maka diperlukan *cross validation* yang akan mengevaluasi model mana yang memiliki tingkat akurasi yang tinggi. Penelitian ini menggunakan *5-fold cross validation* karena *output* yang dihasilkan lebih baik diantara *cross validation* lainnya. Penentuan parameter optimal dilakukan dengan

melihat nilai kesalahan *error* terkecil pada pemodelan SVR, sebagai contoh diambil *data training* sebanyak 65% dengan menggunakan persamaan (2.25) dan persamaan (2.26) maka nilai:

$$MAE = \frac{19,24055}{240} = 0,080$$

$$RMSE = \sqrt{\frac{14,03133}{240}} = 0,242$$

Selanjutnya percobaan atau proses *trial and error* pembagian *data training* dan *data testing* beserta parameter optimal yang dihasilkan oleh algoritma *grid search* adalah:

**Tabel 4.3** Pembagian *Data Training* dan *Data Testing* pada Pemodelan SVR tanpa Faktor yang Mempengaruhi

<i>Data Training</i>	<i>Data testing</i>	RMSE	MAE	Kernel	<i>Cost</i>	<i>Epsilon</i>	<i>Gamma</i>
65%	35%	0,242	0,080	RBF	1	0	0,03125
70%	30%	0,219	0,078	RBF	1	0	0,03125
75%	25%	0,173	0,049	RBF	1	0	0,0625
80%	20%	0,155	0,041	RBF	1	0	0,0625
85%	15%	0,130	0,034	RBF	1	0	0,0625
90%	10%	0,120	0,029	RBF	1	0	0,0625
95%	5%	0,099	0,020	RBF	1	0	0,0625

Berdasarkan tabel 4.3 menunjukkan bahwa model dengan *data training* sebesar 95% dan *data testing* sebesar 5% merupakan model terbaik, karena nilai kesalahan *error* pada *data training* memiliki nilai terkecil diantara *data training* lainnya. Sedangkan *hyperparameter* yang dihasilkan dari algoritma *grid search* yaitu parameter *Cost* (*C*) sebesar 1, parameter *epsilon* ( $\epsilon$ ) sebesar 0 dan parameter *gamma* ( $\gamma$ ) sebesar 0,0625.

#### 4.5.2 Proses Pemodelan Curah Hujan dengan Memperhitungkan Faktor Meteorologi.

Sama halnya pada proses pemodelan curah hujan tanpa faktor yang memengaruhi, baik proses, kernel dan *range* parameter yang digunakan pada prediksi curah hujan dengan memperhitungkan faktor meteorologi yaitu menggunakan *range* parameter  $Cost (C) = 2^{-5}$  hingga  $2^5$  dan  $gamma (\gamma) = 2^{-5}$  hingga  $2^5$ , serta *cross validation* sebesar 5. Sedangkan pemilihan fitur pada kasus ini yaitu variabel dependen adalah curah hujan dan variabel independen terdiri dari: suhu rata-rata, kecepatan angin rata-rata, kelembaban angin rata-rata, dan lamanya penyinaran matahari. Percobaan atau proses *trial and error* pembagian *data training* dan *data testing* beserta parameter optimal yang dihasilkan oleh algoritma *grid search* dalam pembangunan model SVR pada prediksi curah hujan dengan mempertimbangkan faktor meteorologi adalah:

**Tabel 4.4** Pembagian *Data Training* dan *Data Testing* pada Pemodelan SVR dengan Memperhitungkan Faktor Meteorologi

<i>Data Training</i>	<i>Data testing</i>	RMSE	MAE	Kernel	<i>Cost</i>	<i>Epsilon</i>	<i>Gamma</i>
65%	35%	2,107	1,587	RBF	1	0,4	0,03125
70%	30%	2,129	1,585	RBF	1	0,4	0,03125
75%	25%	2,228	1,729	RBF	1	0,5	0,03125
80%	20%	2,038	1,607	RBF	1	0,5	0,125
85%	15%	2,077	1,576	RBF	1	0,3	0,0625
90%	10%	2,117	1,620	RBF	1	0,3	0,0625
95%	5%	2,150	1,634	RBF	1	0,3	0,0625

Berdasarkan tabel 4.4 peneliti memilih model terbaik diperoleh ketika *data training* sebesar 80% (294 data) dan *data testing* sebesar 20% sebanyak (72 data). Pembagian data ini dianggap optimal, karena memiliki nilai RMSE terkecil dibandingkan pembagian data lainnya. Nilai RMSE terkecil menunjukkan model memiliki kesalahan besar, sehingga lebih cocok digunakan pada data yang sensitif

terhadap *outlier*, di mana kesalahan besar perlu dihindari. Sedangkan parameter yang diperoleh dari proses algoritma *grid search* yaitu  $Cost = 1$ ,  $gamma = 0,125$ , dan  $epsilon = 0,5$ .

#### 4.6 Pemodelan *Support Vector Regression*

1. Pemodelan SVR pada Prediksi Curah Hujan tanpa Faktor yang Mempengaruhi

Setelah *tuning* atau penyetelan parameter *Support Vector Regression* menggunakan algoritma *grid search* dengan menggunakan kernel *Radial Basis Function* (RBF) diperoleh parameter terbaik prediksi curah hujan tanpa faktor meteorologi sebesar parameter  $Cost (C) = 1$ , menunjukkan keseimbangan yang baik antara kompleksitas model dengan kemampuan generalisasi, parameter  $epsilon (\epsilon) = 0$  mengidentifikasi tingkat toleransi kesalahan yang dapat diterima, dan parameter  $gamma (\gamma) = 0,0625$ , menunjukkan bahwa sebaran data tidak terlalu jauh dan model perlu memberikan perhatian serupa terhadap semua *data training*.

Parameter tersebut merupakan parameter terbaik karena hasil dari metrik evaluasinya paling kecil di antara parameter lainnya. Kemudian nilai parameter yang telah diperoleh digunakan mencari matriks kernel *Radial Basis Function* (RBF), pada penelitian ini diambil contoh sebanyak empat *data training* yaitu data ke-8 hingga data ke-11 dengan menggunakan persamaan (2.23) dan parameter  $\gamma = 0,0625$  sehingga diperoleh:

$$K(\mathbf{x}_1, \mathbf{x}_1) = e \left( -0,0625 \cdot \sum (0,476 - 0,476)^2 \right) = 1$$

$$K(\mathbf{x}_1, \mathbf{x}_2) = e \left( -0,0625 \cdot \sum (0,476 - 0,348)^2 \right) = 0,9990$$

$$K(\mathbf{x}_1, \mathbf{x}_3) = e\left(-0,0625 \cdot \sum (0,476 - 0,278)^2\right) = 0,9975$$

$$K(\mathbf{x}_1, \mathbf{x}_4) = e\left(-0,0625 \cdot \sum (0,476 - 0,412)^2\right) = 0,9997$$

Perhitungan pada data selanjutnya dilakukan dengan cara yang sama akan memperoleh kernel  $K$  dengan ukuran  $n \times n$ , dengan  $n$  adalah jumlah data.

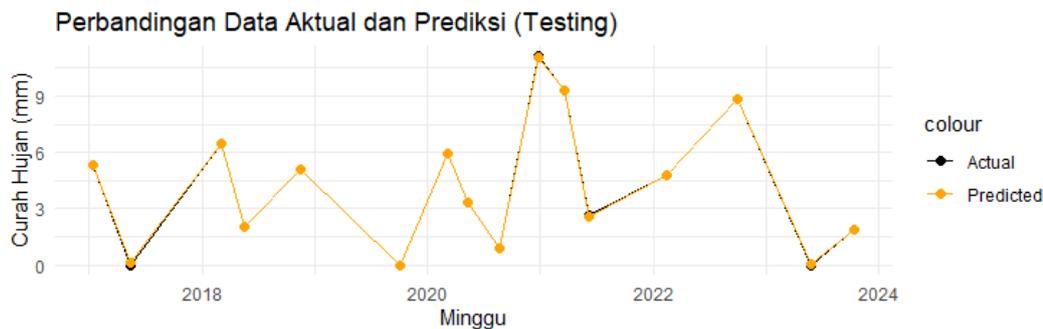
Adapun hasil yang diperoleh dari perhitungan kernel  $K$  adalah:

$$K = \begin{bmatrix} 1 & 0,9990 & 0,9975 & 0,9997 \\ 0,9990 & 1 & 0,9997 & 0,9997 \\ 0,9975 & 0,9997 & 1 & 0,9989 \\ 0,9997 & 0,9997 & 0,9989 & 1 \end{bmatrix}$$

Kemudian dengan bantuan *software R* diperoleh nilai koefisien *lagrange* yang digunakan untuk menentukan nilai  $w$  sebesar 28,32401. Nilai *weight* mencerminkan sensitivitas atau pengaruh *input* model curah hujan terhadap *output* berupa hasil prediksi. Nilai *weight* yang tinggi menunjukkan model SVR memberikan respon yang kuat terhadap *output*. Setelah memperoleh nilai  $w$  tersebut dengan bantuan *software R* didapatkan nilai bias sebesar  $-0,4258008$ . Nilai yang dihasilkan adalah negatif kecil, artinya fungsi prediksi sedikit bergeser ke bawah untuk mencocokkan nilai aktual curah hujan. Sehingga persamaan (2.17) model *Support Vector Regression* (SVR) pada prediksi curah tanpa faktor berpengaruh adalah:

$$y(\mathbf{x}) = \sum 28,32401 \cdot K(\mathbf{x}_i, \mathbf{x}_j) + (-0,4258008)$$

Berikut adalah plot perbandingan data aktual dan hasil prediksi curah hujan tanpa mempertimbangkan faktor yang mempengaruhi:



**Gambar 4.4** Plot Prediksi Curah Hujan tanpa Faktor Meteorologi

Gambar di atas menunjukkan garis dengan titik berwarna hitam sebagai data aktual curah hujan dan garis serta titik kuning sebagai data prediksi curah hujan dari tahun 2017 hingga tahun 2024. Garis prediksi (kuning) mengikuti pola garis aktual (hitam), artinya model memiliki kinerja yang baik dalam menangkap tren curah hujan aktual. Disamping itu, kedua garis memiliki fluktuasi curah hujan yang serupa dari tahun ke tahun, dengan beberapa periode curah hujan tinggi diikuti oleh periode curah hujan yang rendah.

