

**PREDIKSI PRODUKSI TANAMAN PADI DI PROVINSI JAWA
TENGAH MENGGUNAKAN ALGORITMA LINEAR
REGRESSION**

SKRIPSI

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



disusun oleh

IMEL

21.12.2182

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

**PREDIKSI PRODUKSI TANAMAN PADI DI PROVINSI JAWA
TENGAH MENGGUNAKAN ALGORITMA LINEAR
REGRESSION**

SKRIPSI

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



disusun oleh

IMEL

21.12.2182

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

**PREDIKSI PRODUKSI TANAMAN PADI DI PROVINSI JAWA
TENGAH MENGGUNAKAN ALGORITMA LINEAR
REGRESSION**

yang disusun dan diajukan oleh

IMEL

21.12.2182

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 21 Februari 2025

Dosen Pembimbing,



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

HALAMAN PENGESAHAN

SKRIPSI

**PREDIKSI PRODUKSI TANAMAN PADI DI PROVINSI JAWA
TENGAH MENGGUNAKAN ALGORITMA LINEAR
REGRESSION**

yang disusun dan diajukan oleh

IMEL

21.12.2182

Telah dipertahankan di depan Dewan Penguji
pada tanggal 21 Februari 2025

Susunan Dewan Penguji

Nama Penguji

Ali Mustopa, S.Kom., M.Kom
NIK. 190302192

Agung Nugroho, S.Kom., M.Kom
NIK. 190302242

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

Tanda Tangan



Skrripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 21 Februari 2025

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom., Ph.D
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : IMEL
NIM : 21.12.2182

Menyatakan bahwa Skripsi dengan judul berikut:

Prediksi Produksi Tanaman Padi di Provinsi Jawa Tengah Menggunakan Algoritma Linear Regression

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, 21 Februari 2025

Yang Menyatakan,

IMEL

HALAMAN PERSEMBAHAN

Segala puji dan syukur saya panjatkan ke hadirat Allah SWT atas rahmat dan karunia-Nya, sehingga penelitian ini dapat diselesaikan dengan baik. Dengan penuh rasa hormat dan terima kasih, penelitian ini saya persembahkan kepada:

1. Allah SWT, atas segala berkah, kemudahan, dan petunjuk-Nya dalam setiap langkah perjalanan akademik ini.
2. Ibu dan paman tercinta, yang telah mengusahakan segala hal untuk penulis, mencurahkan kasih sayang, doa, serta jerih payah yang tak terhitung nilainya. Semangat dan perjuangan mereka menjadi sumber motivasi terbesar bagi penulis untuk menyelesaikan skripsi ini dan melangkah ke tahap pendewasaan berikutnya.
3. Saudara-saudara saya, yang selalu memberikan motivasi, kebersamaan, dan semangat dalam menjalani proses akademik ini.
4. Dosen pembimbing dan pengajar, yang telah membimbing dengan penuh kesabaran serta memberikan ilmu dan arahan berharga.
5. Teman-teman seperjuangan, yang telah menjadi bagian dari perjalanan ini dengan dukungan, semangat, dan kebersamaan.
6. Diri saya sendiri, atas usaha, kerja keras, dan kesabaran dalam menghadapi berbagai tantangan selama penyelesaian penelitian ini.

Semoga segala usaha dan ilmu yang diperoleh dapat menjadi berkah serta memberikan manfaat bagi diri sendiri, masyarakat, dan perkembangan ilmu pengetahuan.

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Tuhan Yang Maha Esa, yang telah melimpahkan rahmat, hidayah, serta karunia-Nya, sehingga penulis dapat menyelesaikan skripsi ini dengan lancar. Skripsi 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 skripsi ini. Untuk itu penulis mengucapkan terima kasih kepada :

1. Bapak Prof. Dr. M. Suyanto, M.M., selaku Rektor Universitas Amikom Yogyakarta.
2. Bapak Hanif Al Fatta, S.Kom., M.Kom., Ph.D., selaku Dekan Program Fakultas Ilmu Komputer .
3. Bapak Anggit Dwi Hartanto, M.Kom., selaku Ketua Program Studi Sistem Informasi.
4. Ibu Norhikmah, M.Kom., selaku dosen pembimbing yang telah sabar memberikan arahan, saran, dan motivasi terhadap penulis selama proses penelitian dan penulisan skripsi ini.
5. Ibu Darnis, ibu tercinta yang senantiasa memberikan kasih sayang, doa, motivasi, dan selalu menyemangati dalam setiap perjalanan saya.
6. Ayah Yuharman, ayah saya yang senantiasa memberikan kasih sayang, doa, dan selalu menyemangati saya.
7. Paman Ardun, paman tercinta, yang selalu memberikan dukungan, semangat, motivasi serta membiayai kuliah saya selama ini.
8. Adik Dayu & Rifani Maruba, adik tersayang, yang selalu memberi semangat dan dukungan dalam perjalanan ini.
9. Innaka Kubro Dani Istono, yang selalu memberikan semangat, kebersamaan, bantuan, dan menemani penulis dalam mengerjakan skripsi ini.

