

**ANALISIS DAN IMPLEMENTASI MODEL REGRESI UNTUK
MEMPREDIKSI PRODUKSI TOMAT DI PROVINSI JAWA BARAT
BERBASIS DATA *OPEN GOVERNMENT***

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana

Program Studi Sistem Informasi



disusun oleh

DESTYAN YOGA ANGGRIAWAN

22.12.2445

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

**ANALISIS DAN IMPLEMENTASI MODEL REGRESI UNTUK
MEMPREDIKSI PRODUKSI TOMAT DI PROVINSI JAWA
BARAT BERBASIS DATA *OPEN GOVERNMENT***

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Sistem Informasi



disusun oleh

DESTYAN YOGA ANGGRIAWAN

22.12.2445

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS DAN IMPLEMENTASI MODEL REGRESI UNTUK
MEMPREDIKSI PRODUKSI TOMAT DI PROVINSI JAWA BARAT
BERBASIS DATA *OPEN GOVERNMENT***

yang disusun dan diajukan oleh

Destyan Yoga Anggriawan

22.12.2445

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 27 Februari 2026

Dosen Pembimbing,



Norhikmah, S.Kom., M.Kom

NIK. 190302245

HALAMAN PENGESAHAN
SKRIPSI
ANALISIS DAN IMPLEMENTASI MODEL REGRESI UNTUK
MEMPREDIKSI PRODUKSI TOMAT DI PROVINSI JAWA BARAT
BERBASIS DATA *OPEN GOVERNMENT*

yang disusun dan diajukan oleh

Destyan Yoga Anggriawan

22.12.2445

Telah dipertahankan di depan Dewan Penguji
pada tanggal 27 Februari 2026

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Afrig Aminuddin, S.Kom., M. Eng., Ph.D
NIK. 190302351



Irma Rofni Wulandari, S.Pd., M.Eng
NIK. 190302329



Norhikmah, S.Kom., M.Kom
NIK. 190302245



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 27 Februari 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Destyan Yoga Anggriawan
NIM : 22.12.2445

Menyatakan bahwa Skripsi dengan judul berikut:

Analisis dan Implementasi Model Regresi Untuk Memprediksi Produksi Tomat di Provinsi Jawa Barat Berbasis Data *Open Government*

Dosen Pembimbing : Norhikmah, S.Kom., M.Kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 27 Februari 2026

Yang Menyatakan,



Destyan Yoga Anggriawan

HALAMAN PERSEMBAHAN

Dengan memanjatkan puji dan syukur ke hadirat Allah SWT atas segala rahmat, nikmat, dan karunia-Nya, penulis akhirnya dapat menyelesaikan skripsi ini. Karya ini bukan sekadar hasil akademik, melainkan rangkuman dari doa, air mata, perjuangan, dan harapan yang menyertai setiap langkah penulis. Dengan penuh rasa syukur dan ketulusan hati, skripsi ini penulis persembahkan kepada:

1. Ayah tercinta, terima kasih atas keteguhan, kerja keras, dan doa yang tak pernah terucap namun selalu terasa. Setiap nasihat dan pengorbanan Ayah menjadi pijakan kuat bagi penulis untuk terus melangkah, meskipun jalan yang ditempuh tidak selalu mudah. Keteladanan Ayah dalam menghadapi kehidupan dengan penuh tanggung jawab dan kesabaran menjadi inspirasi terbesar bagi penulis untuk tidak mudah menyerah dalam menghadapi tantangan. Dukungan moral maupun pengorbanan yang diberikan dengan tulus menjadi alasan utama penulis untuk terus berjuang hingga mencapai titik ini.
2. Ibu tercinta, terima kasih atas kasih sayang yang tak pernah berkurang, doa yang selalu mengalir di setiap waktu, serta kesabaran yang tiada batas. Terima kasih atas setiap pengorbanan yang Ibu lakukan, melangkah ke sana kemari, menghadapi lelah dan berbagai rintangan demi mencari rezeki agar penulis dapat menempuh pendidikan hingga bangku perkuliahan. Tidak terhitung waktu, tenaga, dan air mata yang Ibu korbankan demi masa depan penulis. Semua jerih payah dan ketulusan Ibu menjadi sumber kekuatan terbesar bagi penulis untuk bangkit, bertahan, dan menyelesaikan skripsi ini. Setiap doa yang Ibu panjatkan menjadi penenang dan penguat di saat penulis merasa lemah.
3. Adik tercinta dan sahabat-sahabat terbaik, terima kasih atas dukungan, doa, semangat, serta kebersamaan yang selalu menguatkan. Kehadiran kalian menjadi pengingat bahwa penulis tidak pernah berjalan sendiri dalam menghadapi setiap proses kehidupan. Tawa, cerita, diskusi, serta motivasi yang kalian berikan menjadi energi tambahan di saat penulis merasa lelah dan hampir menyerah. Kebersamaan yang

terjalin menjadi warna tersendiri dalam perjalanan akademik ini dan memberikan semangat untuk terus melangkah maju.