## 2. Pemodelan SVR pada Prediksi Curah Hujan dengan Memperhitungkan Faktor Meteorologi

Pemodelan curah hujan dengan faktor meteorologi diperoleh parameter terbaik sebesar parameter  $Cost (C) = 1$ , menunjukkan keseimbangan yang baik antara kompleksitas model dan toleransi terhadap kesalahan, parameter  $epsilon (\epsilon) = 0,5$  mengidentifikasi tingkat toleransi kesalahan yang dapat diterima, dan parameter  $gamma (\gamma) = 0,125$ , menunjukkan bahwa sebaran data tidak terlalu jauh dan model perlu memberikan perhatian serupa terhadap semua *data training*. Parameter tersebut merupakan parameter terbaik karena hasil dari metrik evaluasinya paling kecil di antara parameter lainnya. Kemudian nilai parameter yang telah diperoleh digunakan mencari matriks kernel *Radial Basis Function (RBF)*. Adapun matriks kernel *Radial*

*Basis Function* (RBF) dengan menggunakan persamaan (2.23) pada data ke-21 hingga data ke -24 dengan parameter  $\gamma = 0,125$  sehingga diperoleh:

$$K(\mathbf{x}_1, \mathbf{x}_1) = e\left(-0,125 \cdot \sum((0,264 - 0,264)^2)\right) = 1$$

$$K(\mathbf{x}_1, \mathbf{x}_2) = e(-0,125 \cdot \sum((0,264 - 0,178)^2 + (0,526 - 0,526)^2 + (0,637 - 0,642)^2 + (0,516 - 0,645)^2)) = 0,9970$$

$$K(\mathbf{x}_1, \mathbf{x}_3) = e(-0,125 \cdot \sum((0,264 - 0,152)^2 + (0,526 - 0,552)^2 + (0,637 - 0,663)^2 + (0,516 - 0,645)^2)) = 0,9962$$

$$K(\mathbf{x}_1, \mathbf{x}_4) = e(-0,125 \cdot \sum((0,264 - 0,063)^2 + (0,526 - 0,615)^2 + (0,637 - 0,342)^2 + (0,516 - 0,968)^2)) = 0,9584$$

Perhitungan pada data selanjutnya dilakukan dengan cara yang sama akan memperoleh kernel K dengan ukuran  $n \times n$ , dengan n adalah jumlah data. Hasil yang diperoleh dari perhitungan kernel K adalah:

$$K = \begin{bmatrix} 1 & 0,9970 & 0,9962 & 0,9584 \\ 0,9970 & 1 & 0,9998 & 0,9734 \\ 0,9962 & 0,9998 & 1 & 0,9730 \\ 0,9584 & 0,9734 & 0,9730 & 1 \end{bmatrix}$$

Pemodelan *Support Vector Regression* (SVR) dengan bantuan *software R* diperoleh nilai koefisien *lagrange* yang akan digunakan untuk mencari nilai  $w$ , menghasilkan:

**Tabel 4.5** Nilai *Weight*

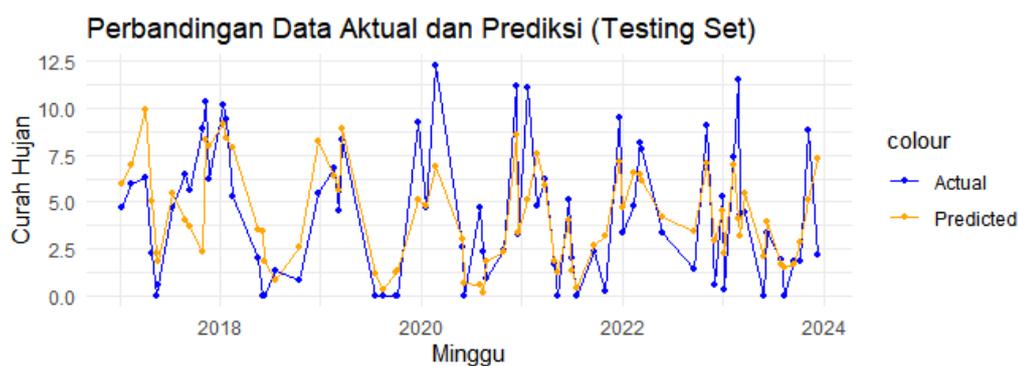
$w_i$	Nilai <i>weight</i>
$w_1$	3,724
$w_2$	6,700
$w_3$	-3,790
$w_4$	-2,412

Nilai *weight* pada suhu rata-rata sebesar 3,724, artinya memiliki kontribusi positif terhadap prediksi, yaitu cenderung meningkatkan nilai curah hujan yang

diprediksi. Nilai *weight* kelembaban rata-rata sebesar 6,700 memiliki kontribusi yang signifikan terhadap prediksi curah hujan, karena memiliki bobot tertinggi di antara semua faktor. Nilai *weight* kecepatan angin rata-rata dan lamanya penyinaran matahari memiliki nilai negatif, artinya jika kedua faktor memiliki nilai tinggi maka nilai model curah hujan cenderung rendah. Adapun nilai bias (*b*) dengan menggunakan perhitungan pada persamaan (2.16) diperoleh sebesar  $-0,3630018$ . Nilai bias yang dihasilkan menunjukkan model memerlukan penyesuaian pada prediksi curah hujan. Berdasarkan hasil perhitungan nilai *weight* (*w*) dan bias (*b*) maka diperoleh model *Support Vector Regression* (SVR) *nonlinear* dengan memperhitungkan faktor meteorologi menggunakan persamaan (2.17) adalah:

$$y(x) = \sum 3,724 + 6,700x_2 + (-3,790) + (-2,412x_4) - (-0,363)$$

Berikut plot perbandingan data aktual dan prediksi curah hujan dengan mempertimbangkan faktor meteorologi:



**Gambar 4.5** Plot Prediksi Curah Hujan dengan Memperhitungkan Faktor Meteorologi

Gambar di atas menunjukkan garis dan titik berwarna biru sebagai data aktual curah hujan, serta garis dan titik kuning sebagai data hasil prediksi curah hujan dari tahun 2017 hingga tahun 2023. Garis kuning mengikuti pola garis biru

dengan cukup baik, meskipun terdapat beberapa perbedaan pada puncak curah hujan. Hal tersebut menunjukkan bahwa model prediksi mampu menangkap pola curah hujan dengan fluktuasi yang serupa setiap tahunnya. Perbedaan yang terlihat pada puncak curah hujan kemungkinan disebabkan oleh faktor meteorologi tertentu yang tidak sepenuhnya terwakili dalam model prediksi.

#### 4.7 Analisis Evaluasi Metode *Support Vector Regression*

##### 1. Evaluasi Kesalahan Prediksi Curah Hujan tanpa Faktor Meteorologi

Sebelum menghitung evaluasi kesalahan *error* prediksi, data harus dinormalisasikan terlebih dahulu, yaitu mengembalikan data ke bentuk awal setelah dinormalisasi. Perhitungan denormalisasi data dengan menggunakan persamaan (2.4):

$$X_1 = 0,330(16,210 - 0) + 0 = 5,344$$

$$X_2 = 0,006(16,210 - 0) + 0 = 0,096$$

$$X_3 = 0,399(16,210 - 0) + 0 = 6,463$$

Hasil denormalisasi data selanjutnya dapat dilihat pada lampiran 5. Alat ukur metrik evaluasi yang digunakan pada penelitian ini yaitu *Mean Absolute Error* (MAE) dan *Root Mean Square Error* (RMSE). Berdasarkan Lampiran 5. Proses Perhitungan RMSE dan MAE Prediksi Curah Hujan tanpa Faktor yang Mempengaruhi dan Data Denormalisasi, perhitungan MAE menggunakan persamaan (2.25) dan RMSE dengan menggunakan persamaan (2.26) menunjukkan nilai kesalahan prediksi curah hujan tanpa faktor meteorologi:

$$MAE = \frac{1}{16} \times 0,271 = 0,02$$

$$RMSE = \sqrt{\frac{0,015}{16}} = 0,03$$

Nilai *Mean Absolute Error* (MAE) yang diperoleh sebesar 0,02 menunjukkan rata-rata selisih absolut antara nilai prediksi dengan nilai sebenarnya memiliki kesalahan prediksi yang kecil. Sedangkan nilai RMSE yang telah diperoleh yaitu sebesar 0,03 menunjukkan model memiliki nilai kesalahan prediksi yang kecil dibandingkan dengan model pada *data testing* lainnya. Hal tersebut mengindikasikan bahwa model mampu memprediksi curah hujan dengan akurasi yang akurat.