10. Tim Konten Prodi SI, atas segala pengalaman, dukungan, dan fasilitas yang telah diberikan, yang sangat membantu dalam penyelesaian skripsi ini.
11. BPC Amikom, atas segala dukungan, dan fasilitas yang telah diberikan, yang sangat membantu dalam penyelesaian skripsi ini.
12. Rekan-rekan seperjuangan, atas segala dukungan, kebersamaan, dan semangat yang diberikan selama proses akademik ini.
13. Teman-teman terbaik, yang telah memberikan inspirasi, motivasi, dan bantuan dalam menyelesaikan skripsi ini.
14. Seluruh pihak yang tidak dapat disebutkan satu per satu, yang telah memberikan dukungan dan kontribusi dalam penyelesaian skripsi ini.

Harapannya, skripsi ini dapat memberikan manfaat serta menjadi kontribusi yang berharga bagi pengembangan ilmu pengetahuan. Penulis juga menyadari bahwa skripsi ini masih jauh dari kesempurnaan. Oleh karena itu, segala kritik dan saran yang membangun sangat diharapkan untuk penyempurnaan di masa yang akan datang.

Yogyakarta, 21 Februari 2025

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMBANG DAN SINGKATAN	xiv
DAFTAR ISTILAH.....	xv
INTISARI	xvi
ABSTRACT.....	xvii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	6
1.3 Batasan Masalah	6
1.4 Tujuan Penelitian	6
1.5 Manfaat Penelitian	7
1.5.1 Manfaat Teoritis	7
1.5.2 Manfaat Praktisi	7
1.6 Sistematika Penulisan	8
BAB II TINJAUAN PUSTAKA	10

2.1	Studi Literatur	10
2.2	Dasar Teori.....	20
2.2.1	<i>Machine Learning</i>	20
2.2.2	Data Mining	20
2.2.3	Algoritma Linear Regression.....	22
2.2.4	Prediksi	23
2.2.5	Padi	24
2.2.6	Uji Performa.....	24
2.2.7	<i>R-Square Error</i>	25
2.2.7	<i>Root Mean Square Error (RMSE)</i>	26
2.2.8	<i>Mean Absolute Error (MAE)</i>	27
2.2.9	<i>Mean Squared Error (MSE)</i>	28
2.2.10	<i>Python</i>	28
2.2.11	<i>Hyperparameter Tuning</i>	29
2.2.12	<i>Streamlit Cloud</i>	30
BAB III METODE PENELITIAN		32
3.1	Objek Penelitian	32
3.2	Alur Penelitian	32
3.2.1	Pengumpulan Data	34
3.2.2	Preprocessing Data	35
3.2.3	Pemodelan Data	36
3.2.4	<i>Hyperparameter Tuning</i>	39
3.2.5	Evaluasi Model	45
3.2.6	<i>Deployment Model</i>	46
3.3	Alat dan Bahan	47

3.3.1	Data Penelitian	47
3.3.2	Alat/instrumen.....	47
BAB IV HASIL DAN PEMBAHASAN		48
4.1	Pengumpulan Data	48
4.2	Preprocessing Data.....	59
4.3	Pemodelan Data	65
4.4	<i>Hyperparameter Tuning</i>	68
4.5	Evaluasi Model	72
4.6	<i>Deployment Model</i>	74
BAB V PENUTUP		80
5.1	Kesimpulan	80
5.2	Saran	80
REFERENSI		82
DAFTAR LAMPIRAN.....		88