4. Ibu Norhikmah, S.Kom., M.Kom, selaku dosen pembimbing, terima kasih atas kesabaran, perhatian, dan bimbingan yang diberikan dengan tulus. Setiap arahan, koreksi, dan motivasi yang Ibu berikan menjadi pedoman yang jelas bagi penulis dalam menyelesaikan skripsi ini dengan lebih terarah dan sistematis. Kesediaan Ibu meluangkan waktu untuk membimbing dan memberikan masukan yang membangun sangat berarti bagi proses penyempurnaan penelitian ini. Ilmu dan pengalaman yang dibagikan menjadi bekal berharga bagi penulis ke depannya.
5. Seluruh dosen Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta, terima kasih atas ilmu, inspirasi, dan pengalaman berharga yang telah diberikan selama masa perkuliahan, yang menjadi bekal penting bagi perjalanan penulis ke depan. Setiap materi, diskusi, serta pembelajaran yang diberikan telah membentuk pola pikir dan kemampuan analitis penulis hingga mampu menyelesaikan penelitian ini. Semoga ilmu yang telah diberikan menjadi amal jariyah dan bermanfaat bagi banyak orang.
6. Untuk diri sendiri, Destyan Yoga Anggriawan, terima kasih karena telah bertahan di saat lelah, tetap melangkah ketika ingin menyerah, dan percaya bahwa semua proses ini akan berbuah baik. Terima kasih karena tidak menyerah pada keadaan, tetap bangkit setiap kali mengalami kegagalan, dan terus berusaha menjadi pribadi yang lebih baik. Semoga segala perjuangan, pengorbanan, dan air mata yang telah dilalui menjadi awal dari langkah-langkah besar menuju masa depan yang lebih baik, penuh keberkahan, dan kesuksesan.

KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Tuhan Yang Maha Esa, yang telah melimpahkan rahmat, hidayah, serta karunia-Nya, sehingga penulis dapat menyelesaikan laporan tugas akhir ini dengan lancar. Tugas akhir ini disusun untuk memenuhi salah satu syarat kelulusan Program Sarjana pada Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.

Selain itu penulis dengan segala kerendahan hati ingin menyampaikan rasa terima kasih kepada semua pihak yang telah berjasa memberikan dukungan dan bantuan untuk menyelesaikan laporan tugas akhir ini. Untuk itu penulis mengucapkan terima kasih kepada :

1. Bapak Prof. Dr. M. Suyanto, M.M selaku Rektor Universitas Amikom Yogyakarta
2. Ibu Prof. Dr. Kusri, M. Kom selaku Dekan Program Fakultas Ilmu Komputer
3. Bapak Anggit Dwi Hartono, M.Kom selaku Ketua Program Studi Sistem Informasi
4. Ibu Norhikmah, S.Kom., M.Kom selaku Dosen Pembimbing yang memberikan arahan, saran, dan motivasi terhadap penulis
5. Kedua orang tua, keluarga besar serta teman-teman tercinta yang memberikan semangat,dukungan,kasih sayang dan doa kepada penulis.
6. Semua pihak yang tidak dapat disebutkan satu persatu yang telah membantu dalam penyelesaian tugas akhir ini

Akhir kata, penulis berharap skripsi ini dapat memberikan manfaat dan kontribusi dalam pengembangan ilmu pengetahuan, khususnya di bidang machine learning dan analisis data pada sektor pertanian. Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan demi perbaikan di masa mendatang.