## 2. Evaluasi Kesalahan Prediksi Curah Hujan Berdasarkan Faktor Meteorologi

Proses denormalisasi data pada prediksi curah hujan dengan memperhitungkan faktor meteorologi menggunakan persamaan (2.4):

$$X_1 = 0,366(16,210 - 0) + 0 = 5,926$$

$$X_2 = 0,429(16,210 - 0) + 0 = 6,956$$

$$X_3 = 0,610(16,210 - 0) + 0 = 9,882$$

Adapun perhitungan *Mean Absolute Error* (MAE) dengan menggunakan persamaan (2.25) serta *Root Mean Square Error* (RMSE) menggunakan persamaan (2.26) untuk menentukan kesalahan *error* prediksi curah hujan berdasarkan faktor meteorologi yaitu:

$$MAE = \frac{1}{72} \times 129,670 = 1,801$$

$$RMSE = \sqrt{\frac{404,432}{72}} = 2,370$$

Berdasarkan hasil yang diperoleh, MAE sebesar 1,801 dan RMSE sebesar 2,370 menunjukkan bahwa kesalahan prediksi curah hujan dengan mempertimbangkan faktor meteorologi memiliki performa yang cukup baik.

#### 4.8 Analisis Tingkat Akurasi Metode *Support Vector Regression*

Uji akurasi model prediksi dilakukan menggunakan metode *R-squared* ( $R^2$ ) sebagai salah satu metrik utama untuk mengevaluasi performa model. *R-squared* merupakan ukuran statistik yang digunakan untuk menilai sejauh mana variabilitas data dapat dijelaskan oleh model. Pada prediksi curah hujan tanpa faktor meteorology, rata-rata nilai aktual yang dihasilkan sebesar 2,234 mm, sedangkan pada prediksi dengan memperhitungkan faktor meteorologi rata-rata nilai aktual sebesar 4,225 mm. Berdasarkan persamaan (2.27), nilai *R-squared* untuk kedua prediksi adalah:

1. Prediksi curah hujan tanpa faktor yang mempengaruhi:

$$R^2 = 1 - \frac{0,015}{183,777} = 1 - 0,0001 = 0,9999 \times 100\% = 99,99\%$$

2. Prediksi curah hujan dengan mempertimbangkan faktor meteorologi:

$$R^2 = 1 - \frac{404,432}{890,083} = 1 - 0,4544 = 0,5456 \times 100\% = 54,56\%$$

Nilai *R-squared* pada prediksi curah hujan tanpa faktor yang mempengaruhi sebesar 99,99% menunjukkan bahwa model yang digunakan sangat baik dalam menjelaskan variabilitas data curah hujan. Sedangkan nilai *R-squared* pada prediksi curah hujan dengan mempertimbangkan faktor meteorologi sebesar 54,56% memiliki performa yang lebih rendah dibandingkan prediksi tanpa faktor meteorologi. Walaupun *R-squared* lebih rendah, model ini mungkin lebih realistis karena curah hujan secara fisik dipengaruhi oleh parameter meteorologi.

#### 4.9 Kajian Penelitian dalam Perspektif Islam

Secara implisit dalam Q.S Yunus [10]:92 menjelaskan tentang prediksi Al-Qur'an terhadap kejadian di masa depan yaitu terjaganya jasad Fir'aun:

Artinya: *“Pada hari ini Kami selamatkan jasadmu agar kamu menjadi pelajaran bagi orang-orang yang datang setelah kamu. Sesungguhnya kebanyakan manusia benar-benar lengah (tidak mengindahkan) tanda-tanda (kekuasaan) Kami” (Al-Qur’an dan Terjemahannya, 2019).*

Ayat di atas adalah menjelaskan bahwa Fir’aun mengaku sebagai tuhan, kemudian pada hari kematiannya jenazah Fir’aun akan dikeluarkan dari dasar laut dan dilemparkan ke daratan agar mereka yang meragukan kematiannya menjadi yakin dan sebagai pelajaran bagi manusia sesudahnya. Bagaimana besar dan luasnya kekuasaan seseorang, jika ia menentang perintah-perintah Allah dan meninggalkan petunjuk-petunjuk Rasul-Nya, niscaya dia akan mengalami kehancuran. Ayat di atas memperingati kepada manusia selanjutnya bahwasanya barangsiapa yang menyekutukan, dan mendustakan Allah maka kelak akan menimpa azab atas apa yang telah diperbuat.

Allah SWT memberikan manusia akal, kebebasan berpikir, kemampuan untuk merencanakan, mengambil keputusan, dan bertindak sesuai dengan kemampuan dan pengetahuannya. Meskipun manusia memiliki kebebasan-kebebasan tersebut, hasil dari rencana tersebut ditentukan oleh Allah. Artinya keputusan akhir ada di tangan Allah dan segala sesuatu terjadi sesuai dengan ketentuan Allah yang terbaik untuk umat-Nya (*Al-Qur’an dan Terjemahannya, 2019*). Kemudian dalam Q.S an-Nahl ayat 97 menjelaskan tentang Allah menjanjikan kepada hamba-Nya akan memberikan kehidupan yang baik di dunia dan balasan pahala berlipat ganda di akhirat atas amal sholeh yang telah mereka kerjakan, yang berbunyi:

*“Barangsiapa yang mengerjakan kebajikan, baik dari laki-laki maupun perempuan, sedangkan ia seorang beriman, sungguh kami pasti akan berikan kepadanya kehidupan yang baik dan akan Kami beri balasan dengan pahala yang lebih baik daripada apa yang selalu mereka kerjakan” (Al-Qur’an dan Terjemahannya, 2019).*

Nabi Muhammad senantiasa mengingatkan umatnya betapa pentingnya waktu dan bagaimana seharusnya menggunakan waktu tersebut dengan semaksimal mungkin agar tidak berlalu begitu saja yang tercantum dalam hadist:

من كان يومه خيرامن أمسه فهو رابح. ومن كان يومه مثل أمسه فهو مغبون. و من كان يومه شرامن أمسه فهو ملعون. (رواه الحاكم)

*Barangsiapa hari ini lebih baik dari hari kemarin, dialah tergolong orang yang beruntung. Barangsiapa yang hari ini sama dengan hari kemarin dialah tergolong orang yang merugi dan barangsiapa yang hari ini lebih buruk dari hari kemarin dialah tergolong orang yang celaka” (HR. Al-Hakim).*

Hadist tersebut memberikan motivasi untuk selalu berusaha menjadi pribadi yang lebih baik dan tidak puas dengan kondisi saat ini. Selalu ada ruang untuk peningkatan dan pembelajaran. Selalu introspeksi diri, mengevaluasi perbuatan dan terus berusaha meningkatkan kualitas hidup secara keseluruhan (Muchlis, 2024).

## **BAB V**

### **PENUTUP**

#### **5.1 Kesimpulan**

Berdasarkan rumusan masalah dan pembahasan pada bab sebelumnya, maka dapat ditarik beberapa kesimpulan yaitu:

1. Tingkat akurasi penerapan metode *Support Vector Regression* (SVR) dengan algoritma *grid search optimization* dalam memprediksi curah hujan di Jawa Timur pada tahun 2017-2023 tanpa mempertimbangkan faktor meteorologi memiliki akurasi yang akurat. Sedangkan tingkat akurasi prediksi curah hujan di Jawa Timur pada tahun 2017-2023 dengan mempertimbangkan faktor meteorologi menggunakan metode yang sama memiliki tingkat akurasi yang moderat.
2. Metode *Support Vector Regression* (SVR) dengan algoritma *grid search optimization* tanpa mempertimbangkan faktor meteorologi memiliki tingkat akurasi yang lebih akurat dalam memprediksi curah hujan di Jawa Timur pada tahun 2017-2023.

#### **5.2 Saran**

Berdasarkan hasil analisis yang telah dilakukan, maka peneliti memiliki beberapa saran untuk penelitian selanjutnya untuk mengembangkan dan menyempurnakan penelitian ini sehingga menjadi lebih baik. Adapun saran tersebut adalah sebagai berikut:

1. Diharapkan untuk peneliti selanjutnya untuk menambah variasi atau mengoptimalkan parameter *Support Vector Regression* (SVR) agar mendapatkan hasil yang maksimal.
2. Peneliti selanjutnya diharapkan menggunakan algoritma optimasi lainnya seperti PSO dan IPSO untuk memperoleh model yang terbaik.
3. Peneliti selanjutnya diharapkan menambah dan mempertimbangkan faktor lain yang mempengaruhi curah hujan.
4. Pada penelitian berikutnya diharapkan menggunakan metode prediksi lainnya agar dapat diketahui perbedaan tingkat keakuratan antar model.