Yogyakarta, 27 Februari 2026

Penulis

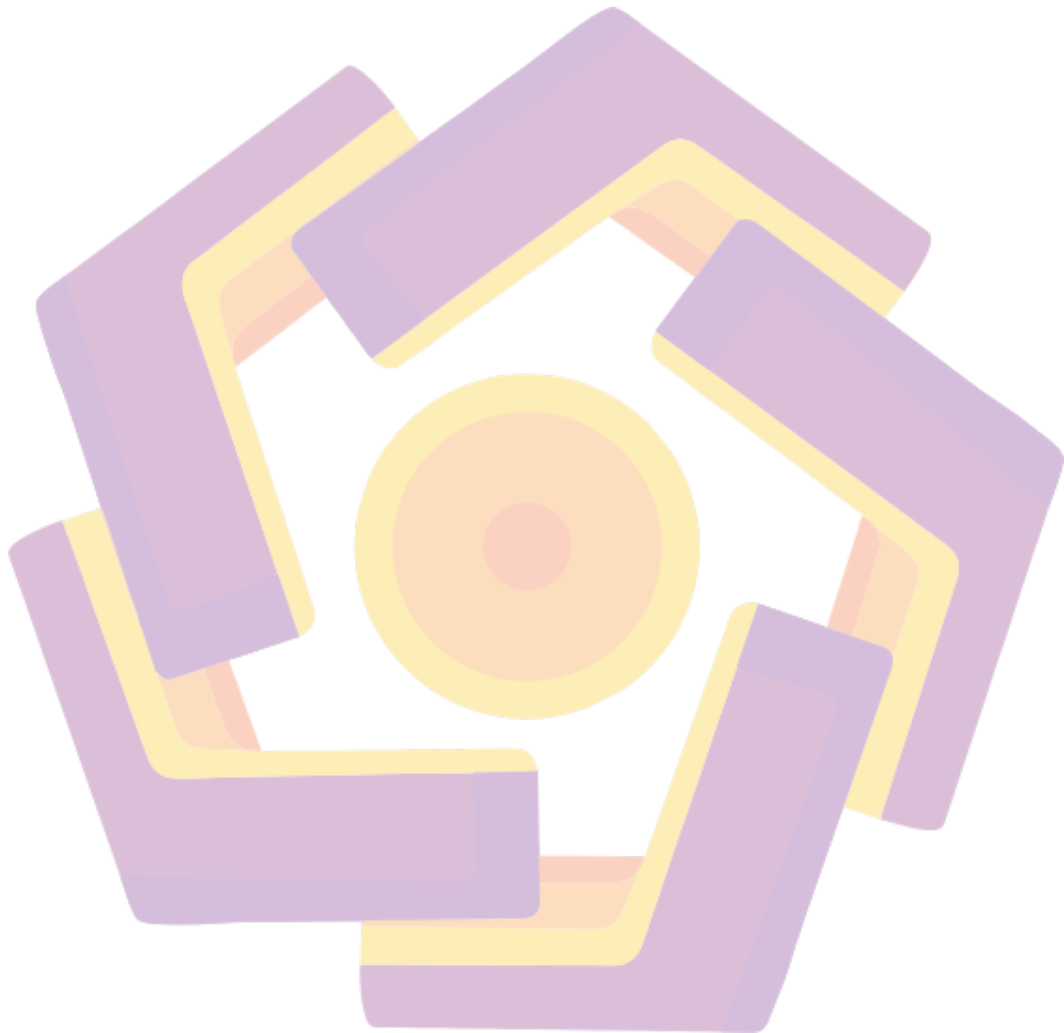
DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI.....	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xii
DAFTAR GAMBAR.....	xiii
DAFTAR LAMBANG DAN SINGKATAN	xvi
DAFTAR ISTILAH.....	xvii
INTISARI	xviii
<i>ABSTRACT</i>	xix
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	5
1.6 Sistematika Penulisan.....	6
BAB II TINJAUAN PUSTAKA	8
2.1 Studi Literatur	8

2.2	Dasar Teori	20
2.2.1	Prediksi dan Peramalan	20
2.2.2	Regresi Linear	20
2.2.3	<i>Support Vector Regression (SVR)</i>	21
2.2.4	<i>Random Forest Regressor</i>	22
2.2.5	<i>Open Government Data (OGD)</i> dan Satu Data Indonesia	23
2.2.6	Evaluasi Model Regresi	23
2.2.7	<i>K-fold Cross Validation</i>	23
2.2.8	Implementasi Model Prediksi	24
BAB III METODE PENELITIAN		25
3.1	Objek Penelitian	25
3.2	Alur Penelitian	25
3.3	Alat dan Bahan	36
3.3.1	Data Penelitian	36
3.3.2	Alat/instrumen	36
BAB IV HASIL DAN PEMBAHASAN		38
4.1	Dataset	38
4.2	Pembersihan Data	42
4.2.1	Hapus Duplikat	42
4.2.2	Cek Missing Value	43
4.2.3	Hapus Baris Mengandung NaN	44
4.3	Eksplorasi Data	45
4.4	Distribusi Fitur Kategorikal	47
4.5	Korelasi Fitur Numerik	48
4.6	Deteksi <i>Outlier Boxplot</i> (Sebelum Dihapus)	50

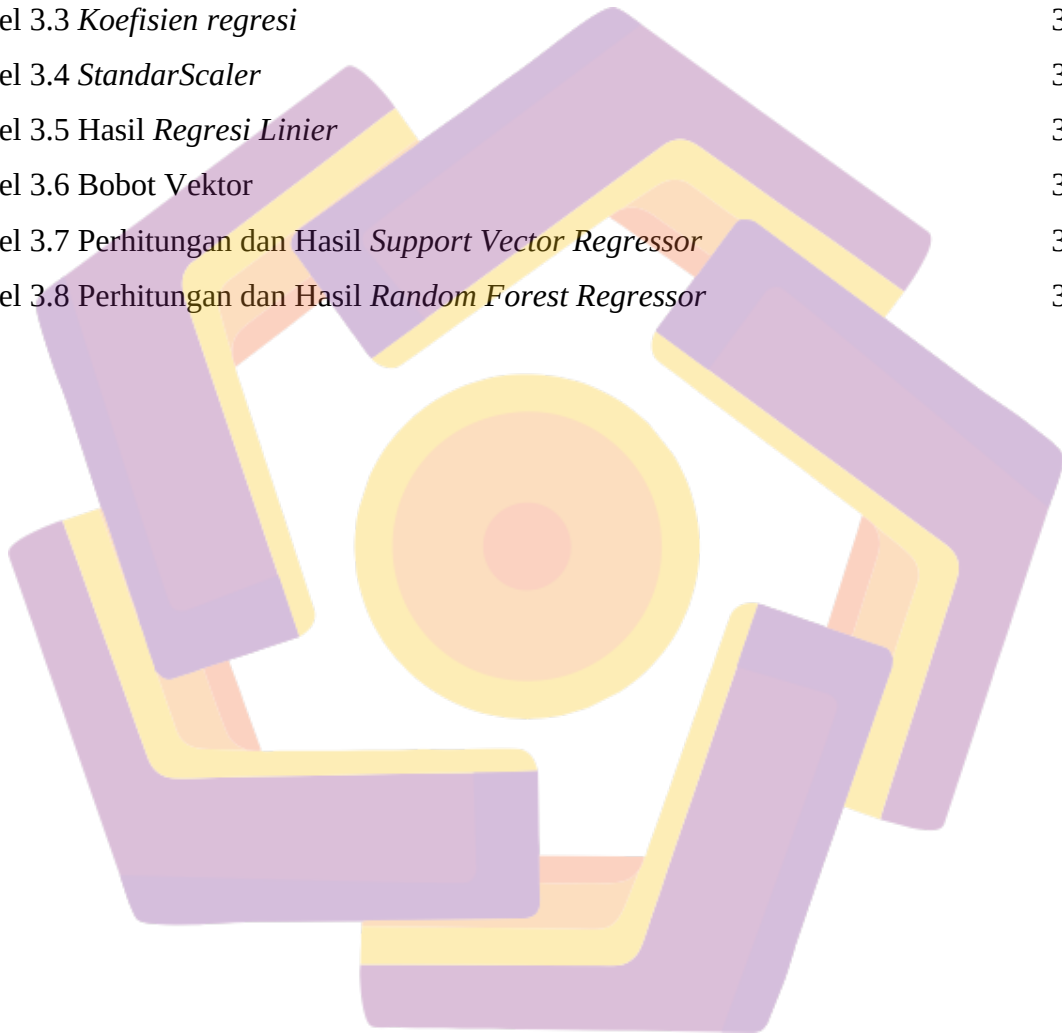
4.7	Distribusi Fitur Numerik (Sebelum Dihapus)	52
4.8	Hapus <i>Outlier</i> dengan IQR.....	53
4.9	Normalisasi Target	54
4.10	Deteksi <i>Outlier Boxplot</i> (Sesudah Dihapus).....	55
4.11	Distribusi Fitur Numerik (Sesudah Dihapus)	56
4.12	Pembagian Variabel X dan Y	57
4.13	<i>Encoding</i> dan Normalisasi.....	59
4.14	<i>Train-Test Split</i>	60
4.15	Definisi dan Pelatihan Model	61
4.15.1	<i>K-fold Cross Validation</i>	62
4.15.2	Metrik Evaluasi.....	62
4.16	Pemilihan Model Terbaik dan Visualisasi Evaluasi	64
4.17	Analisis Fitur vs Target	66
4.18	Hitung Korelasi <i>Pearson</i> per fitur	68
4.19	<i>Koefisien Regresi Linear</i>	70
4.20	R^2 Setiap Fitur.....	72
4.21	Scatter Plot Aktual vs Prediksi	73
4.22	<i>Residual Plot</i>	75
4.23	Simpan Model.....	77
4.24	Implementasi Model	77
4.21.1	<i>Source Code App</i> Prediksi Produksi Tomat	78
4.21.2	Membuat <i>Repository</i> Github	86
4.21.3	Membuat App <i>Streamlit</i>	86
4.21.4	Hasil Pengujian.....	87
BAB V PENUTUP		89

5.1	Kesimpulan.....	89
5.2	Saran.....	90
REFERENSI		92



DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian	11
Tabel 3.1 Perhitungan Regresi Linear	30
Tabel 3.2 <i>MinMaxScaler</i>	30
Tabel 3.3 <i>Koefisien regresi</i>	30
Tabel 3.4 <i>StandarScaler</i>	30
Tabel 3.5 Hasil <i>Regresi Linier</i>	31
Tabel 3.6 Bobot Vektor	33
Tabel 3.7 Perhitungan dan Hasil <i>Support Vector Regressor</i>	34
Tabel 3.8 Perhitungan dan Hasil <i>Random Forest Regressor</i>	35