## DAFTAR PUSTAKA

- Al-Azizi, A. S. (2018). *Hadits-Hadits Sains* (Kamalulfik (ed.); 1 ed.). Yogyakarta: Laksana.
- Al-Qur'an dan Terjemahannya*. (2019). Kementrian Agama RI.
- Anggito, A., & Setiawan, J. (2018). *Metodologi Penelitian Kualitatif* (E. Lestari (Ed.); 1 ed.). Sukabumi: Jejak Publisher.
- Arisandi, R. R. R., & Hakim, A. R. (2022). Aplikasi Naive Bayes Classifier (NBC) pada Klasifikasi Status Gizi Balita Stunting dengan Pengujian K-Fold Cross Validation. *Jurnal Gaussian*, *11*(1), 130–139. <https://doi.org/https://doi.org/10.14710/j.gauss.v11i1.33991>
- Bayu, D. A. N. (2021). *Penerapan Klimatologi Dalam Pertanian* (1 ed.). Yogyakarta: Deepublish.
- Benhar, H., Idri, A., & Aleman, J. . F. (2020). Data preprocessing for heart disease classification: A systematic literature review. *Comput. Methods Programs Biomed*, *195*, 105635. <https://doi.org/10.1016/j.cmpb.2020.105635>
- BPS. (2022). *Banyaknya Desa/Kelurahan Menurut Jenis Bencana Alam dalam Tiga Tahun Terakhir (Desa), 2018*. BPS. <https://www.bps.go.id/indicator/168/954/1/banyaknya-desa-kelurahan-menurut-jenis-bencana-alam-dalam-tiga-tahun-terakhir.html>
- Chen, N., Lu, W., Yang, J., & Li, G. (2004). *Support Vector Machine in Chemistry*. Singapore: World Scientific Publishing.
- Dihni, V. A. (2022). *10 Negara Ini Memiliki Curah Hujan Tertinggi di Dunia, Ada Indonesia?* Kadata Media Network. <https://databoks.katadata.co.id/datapublish/2022/03/24/10-negara-ini-miliki-curah-hujan-tertinggi-di-dunia-ada-indonesia#:~:text=Rata-rata curah hujan di,untuk menyebabkan lebih banyak hujan.>
- Gandhi, R. (2018). *Support Vector Machine-Introduction to Machine Learning Algorithms*. Towards Data Science. <https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47>
- Hasanah, M. A., Soim, S., & Handayani, A. S. (2021). Implementasi CRISP-DM Model Menggunakan Metode Decision Tree dengan Algoritma CART untuk Prediksi Curah Hujan Berpotensi Banjir. *Journal of Applied Informatics and Computing*, *5*(2), 103–108. <https://doi.org/10.30871/jaic.v5i2.3200>
- Iswandi, U., & Indang, D. (2020). *Pengelolaan Sumber Daya Alam* (1 ed.). Yogyakarta: Deepublish.

- Khatimi, H., Alkaff, M., & Najipah, D. R. (2017). Penerapan Support Vector Regression (SVR) untuk Peramalan Inflasi Bulanan Nasional. *Jurnal Teknologi Informasi Universitas Lambung Mangkurat (JTIULM)* (2017), 02(2), 29–34. <https://doi.org/https://doi.org/10.20527/jtiulm.v2i2.21>
- Laia, M. L., & Setyawan, Y. (2020). Perbandingan Hasil Klasifikasi Curah Hujan Menggunakan Metode SVM dan NBC. *Jurnal Statistika Industri dan Komputasi*, 05(2), 51–61.
- Luz, H. (2009). *Knowledge Discovery With Support Vector Machines*. America: University of Rhode Island.
- Muchlis, H. (2024). *Hari ini Harus Lebih Baik dari Hari Kemarin dan Hari Esok Harus Lebih Baik dari Hari Ini*. Mahkamah Agung Republik Indonesia. <https://badilag.mahkamahagung.go.id/pojok-dirjen/pojok-dirjen-badilag/hari-ini-harus-lebih-baik-dari-hari-kemarin-dan-hari-esok-harus-lebih-baik-dari-hari-ini-oleh-drs-h-muchlis-s-h-m-h-27-09>
- Muflikhah, L., Widodo, Mahmudy, F. W., & Solimun. (2021). *Machine Learning dalam Bioinformatika*. Malang: UB Press.
- Mukid, M. A., & Sugito. (2013). Model Prediksi Curah Hujan Dengan Pendekatan Regresi Proses Gaussian (Studi Kasus di Kabupaten Grobogan). *Medika Statistika*, 6, 113–122.
- Mustajab, R. (2023). *BNPB: Indonesia Alami 3.522 Bencana Alam pada 2022*. Data Indonesia. <https://dataindonesia.id/varia/detail/bnpb-indonesia-alami-3522-bencana-alam-pada-2022>
- Mustakim, Buono, A., & Irman, H. (2015). Support Vector Regression untuk Prediksi Produktivitas Kelapa Sawit di Provinsi Riau. *Jurnal Sains, Teknologi, dan Industri*, 12(2), 179–188.
- Mustika, Ardilla, Y., Manuhutu, A., Ahmad, N., & Hasbi, I. (2021). *Data Mining dan Aplikasinya* (N. Rismawati (Ed.)). Bandung: Widina Bhakti Persada Bandung.
- Putra, R. R., Isa, I. G., & Malyan, A. B. (2023). *Buku Ajar Pengantar Deep Learning dalam Pemrosesan Citra* (M. Nasrudin (Ed.)). Pekalongan: PT. Nasya Expanding Management.
- Ramadhan, R. F., Saputri, F. R., Jarwo, & Risanto, D. (2023). *Kecerdasan Buatan Digital* (A. Yanto (Ed.)). Padang: Global Eksekutif Teknologi.
- Restu, H., Saputra, M. I., Triyono, A., & Suwaji. (2021). *Metode Penelitian* (D. Amidasti (Ed.)). Yogyakarta: Deepublish Publisher.
- Riyanarto, S., Shoffi, S. I., Malikhah, Doni, P. P., & Syauqi, A. H. (2023). *Machine*

*Learning Deep Learning Konsep dan Pemograman Python Studi Kasus Jurnal Bereputasi Scopus dan Berimpact Factor Web of Science*. Yogyakarta: Penerbit Andi.

- Safira, A. N., Warsito, B., & Rusgiyono, A. (2022). Analisis Support Vector Regression (SVR) dengan Algoritma Grid Search Time Series Cross Validation untuk Prediksi Jumlah Kasus Terkonfirmasi COVID-19 di Indonesia. *Jurnal Gaussian*, 11(4), 512–521. <https://doi.org/https://ejournal3.undip.ac.id/index.php/gaussian/>
- Santoso, S. (2018). *Mahir Statistik Multivariat dengan SPSS*. Jakarta: PT. Alex Media Komputindo.
- Siagian, D., & Sugiarto. (2000). *Metode Statistika untuk Bisnis dan Ekonomi*. Jakarta: Penerbit PT Gramedia Pustaka Utama.
- Sigit, N. (2008). *Dasar-Dasar Metode Statistik* (1 ed.). Jakarta: Grasindo.
- Sulianta, F. (2023). *Basic Data Mining from A to Z Dasar Membangun Tindakan Bisnis*. Bandung: Universitas Widyatama.
- Suparwito, H., Gunawan, R., & Binanto, I. (2023). *Pengantar Pembelajaran Mesin Menggunakan Bahasa Pemograman Python* (A. R. Widiarti (Ed.)). Yogyakarta: Sanata Dharma University Press.
- Suroyo, K. N. (2022). *Implementasi Metode SVR (Support Vector Regression) untuk Prediksi Jumlah Positif COVID-19 di Jawa Timur*. Universitas Islam Negeri Sunan Ampel Surabaya.
- Trilaksono, B. R., Riza, H., & Jarin, A. (2023). *Prosiding Use Cases Artificial Intelligence Indonesia: Embracing Collaboration for Research and Industrial Innovation in Artificial Intelligence*. Jakarta: Badan Riset dan Inovasi Nasional.
- Vapnik, V. N. (2000). *The Nature of Statistical Learning Theory* (2nd ed.). New York: Springer-Verlag.
- Wadi, H. (2023). *Prediksi Harga Emas dengan Jaringan Syaraf Tiruan Backpropagation menggunakan MATLAB GUI*. Mataram: Turida Publisher.
- Werdiningsih, I., Novitasari, D. C. R., & Haq, D. Z. (2022). *Pengelolaan Data Mining dengan Pemograman Matlab*. Surabaya: Airlangga University Press.
- Yudistira, N., Alfiansih, L. M., Andriyani, N. I., & Maulida, N. (2023). *Prediksi Deret Waktu Menggunakan Deep Learning* (1 ed.). Malang: UB Press.
- Yunardi, T. R., & Dina, Z. N. (2022). *Data Mining dan Machine Learning dengan Orange3 Tutorial dan Aplikasinya* (Abadi Anas (Ed.)). Surabaya: Airlangga

University Press.

## LAMPIRAN

**Lampiran 1.** Data Curah Hujan Harian Jawa Timur pada Tahun 2017-2023

No	Tanggal	Suhu udara rata-rata (°C)	Kelembaban udara rata-rata (%)	Lamanya penyinaran matahari (jam)	Kecepatan angin rata-rata (km/h)	Curah hujan (mm)
1	01-01-2017	24.4	87	0	2	
2	02-01-2017	24.2	85	4.6	1	1
3	03-01-2017	23.3	90	2.5	2	0
4	04-01-2017	24.7	84	3.3	1	5.3
5	05-01-2017	24.3	80	6	1	3.6
6	06-01-2017	24	86	3.5	1	17.6
7	07-01-2017			3	1	
8	08-01-2017	24.8	83	2.6	2	1.1
9	09-01-2017	24	82	7.8	2	0
10	10-01-2017	25.2	79	6	1	1.3
11	11-01-2017	24.9	82	6.3	1	10
12	12-01-2017	23.7	88	7.2	1	42.8
13	13-01-2017	24.6	81	3.7	1	1.8
14	14-01-2017	23.8	87	3.3	2	2.2
15	15-01-2017	22.7	90	3	1	15.8
16	16-01-2017	23	89	0	1	1.8
17	17-01-2017	23.1	91	0.8	1	0.7
18	18-01-2017	23.9	85	3.5	1	47.8
19	19-01-2017			2.7	1	1.7
20	20-01-2017	24	84	4	2	0.2
21	21-01-2017	23.7	87	1.5	1	1
22	22-01-2017	23.4	91	4.5	1	47
23	23-01-2017	22.5	91	4.4	1	5.2
24	24-01-2017	23.6	84	2	2	10.4
25	25-01-2017	23.3	90	5.6	1	68.2
26	26-01-2017	23.2	86	1	2	7.5
27	27-01-2017	24.7	80	0.6	2	20.7
28	28-01-2017	23.9	82	8.2	1	14.3
29	29-01-2017	23.6	85	3.4	2	12.8
30	30-01-2017	23.6	86	0	1	15.7

No	Tanggal	Suhu udara rata-rata (°C)	Kelembaban udara rata-rata (%)	Lamanya penyinaran matahari (jam)	Kecepatan angin rata-rata (m/s)	Curah hujan (mm)
31	31-01-2017			2.3	1	16.8
32	01-02-2017	23.1	89	0	2	5.4
33	02-02-2017	22.1	93	8	1	40.8
34	03-02-2017	22.6	90	0	2	14.5
35	04-02-2017	23.9	81	0.6	2	11.5
36	05-02-2017	25.5	64	3.7	3	
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
2550	25-12-2023	25,6	74	8.6	1	0
2551	26-12-2023	25.5	82	4	1	13
2552	27-12-2023	25,9	78	5.5	1	25.3
2553	28-12-2023	25,4	86	4.5	1	19.3
2554	29-12-2023	27.1	79	5.4	2	27.5
2555	30-12-2023	26,1	82	8.7	2	1.4
2556	31-12-2023	26,6	81	4	2	34

**Lampiran 2. Data Cleaning dari Data Missing**

No	Tanggal	Suhu udara rata-rata (°C)	Kelembaban udara rata-rata (%)	Lamanya penyinaran matahari (jam)	Kecepatan angin rata-rata (m/s)	Curah hujan (mm)
1	01-01-2017	24.40	87.00	0.00	2.00	6.48
2	02-01-2017	24.20	85.00	4.60	1.00	1.00
3	03-01-2017	23.30	90.00	2.50	2.00	0.00
4	04-01-2017	24.70	84.00	3.30	1.00	5.30
5	05-01-2017	24.30	80.00	6.00	1.00	3.60
6	06-01-2017	24.00	86.00	3.50	1.00	16.21
7	07-01-2017	24.06	78.33	3.00	1.00	6.48
8	08-01-2017	24.80	83.00	2.60	2.00	1.10
9	09-01-2017	24.00	82.00	7.80	2.00	0.00
10	10-01-2017	25.20	79.00	6.00	1.00	1.30
11	11-01-2017	24.90	82,00	6.30	1.00	10.00
12	12-01-2017	23.70	88.00	7.20	1.00	16.21

No	Tanggal	Suhu udara rata-rata (°C)	Kelembaban udara rata-rata (%)	Lamanya penyinaran matahari (jam)	Kecepatan angin rata-rata (m/s)	Curah hujan (mm)
13	13-01-2017	24.60	81.00	3.70	1.00	1.80
14	14-01-2017	23.80	87.00	3.30	2.00	2.20
15	15-01-2017	22.70	90.00	3.00	1.00	15.80
16	16-01-2017	23.00	89.00	0.00	1.00	1.80
17	17-01-2017	23.10	91.00	0.80	1.00	0.70
18	18-01-2017	23.90	85.00	3.50	1.00	16.21
19	19-01-2017	24.06	78.33	2.70	1.00	1.70
20	20-01-2017	24.00	84.00	4.00	2.00	0.20
21	21-01-2017	23.70	87.00	1.50	1.00	1.00
22	22-01-2017	23.40	91.00	4.50	1.00	16.21
23	23-01-2017	22.50	91.00	4.40	1.00	5.20
24	24-01-2017	23.60	84.00	2.00	2.00	10.40
25	25-01-2017	23.30	90.00	5.60	1.00	16.21
26	26-01-2017	23.20	86.00	1.00	2.00	7.50
27	27-01-2017	24.70	80.00	0.60	2.00	16.21
28	28-01-2017	23.90	82.00	8.20	1.00	14.30
29	29-01-2017	23.60	85.00	3.40	2.00	12.80
30	30-01-2017	23.60	86.00	0.00	1.00	15.70
31	31-01-2017	24.06	78.33	2.30	1.00	16.21
32	01-02-2017	23.10	89.00	0.00	2.00	5.40
33	02-02-2017	22.10	93.00	8.00	1.00	16.21
34	03-02-2017	22.60	90.00	0.00	2.00	14.50
35	04-02-2017	23.90	81.00	0.60	2.00	11.50
36	05-02-2017	25.50	64.00	3.70	3.00	6.48
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
2550	25-12-2023	25.60	74.00	8.60	1.00	0.00
2551	26-12-2023	25.50	82.00	4.00	1.00	13.00
2552	27-12-2023	25.90	78.00	5.50	1.00	16.21
2553	28-12-2023	25.40	86.00	4.50	1.00	16.21
2554	29-12-2023	26.90	79.00	5.40	2.00	16.21
2555	30-12-2023	26.10	82.00	8.70	2.00	1.40
2556	31-12-2023	26.60	81.00	4.00	2.00	16.21

**Lampiran 3.** Data Curah Hujan Setelah Proses *Downsampling*

No	Tanggal	Suhu udara rata-rata (°C)	Kelembaban udara rata-rata (%)	Lamanya penyinaran matahari (jam)	Kecepatan angin rata-rata (m/s)	Curah hujan (mm)
1	01-01-2017	24.14	84.33	3.27	1.26	5.58
2	08-01-2017	24.43	83.24	5.27	1.43	4.66
3	15-01-2017	23.49	86.35	2.21	1.43	5.34
4	22-01-2017	23.51	86.28	3.76	1.43	12.29
5	29-01-2017	23.27	86.06	2.04	1.57	13.19
6	05-02-2017	25.00	74.14	5.54	2.86	7.06
7	12-02-2017	23.91	84.86	4.68	1.71	5.94
8	19-02-2017	24.26	84.71	4.93	1.71	4.82
9	26-02-2017	24.20	82.57	5.77	1.71	7.72
10	05-03-2017	24.10	79.28	5.81	1.28	5.64
11	12-03-2017	24.27	84.00	5.58	1.57	4.50
12	19-03-2017	24.16	82.35	3.20	1.71	6.67
13	26-03-2017	23.48	84.43	4.13	1.28	10.10
14	02-04-2017	23.31	89.28	3.24	1.14	6.30
15	09-04-2017	24.50	79.24	4.81	1.71	5.03
16	16-04-2017	24.62	81.71	7.07	1.42	3.87
17	23-04-2017	24.48	80.57	6.93	1.57	4.10
18	30-04-2017	24.60	82.14	6.13	1.57	2.25
19	07-05-2017	24.13	75.23	6.68	1.71	1.17
20	14-05-2017	24.37	71.14	7.98	2.00	0.00
21	21-05-2017	23.84	73.00	8.23	1.86	0.53
22	28-05-2017	24.04	83.43	3.56	2.00	3.99
23	04-06-2017	23.72	77.57	6.61	1.57	1.64
24	11-06-2017	23.96	80.43	6.08	1.71	0.99
25	18-06-2017	23.40	74.57	7.13	2.21	0.09
.	....	....	....	....	....	....
.	....	....	....	....	....	....
.	....	....	....	....	....	....
362	03-12-2023	25.76	79.57	5.51	1.43	9.09
363	10-12-2023	25.43	76.43	4.73	1.71	2.19
364	17-12-2023	26.29	66.71	9.44	2.29	3.88
365	24-12-2023	25.04	77.43	6.56	1.43	9.93
366	31-12-2023	26.60	81	4	2	16.21