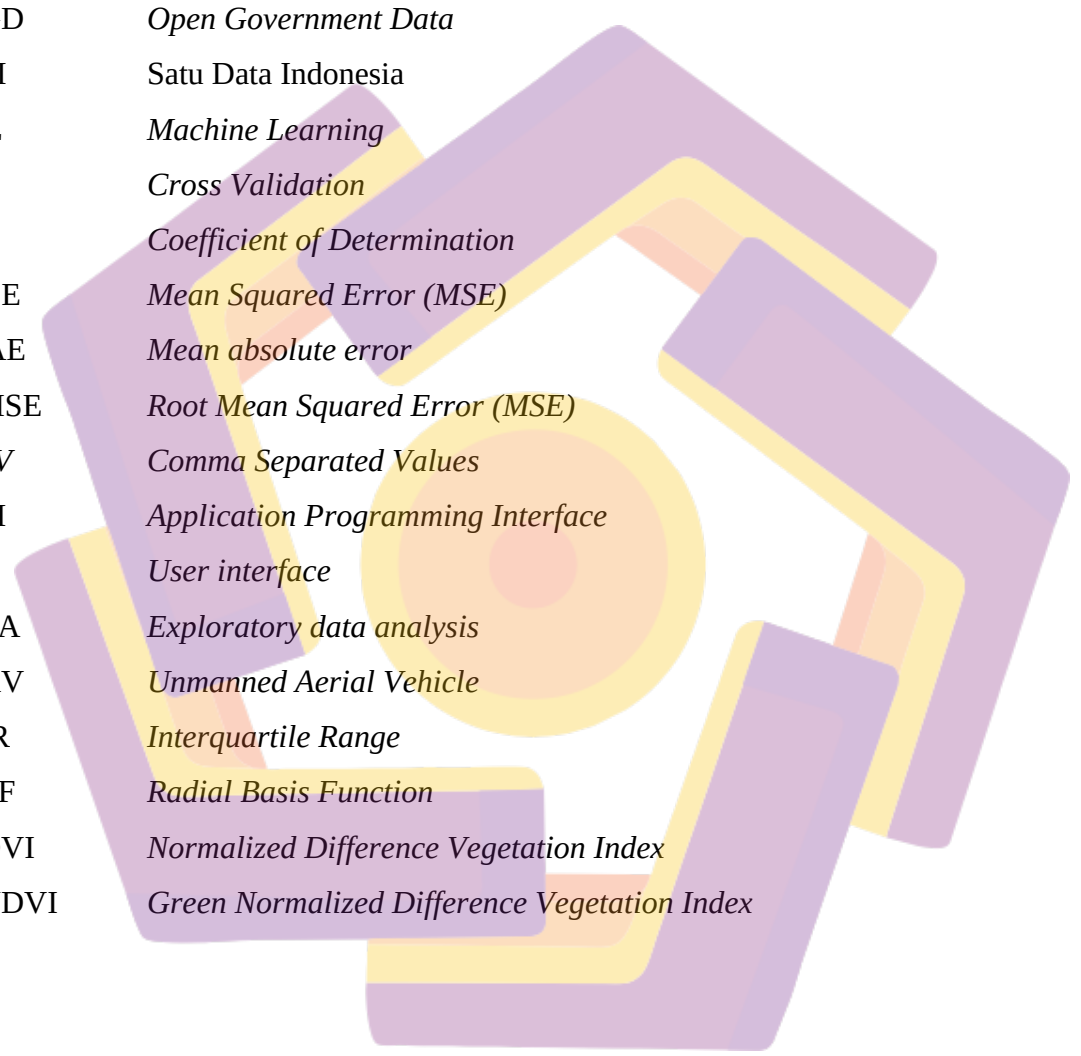
DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian	26
Gambar 4.1 <i>Script</i> Import <i>Library</i>	38
Gambar 4.2 <i>Script</i> Load Dataset	39
Gambar 4.3 Hasil Load Dataset	40
Gambar 4.4 Skript Penggabungan Dataset	41
Gambar 4.5 Hasil Penggabungan Dataset	41
Gambar 4.6 Skript Hapus Duplikat	42
Gambar 4.7 Skript Missing Value	43
Gambar 4.8 Hasil Missing Value	43
Gambar 4.9 <i>Script</i> Hapus Baris Mengandung NaN	44
Gambar 4.10 Hasil Hapus Baris Mengandung NaN	45
Gambar 4.11 <i>Script</i> Info Dataset	45
Gambar 4.12 Hasil Info Dataset	46
Gambar 4.13 <i>Script</i> Statistik Deskriptif	46
Gambar 4.14 Hasil Statistik Deskriptif	47
Gambar 4.15 <i>Script</i> Distribusi Fitur Kategorikal	48
Gambar 4.16 Hasil Distribusi Fitur Kategorikal	48
Gambar 4.17 <i>Script</i> Korelasi Fitur Numerik	49
Gambar 4.18 Hasil Korelasi Fitur Numerik	50
Gambar 4.19 <i>Script</i> Deteksi <i>Outlier Boxplot</i> (Sebelum Dihapus)	51
Gambar 4.20 Hasil Deteksi <i>Outlier Boxplot</i> (Sebelum Dihapus)	51
Gambar 4.21 <i>Script</i> Distribusi Fitur Numerik (Sebelum Dihapus)	52
Gambar 4.22 Hasil Distribusi Fitur Numerik (Sebelum Dihapus)	53
Gambar 4.23 <i>Script</i> Hapus <i>Outlier</i> dengan IQR	54
Gambar 4.24 Hasil Hapus <i>Outlier</i> dengan IQR	54
Gambar 4.25 <i>Script</i> Normalisasi Target	55
Gambar 4.26 <i>Script</i> Deteksi <i>Outlier Boxplot</i> (Sesudah Dihapus)	55
Gambar 4.27 Hasil Deteksi <i>Outlier Boxplot</i> (Sesudah Dihapus)	56
Gambar 4.28 <i>Script</i> Distribusi Fitur Numerik (Sesudah Dihapus)	57

Gambar 4.29 Hasil Distribusi Fitur Numerik (Sesudah Dihapus)	57
Gambar 4.30 <i>Script</i> Pembagian Variabel X dan Y	58
Gambar 4.31 Hasil Pembagian Variabel X dan Y	59
Gambar 4.32 <i>Script Encoding</i> dan Normalisasi	59
Gambar 4.33 Hasil <i>Encoding</i> dan Normalisasi	60
Gambar 4.34 <i>Script</i> dan Hasil <i>Train-Test Split</i>	60
Gambar 4.35 <i>Script</i> Pelatihan Model	61
Gambar 4.36 <i>Script K-fold Cross Validation</i>	62
Gambar 4.37 <i>Script</i> Metrik Evaluasi	63
Gambar 4.38 Hasil Pelatihan Model, <i>K-fold Cross Validation</i> , dan Metrik Evaluasi	64
Gambar 4.39 <i>Script</i> Pemilihan Model dan Visualisasi Evaluasi	65
Gambar 4.40 Hasil Pemilihan Model dan Visualisasi Evaluasi	66
Gambar 4.41 <i>Script</i> Analisis Fitur vs Target	67
Gambar 4.42 Hasil Analisis Fitur vs Target	68
Gambar 4.43 <i>Script</i> Hitung Korelasi <i>Pearson</i> per fitur	69
Gambar 4.44 Hasil Hitung Korelasi <i>Pearson</i> per fitur	70
Gambar 4.45 <i>Script</i> Koefisien Regresi Linear	70
Gambar 4.46 Hasil Koefisien Regresi Linear	71
Gambar 4.47 <i>Script</i> R^2 Setiap Model	72
Gambar 4.48 Hasil R^2 Setiap Model	73
Gambar 4.49 <i>Script</i> Scatter Plot Aktual vs Prediksi	74
Gambar 4.50 Hasil Scatter Plot Aktual vs Prediksi	75
Gambar 4.51 <i>Script Residual Plot</i>	76
Gambar 4.52 Hasil <i>Residual Plot</i>	76
Gambar 4.53 <i>Script</i> dan Hasil Simpan Model	77
Gambar 4.54 <i>Source Code App</i> Prediksi Produksi Tomat	78
Gambar 4.55 Lanjutan <i>Source Code App</i> Prediksi Produksi Tomat	79
Gambar 4.56 Lanjutan <i>Source Code App</i> Prediksi Produksi Tomat	79
Gambar 4.57 Lanjutan <i>Source Code App</i> Prediksi Produksi Tomat	80
Gambar 4.58 Lanjutan <i>Source Code App</i> Prediksi Produksi Tomat	80
Gambar 4.59 Lanjutan <i>Source Code App</i> Prediksi Produksi Tomat	81