**Lampiran 4.** Normalisasi Data *Min-max Scaling*

No	Tanggal	Suhu udara rata-rata (°C)	Kelembaban udara rata-rata (%)	Lamanya penyinaran matahari (jam)	Kecepatan angin rata-rata (m/s)	Curah hujan (mm)
1	01-01-2017	0.51	0.76	0.27	0.27	0.34
2	08-01-2017	0.57	0.72	0.49	0.33	0.29
3	15-01-2017	0.02	0.93	0.00	0.00	0.75
4	22-01-2017	0.39	0.83	0.32	0.33	0.81
5	29-01-2017	0.34	0.82	0.13	0.40	0.44
6	05-02-2017	0.68	0.39	0.52	1.00	0.37
7	12-02-2017	0.46	0.78	0.43	0.47	0.30
8	19-02-2017	0.53	0.78	0.45	0.47	0.48
9	26-02-2017	0.52	0.70	0.55	0.47	0.35
10	05-03-2017	0.50	0.58	0.55	0.27	0.28
11	12-03-2017	0.53	0.75	0.53	0.40	0.41
12	19-03-2017	0.51	0.69	0.26	0.47	0.62
.	....	....	....	....	....	
.	....	....	....	....	....	
.	....	....	....	....	....	
346	03-12-2023	0.83	0.59	0.52	0.34	0.56
347	10-12-2023	0.85	0.47	0.43	0.47	0.13
348	17-12-2023	0.94	0.12	0.97	0.74	0.24
349	24-12-2023	0.89	0.51	0.64	0.34	0.61
350	31-12-2023	1.00	0.64	0.35	0.60	1

**Lampiran 5.** Proses Perhitungan RMSE dan MAE Prediksi Curah Hujan tanpa Faktor yang Mempengaruhi dan Data Denormalisasi

$n$	$y$	$\hat{y}$	$y - \hat{y}$	$(y - \hat{y})^2$	$ y - \hat{y} $	$(y_i - \bar{y})^2$
1	5,344	5,345	-0,001	0,000	0,001	1,233
2	0,000	0,096	-0,096	0,009	0,096	17,927
3	6,459	6,463	-0,005	0,000	0,005	4,949
4	2,029	2,013	0,016	0,000	0,016	4,864
5	5,132	5,130	0,002	0,000	0,002	0,806
6	0,000	0,001	-0,001	0,000	0,001	17,927
7	5,944	5,948	-0,004	0,000	0,004	2,925
8	3,314	3,305	0,008	0,000	0,008	0,847

$n$	$y$	$\hat{y}$	$y - \hat{y}$	$(y - \hat{y})^2$	$ y - \hat{y} $	$(y_i - \bar{y})^2$
9	0,926	0,909	0,017	0,000	0,017	10,941
10	11,118	11,083	0,035	0,001	0,035	47,389
11	9,342	9,344	-0,002	0,000	0,002	26,092
12	2,630	2,616	0,014	0,000	0,014	2,573
13	4,787	4,786	0,001	0,000	0,001	0,306
14	8,860	8,856	0,004	0,000	0,004	21,401
15	0,000	0,061	-0,061	0,004	0,061	17,927
16	1,853	1,858	-0,005	0,000	0,005	5,671
Jumlah			-0,079	0,015	0,017	184,777

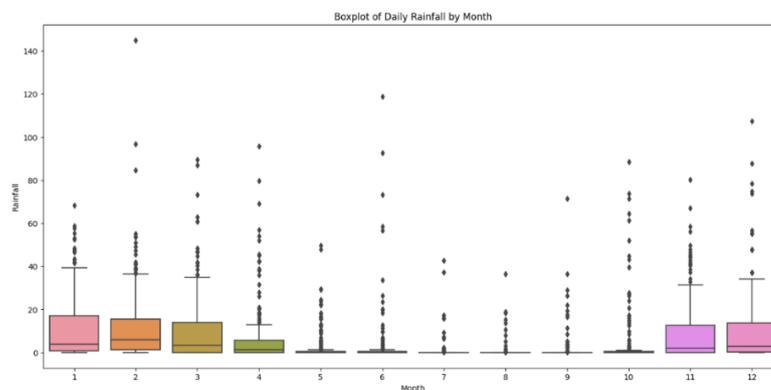
**Lampiran 6.** Proses Perhitungan RMSE dan MAE Prediksi Curah Hujan dengan Memperhitungkan Faktor Meteorologi

$n$	$y$	$\hat{y}$	$y - \hat{y}$	$(y - \hat{y})^2$	$ y - \hat{y} $	$(y_i - \bar{y})^2$
1	4,659	5,926	-1,268	1,607	1,268	0,188
2	5,944	6,956	-1,012	1,023	1,012	2,956
3	6,303	9,882	-3,579	12,813	3,579	4,318
4	2,255	5,020	-2,765	7,645	2,765	3,881
5	0,000	2,246	-2,246	5,044	2,246	17,851
6	0,529	1,861	-1,333	1,776	1,333	13,664
7	4,732	5,470	-0,738	0,545	0,738	0,257
8	6,484	4,008	2,476	6,133	2,476	5,104
9	5,615	3,652	1,963	3,852	1,963	1,932
10	8,901	2,375	6,525	42,579	6,525	21,862
11	10,346	8,286	2,060	4,243	2,060	37,465
12	6,228	7,969	-1,741	3,032	1,741	4,011
13	10,146	9,156	0,990	0,980	0,990	35,056
14	9,428	8,408	1,020	1,040	1,020	27,069
15	5,256	7,932	-2,676	7,161	2,676	1,064
16	2,029	3,548	-1,520	2,310	1,520	4,824
17	0,000	3,426	-3,426	11,737	3,426	17,851
18	0,000	1,830	-1,830	3,347	1,830	17,851
19	1,329	0,851	0,477	0,228	0,477	8,389
20	0,857	2,592	-1,734	3,008	1,734	11,342
21	5,428	8,265	-2,837	8,048	2,837	1,447
22	6,814	6,348	0,465	0,217	0,465	6,700
23	4,559	5,636	-1,077	1,161	1,077	0,111
24	8,303	8,926	-0,623	0,389	0,623	16,630
25	0,000	1,179	-1,179	1,390	1,179	17,851

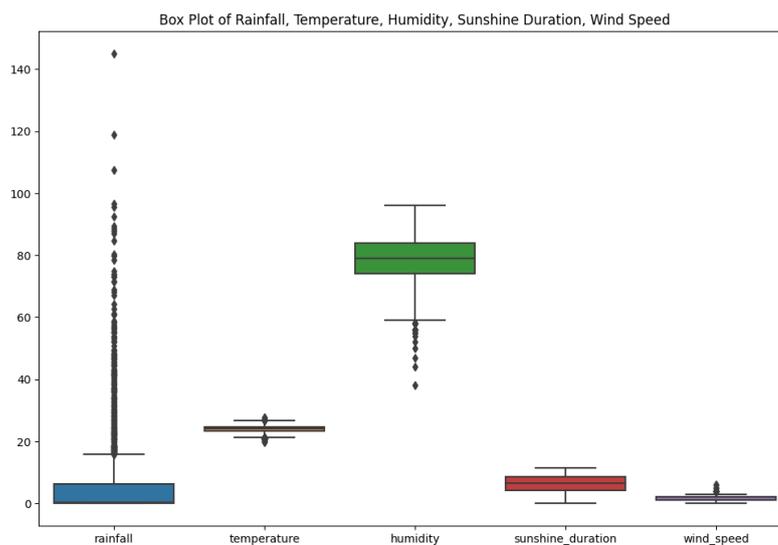
$n$	$y$	$\hat{y}$	$y - \hat{y}$	$(y - \hat{y})^2$	$ y - \hat{y} $	$(y_i - \bar{y})^2$
26	0,000	0,300	-0,300	0,090	0,300	17,851
27	0,000	1,280	-1,280	1,637	1,280	17,851
28	0,000	1,315	-1,315	1,729	1,315	17,851
29	9,217	5,075	4,142	17,157	4,142	24,923
30	4,659	4,784	-0,125	0,016	0,125	0,188
31	12,247	6,918	5,329	28,401	5,329	64,357
32	2,601	2,997	-0,395	0,156	0,395	2,636
33	0,000	0,671	-0,671	0,450	0,671	17,851
34	4,689	0,604	4,085	16,688	4,085	0,215
35	2,316	0,174	2,142	4,587	2,142	3,645
36	0,926	1,807	-0,881	0,776	0,881	10,881
37	2,444	2,337	0,107	0,012	0,107	3,171
38	11,189	8,553	2,637	6,953	2,637	48,503
39	3,242	3,341	-0,099	0,010	0,099	0,966
40	11,119	5,155	5,964	35,566	5,964	47,524
41	4,817	7,527	-2,709	7,340	2,709	0,351
42	6,187	5,878	0,309	0,096	0,309	3,850
43	1,629	1,846	-0,218	0,047	0,218	6,741
44	0,000	1,240	-1,240	1,538	1,240	17,851
45	5,144	3,989	1,155	1,334	1,155	0,845
46	1,957	1,320	0,637	0,406	0,637	5,143
47	0,000	0,362	-0,362	0,131	0,362	17,851
48	2,344	2,639	-0,294	0,087	0,294	3,537
49	0,243	3,136	-2,893	8,368	2,893	15,857
50	9,486	7,129	2,358	5,559	2,358	27,683
51	3,324	4,693	-1,369	1,874	1,369	0,812
52	4,787	6,565	-1,778	3,161	1,778	0,316
53	8,131	6,428	1,703	2,901	1,703	15,255
54	7,774	6,138	1,636	2,678	1,636	12,598
55	3,324	4,197	-0,873	0,762	0,873	0,812
56	1,386	3,399	-2,013	4,051	2,013	8,062
57	9,073	7,090	1,983	3,931	1,983	23,502
58	0,571	2,909	-2,337	5,463	2,337	13,349
59	5,301	4,552	0,750	0,562	0,750	1,159
60	0,329	2,234	-1,906	3,631	1,906	15,182
61	7,373	7,001	0,372	0,139	0,372	9,909
62	11,520	4,076	7,444	55,413	7,444	53,220
63	4,259	3,154	1,104	1,219	1,104	0,001