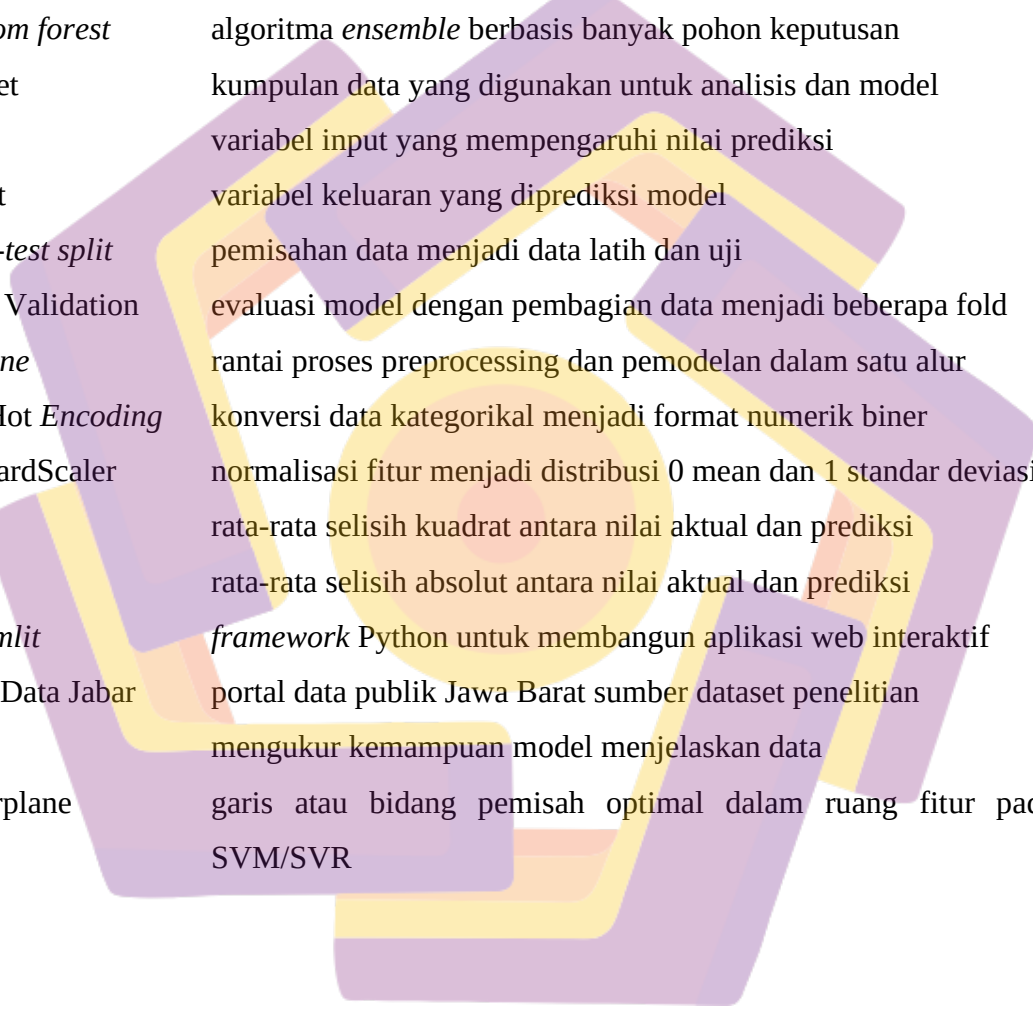
Gambar 4.60 Lanjutan <i>Source Code App</i> Prediksi Produksi Tomat	81
Gambar 4.61 Lanjutan <i>Source Code App</i> Prediksi Produksi Tomat	82
Gambar 4.62 Lanjutan <i>Source Code App</i> Prediksi Produksi Tomat	82
Gambar 4.63 Lanjutan <i>Source Code App</i> Prediksi Produksi Tomat	83
Gambar 4.64 Lanjutan <i>Source Code App</i> Prediksi Produksi Tomat	83
Gambar 4.65 Lanjutan <i>Source Code App</i> Prediksi Produksi Tomat	84
Gambar 4.66 Lanjutan <i>Source Code App</i> Prediksi Produksi Tomat	84
Gambar 4.67 Lanjutan <i>Source Code App</i> Prediksi Produksi Tomat	85
Gambar 4.68 Lanjutan <i>Source Code App</i> Prediksi Produksi Tomat	85
Gambar 4.69 Lanjutan <i>Source Code App</i> Prediksi Produksi Tomat	85
Gambar 4.70 Membuat <i>Repository</i> Github	86
Gambar 4.71 Membuat App <i>Streamlit</i>	87
Gambar 4.72 Hasil Pengujian	87
Gambar 4.73 Lanjutan Hasil Pengujian	88

DAFTAR LAMBANG DAN SINGKATAN



LR	<i>Regresi Linear</i>
SVR	<i>Support Vector Regression</i>
RFR	<i>Random Forest Regressor</i>
OGD	<i>Open Government Data</i>
SDI	<i>Satu Data Indonesia</i>
ML	<i>Machine Learning</i>
CV	<i>Cross Validation</i>
R^2	<i>Coefficient of Determination</i>
MSE	<i>Mean Squared Error (MSE)</i>
MAE	<i>Mean absolute error</i>
RMSE	<i>Root Mean Squared Error (MSE)</i>
CSV	<i>Comma Separated Values</i>
API	<i>Application Programming Interface</i>
UI	<i>User interface</i>
EDA	<i>Exploratory data analysis</i>
UAV	<i>Unmanned Aerial Vehicle</i>
IQR	<i>Interquartile Range</i>
RBF	<i>Radial Basis Function</i>
NDVI	<i>Normalized Difference Vegetation Index</i>
GNDVI	<i>Green Normalized Difference Vegetation Index</i>