$n$	$y$	$\hat{y}$	$y - \hat{y}$	$(y - \hat{y})^2$	$ y - \hat{y} $	$(y_i - \bar{y})^2$
64	4,459	5,419	-0,960	0,922	0,960	0,055
65	0,000	2,042	-2,042	4,170	2,042	17,851
66	3,385	3,909	-0,524	0,274	0,524	0,706
67	1,895	1,689	0,206	0,043	0,206	5,427
68	0,014	1,522	-1,507	2,272	1,507	17,730
69	1,853	1,650	0,202	0,041	0,202	5,628
70	1,853	2,838	-0,985	0,970	0,985	5,628
71	8,800	5,147	3,653	13,345	3,653	20,930
72	2,186	7,302	-5,116	26,173	5,116	4,159
Jumlah			-1,881	404,432	129,670	890,083

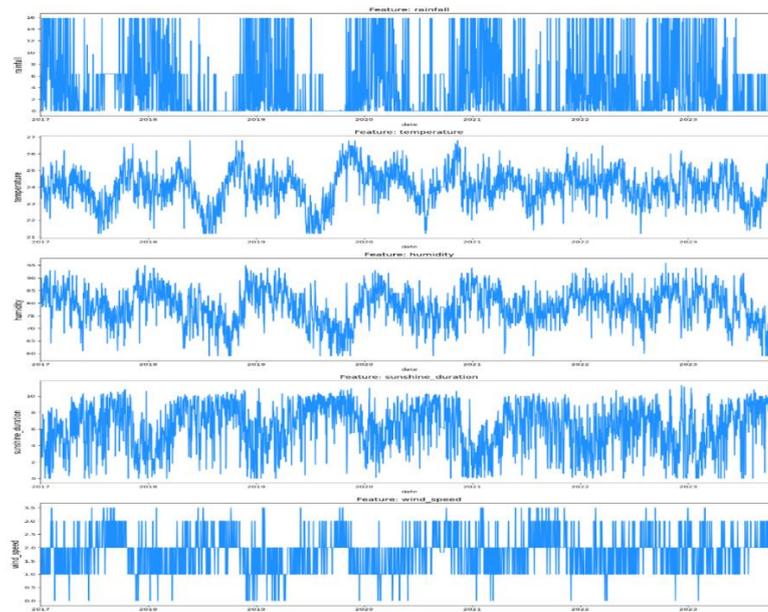
**Lampiran 7.** *Boxplot* Curah Hujan Harian dalam Akurasi Bulanan



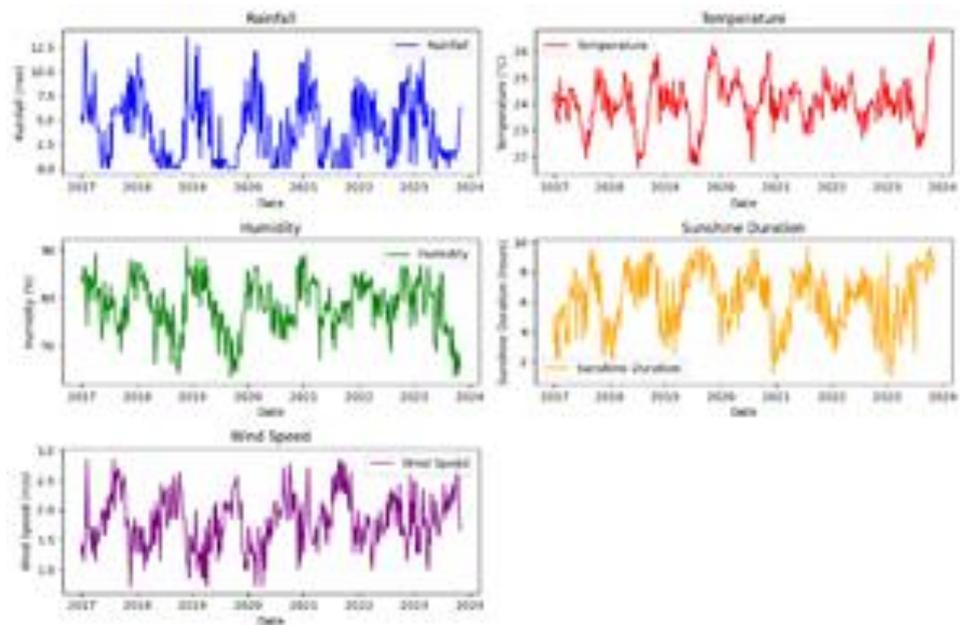
**Lampiran 8.** *Boxplot* Mendeteksi Adanya *Outliers*



### Lampiran 9. Scatterplot Data Curah Hujan Harian Jawa Timur



### Lampiran 10. Scatterplot Data Curah Hujan Mingguan Jawa Timur



**Lampiran 11.** *Syntax* Model SVR pada Prediksi Curah Hujan Menggunakan R

```
# Memuat library yang diperlukan
library(dplyr)
library(caret)
library(e1071)
library(Metrics)
library(ggplot2)
library(tidyr)
library(readxl)
library(data.table)
library(lubridate)
library(writexl)
# Memuat data
data <- read_excel("Bismillah SkripSweet/Data/datahujan23.xlsx")
View(data)
# Analisis Deskriptif
summary(data)
# Ubah kolom 'date' menjadi tipe data Date
data$date <- as.Date(data$date)
# Ekstrak bulan dari kolom 'date'
data$month <- format(data$date, "%m")
#Buat boxplot data curah hujan harian menjadi perbulan
ggplot(data, aes(x = month, y = rainfall)) +
  geom_boxplot() +
  ggtitle("Boxplot of Daily Rainfall by Month") +
  xlab("Month") +
  ylab("Rainfall")
#Hitung rata-rata curah hujan tiap bulan, bandingkan dengan rata-rata
akumulasi curah hujan dan tampilkan perbandingannya
monthly_mean_rainfall <- data %>%
  group_by(month) %>%
  summarise(monthly_mean_rainfall = mean(rainfall, na.rm = TRUE))
total_mean_rainfall <- mean(data$rainfall, na.rm = TRUE)
comparison <- monthly_mean_rainfall$monthly_mean_rainfall >
total_mean_rainfall
cat("Rata-rata Akumulasi Curah Hujan Secara Keseluruhan: ",
round(total_mean_rainfall, 2), "mm\n")
cat("\nPerbandingan Rata-rata Curah Hujan Harian Tiap Bulan:\n")
print(monthly_mean_rainfall)
cat("\nLebih Besar dari Rata-rata Keseluruhan:\n")
print(monthly_mean_rainfall[comparison, ])
cat("\nLebih Kecil atau Sama dengan Rata-rata Keseluruhan:\n")
print(monthly_mean_rainfall[!comparison, ])
# definisikan variabel
vars <- c('rainfall', 'temperature', 'humidity', 'sunshine_duration',
'wind_speed')
```

```

# Melting the data to long format
data_long <- reshape2::melt(data, id.vars = "date", measure.vars = vars)
# Buat box plot menggunakan ggplot2
ggplot(data_long, aes(x = variable, y = value)) +
  geom_boxplot() +
  ggtitle('Box Plot of Rainfall, Temperature, Humidity, Sunshine Duration,
Wind Speed') +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotate x-axis
labels if needed
# Ubah kolom date ke tipe tanggal
data$date <- as.Date(data$date, format="%Y-%m-%d")
# Data Cleaning: Menghapus data yang hilang dan outlier
# mengisi data hilang dengan nilai mean masing2 variabel
data$rainfall[is.na(data$rainfall)] <- mean(data$rainfall, na.rm = TRUE)
data$temperature[is.na(data$temperature)] = mean(data$temperature, na.rm =
TRUE)
data$humidity[is.na(data$humidity)] = mean(data$humidity, na.rm = TRUE)
data$wind_speed[is.na(data$wind_speed)] = mean(data$wind_speed, na.rm =
TRUE)
data$sunshine_duration[is.na(data$sunshine_duration)] = mean(data$sunshine
_duration, na.rm = TRUE)
# Menangani outliers dengan IQR untuk setiap variable
# Fungsi untuk mendeteksi dan menangani outlier dengan IQR
handle_outliers <- function(x) {
  Q1 <- quantile(x, 0.25, na.rm = TRUE)
  Q3 <- quantile(x, 0.75, na.rm = TRUE)
  IQR <- Q3 - Q1
  lower_bound <- Q1 - 1.5 * IQR
  upper_bound <- Q3 + 1.5 * IQR
  x <- ifelse(x < lower_bound, lower_bound, x)
  x <- ifelse(x > upper_bound, upper_bound, x)
  return(x)
}
# Terapkan fungsi untuk setiap variabel
data$rainfall <- handle_outliers(data$rainfall)
data$humidity <- handle_outliers(data$humidity)
data$temperature <- handle_outliers(data$temperature)
data$wind_speed <- handle_outliers(data$wind_speed)
data$sunshine_duration <- handle_outliers(data$sunshine_duration)
# Resampling data harian menjadi mingguan
data_weekly <- data %>%
  mutate(week = floor_date(date, "week")) %>%
  group_by(week) %>%
  summarise(
    rainfall = mean(rainfall, na.rm = TRUE),
    humidity = mean(humidity, na.rm = TRUE),
    temperature = mean(temperature, na.rm = TRUE),

```

```

sunshine_duration = mean(sunshine_duration, na.rm = TRUE),
wind_speed = mean(wind_speed, na.rm = TRUE)
)
View(data_weekly)
# Normalisasi Min-Max Scaling
normalize <- function(x) (x - min(x)) / (max(x) - min(x))
data_normalized <- data_weekly %>%
  mutate(across(-week, normalize))
# Pembagian data menjadi data training dan testing
set.seed(123) # memastikan pengacakan yang sama
trainIndex <- createDataPartition(data_normalized$rainfall, p = .8, list =
FALSE)
dataTrain <- data_normalized[ trainIndex,]
dataTest <- data_normalized[-trainIndex,]
# Ekstrak variabel input dan output
x_train <- as.data.frame(dataTrain[, !colnames(dataTrain) %in%
c("rainfall")])
y_train <- dataTrain$rainfall
x_test <- as.data.frame(dataTest[, !colnames(dataTest) %in% c("rainfall")])
y_test <- dataTest$rainfall
# Optimasi parameter menggunakan Grid Search
tuneResult <- tune(svm, rainfall ~ ., data = dataTrain,
                  ranges = list(epsilon = seq(0,1,0.1), C = 2^(-5:5),
gamma = 2^(-5:5)))
bestModel <- tuneResult$best.model
bestModel
# Prediksi menggunakan model terbaik
pred_train <- predict(bestModel, x_train)
pred_test <- predict(bestModel, x_test)
# Denormalisasi data prediksi dan data actual
denormalize <- function(x, min_x, max_x) x * (max_x - min_x) + min_x
rainfall_min <- min(data_weekly$rainfall, na.rm = TRUE)
rainfall_max <- max(data_weekly$rainfall, na.rm = TRUE)
# Denormalisasi data prediksi
pred_train_denorm <- denormalize(pred_train, rainfall_min, rainfall_max)
pred_test_denorm <- denormalize(pred_test, rainfall_min, rainfall_max)
# Denormalisasi data aktual
y_train_denorm <- denormalize(y_train, rainfall_min, rainfall_max)
y_test_denorm <- denormalize(y_test, rainfall_min, rainfall_max)
# Menghitung metrik evaluasi
rmse_train <- rmse(y_train_denorm, pred_train_denorm)
mae_train <- mae(y_train_denorm, pred_train_denorm)
rmse_test <- rmse(y_test_denorm, pred_test_denorm)
mae_test <- mae(y_test_denorm, pred_test_denorm)
# Menampilkan hasil metrik
cat("Training Metrics:\n")
cat("RMSE: ", rmse_train, "\n")

```

```

cat("MAE: ", mae_train, "\n")
cat("\nTesting Metrics:\n")
cat("RMSE: ", rmse_test, "\n")
cat("MAE: ", mae_test, "\n")
# Plot perbandingan data aktual dan prediksi untuk testing
comparison_test <- data.frame(
  Week = dataTest$week,
  Actual = y_test_denorm,
  Predicted = pred_test_denorm
)
ggplot(comparison_test, aes(x = Week)) +
  geom_line(aes(y = Actual, color = "Actual")) +
  geom_point(aes(y = Actual, color = "Actual"), size = 1) +
  geom_line(aes(y = Predicted, color = "Predicted")) +
  geom_point(aes(y = Predicted, color = "Predicted"), size = 1) +
  labs(
    title = "Perbandingan Data Aktual dan Prediksi (Testing Set)",
    x = "Minggu",
    y = "Curah Hujan"
  ) +
  scale_color_manual(values = c("Actual" = "blue", "Predicted" = "orange"))
+
  theme_minimal()
# Menghitung vektor pembobot
weights <- t(bestModel$coefs) %*% bestModel$SV
cat("Vektor Pembobot: ", weights, "\n")
# Nilai bias
bias <- bestModel$rho
cat("Nilai Bias (Intercept): ", bias, "\n")
# Nilai alpha
alpha <- bestModel$coefs
cat("Nilai Alpha: \n")
View(alpha)
# Menghitung R-Square
# Fungsi untuk menghitung R-Square
calculate_rsquare <- function(actual, predicted) {
  ss_total <- sum((actual - mean(actual))^2)
  ss_residual <- sum((actual - predicted)^2)
  r_square <- 1 - (ss_residual / ss_total)
  return(r_square)
}
# Menghitung R-Square untuk data testing
r_square_test <- calculate_rsquare(y_test_denorm, pred_test_denorm)
# Menampilkan hasil R-Square
cat("Testing R-Square: ", r_square_test, "\n")

```

## RIWAYAT HIDUP



Della Safitri lahir di Malang pada tanggal 21 Desember 2000, biasa dipanggil Della. Penulis tinggal di Jalan Darmo Sastro Sumber Manjing Wetan RT/RW: 10/03, Desa Sumber Manjing Wetan, Kecamatan Sumber Manjing Wetan, Kabupaten Malang. Penulis merupakan anak pertama dari dua bersaudara yakni putri dari Bapak Sholi dan Ibu Siti Khoiriyah.

Penulis telah menempuh Pendidikan formal mulai dari pendidikan dasar di SDN Ringisari 01 (2007-2013). Setelah itu, penulis melanjutkan Pendidikan menengah pertama di MTs Negeri 04 Malang (2013-2016). Kemudian penulis melanjutkan pendidikan menengah atas di SMAS Ibrahimy Sukorejo (2016-2019) dan tahun 2019 penulis menempuh pendidikan tinggi di Universitas Islam Negeri Maulana Malik Ibrahim Malang mengambil Program Studi Matematika, Fakultas Sains dan Teknologi.



KEMENTERIAN AGAMA RI  
UNIVERSITAS ISLAM NEGERI  
MAULANA MALIK IBRAHIM MALANG  
FAKULTAS SAINS DAN TEKNOLOGI  
Jl. Gajayana No.50 Dinoyo Malang Telp. / Fax. (0341)558933

### BUKTI KONSULTASI SKRIPSI

Nama : Della Sufitri  
NIM : 19610030  
Fakultas / Jurusan : Sains dan Teknologi / Matematika  
Judul Skripsi : Penerapan Metode *Support Vector Regression* (SVR) pada  
Prediksi Curah Hujan di Jawa Timur  
Pembimbing I : Dr. Fachrur Rozi, M.Si  
Pembimbing II : Muhammad Nafie Jauhari, M.Si

No	Tanggal	Hal	Tanda Tangan
1.	14 Januari 2024	Konsultasi Bab I	1.
2.	18 Februari 2024	Konsultasi Bab II dan III	2.
3.	12 Maret 2024	Konsultasi Revisi Bab I, II, dan III	3.
4.	5 April 2024	Konsultasi Revisi Bab I, II, dan III	4.
5.	17 April 2024	ACC Bab I, II, dan III	5.
6.	10 Mei 2024	Konsultasi Kajian Agama Bab I dan II	6.
7.	21 Mei 2024	Konsultasi Revisi Kajian Agama Bab I dan II	7.
8.	5 Juni 2024	ACC Kajian Agama Bab I dan II	8.
9.	21 Juli 2024	ACC Seminar Proposal	9.
10.	9 Agustus 2024	Konsultasi Revisi Seminar Proposal	10.



KEMENTERIAN AGAMA RI  
UNIVERSITAS ISLAM NEGERI  
MAULANA MALIK IBRAHIM MALANG  
FAKULTAS SAINS DAN TEKNOLOGI  
Jl. Gajayana No.50 Dinoyo Malang Telp. / Fax. (0341)558933

11.	5 September 2024	Konsultasi Bab IV dan V	11. <i>Fr</i>
12.	14 Oktober 2024	Konsultasi Kajian Agama Bab IV	12. <del><i>Fr</i></del>
13.	7 November 2024	ACC Kajian Agama Bab IV	13. <del><i>Fr</i></del>
14.	20 November 2024	ACC Seminar Hasil	14. <i>Fr</i>
15.	2 Desember 2024	Konsultasi Revisi Seminar Hasil	15. <i>Fr</i>
16.	8 Desember 2024	ACC Sidang Skripsi	16. <i>Fr</i>
17.	19 Desember 2024	ACC Keseluruhan	17. <i>Fr</i>

Malang, 19 Desember 2024

Mengetahui,

Ketua Program Studi Matematika



*Elly Susanti*  
Dr. Elly Susanti, M.Sc.

NIP. 19741129 200012 2 005