DAFTAR ISTILAH



Prediksi	perkiraan nilai masa depan berdasarkan data historis
Regresi	metode statistik untuk menentukan hubungan antar variabel
<i>Regresi Linear</i>	model regresi dengan hubungan linier antar fitur
SVR	<i>Support Vector Regression</i> untuk prediksi berbasis hyperplane
<i>Random forest</i>	algoritma <i>ensemble</i> berbasis banyak pohon keputusan
Dataset	kumpulan data yang digunakan untuk analisis dan model
Fitur	variabel input yang mempengaruhi nilai prediksi
Target	variabel keluaran yang diprediksi model
<i>Train-test split</i>	pemisahan data menjadi data latih dan uji
Cross Validation	evaluasi model dengan pembagian data menjadi beberapa fold
<i>Pipeline</i>	rantai proses <i>preprocessing</i> dan pemodelan dalam satu alur
<i>One Hot Encoding</i>	konversi data kategorikal menjadi format numerik biner
StandardScaler	normalisasi fitur menjadi distribusi 0 mean dan 1 standar deviasi
MSE	rata-rata selisih kuadrat antara nilai aktual dan prediksi
MAE	rata-rata selisih absolut antara nilai aktual dan prediksi
<i>Streamlit</i>	<i>framework</i> Python untuk membangun aplikasi web interaktif
Open Data Jabar	portal data publik Jawa Barat sumber dataset penelitian
R ²	mengukur kemampuan model menjelaskan data
Hyperplane	garis atau bidang pemisah optimal dalam ruang fitur pada SVM/SVR

INTISARI

Produksi tomat di Provinsi Jawa Barat memiliki peran penting dalam mendukung ketahanan pangan dan sektor hortikultura, namun jumlah produksinya cenderung mengalami fluktuasi dari tahun ke tahun. Ketidakakuratan dalam memprediksi produksi tomat dapat berdampak pada ketidakseimbangan pasokan, perencanaan distribusi yang kurang optimal, serta potensi kerugian bagi petani dan pemangku kepentingan. Oleh karena itu, diperlukan pendekatan berbasis data untuk memprediksi produksi tomat secara lebih akurat dan objektif dengan memanfaatkan data historis yang tersedia.

Penelitian ini bertujuan untuk memprediksi produksi tomat di Provinsi Jawa Barat menggunakan algoritma *Regresi Linear*, *Support Vector Regression (SVR)*, dan *Random Forest Regressor* berbasis data *Open Government Data* dari *Open Data Jabar* periode 2013–2024. Data diproses melalui tahapan pembersihan data, penggabungan dataset, dan normalisasi fitur, kemudian dibagi menjadi data latih dan data uji dengan rasio 80:20. Evaluasi model dilakukan menggunakan *K-fold Cross Validation (K=5)* dengan metrik R^2 , *MSE*, dan *MAE*. Hasil penelitian menunjukkan bahwa model *Regresi Linear* memberikan performa terbaik dibandingkan *SVR* dan *Random Forest Regressor*, dengan nilai R^2 pada data uji sebesar 0,912, yang menunjukkan bahwa 91,2% variasi produksi tomat dapat dijelaskan oleh variabel dalam model. Temuan ini mengindikasikan bahwa hubungan antar variabel dalam dataset cenderung bersifat linier setelah proses pembersihan data dilakukan, serta menunjukkan bahwa *Regresi Linear* efektif digunakan sebagai model prediksi produksi tomat di Provinsi Jawa Barat.

Kata kunci: *Prediksi Produksi Tomat, Regresi Linear, Support Vector Regression, Random Forest Regressor, Open Government Data*

ABSTRACT

Tomato production in West Java Province plays an important role in supporting food security and the horticulture sector, but production volumes tend to fluctuate from year to year. Inaccurate tomato production forecasts can lead to supply imbalances, suboptimal distribution planning, and potential losses for farmers and stakeholders. Therefore, a data-driven approach is needed to predict tomato production more accurately and objectively by utilizing available historical data.

This study aims to predict tomato production in West Java Province using Regresi Linear, Support Vector Regression (SVR), and Random Forest Regressor algorithms based on Open Government Data from Open Data Jabar for the period 2013–2024. The data was processed through data cleaning, dataset merging, and feature normalization stages, then divided into training and test data with a ratio of 80:20. Model evaluation was performed using K-fold Cross Validation (K=5) with R^2 , MSE, and MAE metrics. The results showed that the Regresi Linear model performed best compared to SVR and Random Forest Regressor, with an R^2 value of 0.912 on the test data, indicating that 91.2% of the variation in tomato production could be explained by the variables in the model. These findings indicate that the relationship between variables in the dataset tends to be linear after the data cleaning process, and show that Regresi Linear is effective as a model for predicting tomato production in West Java Province.

Keyword: Tomato Production Prediction, Regresi Linear, Support Vector Regression, Random Forest Regressor, Open Government Data