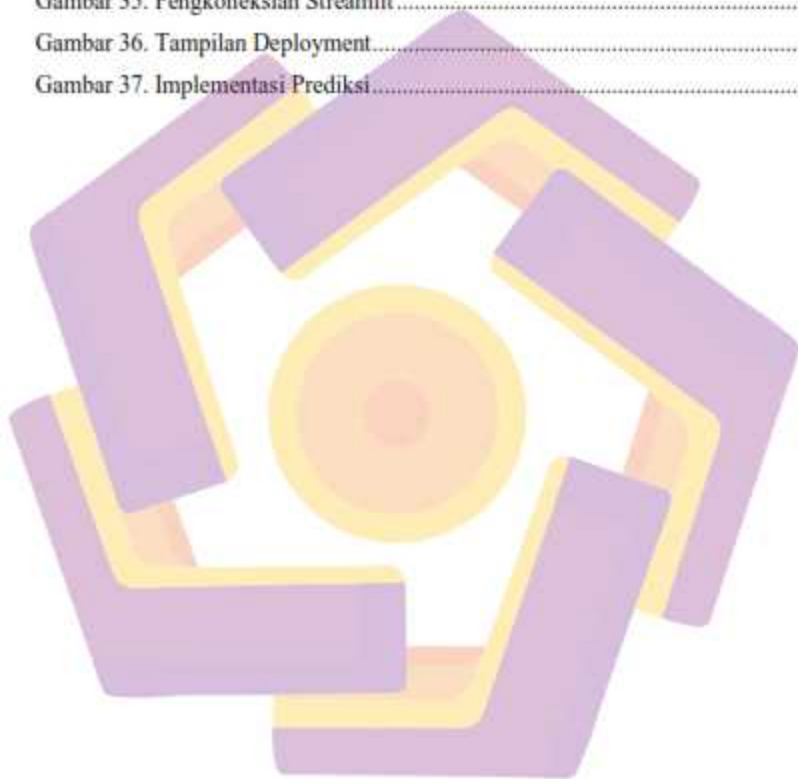
DAFTAR TABEL

Tabel 1. Contoh Dataset Perhitungan Manual	37
Tabel 2. Mengelompokkan Nilai	38
Tabel 3. Menentukan Matriks A	39
Tabel 4. Menentukan Matriks A1	39
Tabel 5. Menentukan Matriks A2	40
Tabel 6. Menentukan Matriks A3	41
Tabel 7. Menentukan Matriks A4	41
Tabel 8. Dataset Padi	48
Tabel 9. Nilai Koefisien Antar Variabel	58
Tabel 10. Hasil Hyperparameter Tuning	71
Tabel 11. Hasil Evaluasi Model	72

DAFTAR GAMBAR

Gambar 1. Alur Penelitian	33
Gambar 2. Prediksi pada Kode Program.....	43
Gambar 3. Code Program Proses Membaca Data.....	50
Gambar 4. Menampilkan Dataset Padi	50
Gambar 5. Melihat Informasi Data	51
Gambar 6. Deskripsi Dataset	52
Gambar 7. Visualisasi Raya-rata Produksi Per Kabupaten.....	53
Gambar 8. Visualisasi Rata-rata Produksi Setiap Kabupaten	55
Gambar 9. Visualisasi Rata-rata Produksi Per Tahun.....	56
Gambar 10. Kode Program Heatmap	57
Gambar 11. Heatmap Korelasi Antar Variabel.....	57
Gambar 12. Code Program Data Cleaning	59
Gambar 13. Code Program Mengecek Missing Value	60
Gambar 14. Pengecekan Outlier	60
Gambar 15. Menampilkan Data Outlier.....	61
Gambar 16. Data Tanpa Outlier.....	62
Gambar 17. Boxplot Baru	62
Gambar 18. Hasil Outlier	63
Gambar 19. Train and Test Split.....	64
Gambar 20. Scaling Data Train	65
Gambar 21. Scaling Data Test	65
Gambar 22. Code Program Pembuatan Model Linear Regression	66
Gambar 23. Kode Program Proses Pelatihan Algoritma.....	66
Gambar 24. Nilai Aktual dan Prediksi	67
Gambar 25. Akurasi Cross Validation	68
Gambar 26. Kode Program Nilai Terkini Hyperparameter.....	69
Gambar 27. Hasil Parameter Optimal GridSearchCV	69
Gambar 28. Grafik Linear Regression dengan GridSearchCV	70
Gambar 29. Cross Validation Linear Regression dengan GridSearchCV	70

Gambar 30. Kode Program Proses Prediksi Baru	73
Gambar 31. Prediksi Menggunakan Data Baru	74
Gambar 32. Membuat Repository di GitHub.....	75
Gambar 33. Source Code Tampilan Antarmuka di Streamlit.....	75
Gambar 34. Mengekspor File ke dalam Repository	76
Gambar 35. Pengkoneksian Streamlit.....	77
Gambar 36. Tampilan Deployment.....	78
Gambar 37. Implementasi Prediksi.....	79



DAFTAR LAMBANG DAN SINGKATAN

BPS	Badan Pusat Statistik
DM	Data Mining
ML	Machine Learning
LR	<i>Linear Regression</i>
R ²	<i>R-Squared</i>
MAE	<i>Mean Absolute Error</i>
MSE	<i>Mean Squared Error</i>
RMSE	<i>Root Mean Squared Error</i>
KDD	<i>Knowledge Discovery in Database</i>
IQR	<i>Interquartile Range</i>



DAFTAR ISTILAH

<i>Linear Regression</i>	Metode statistik yang digunakan untuk memodelkan hubungan antara variabel independen dan variabel dependen dalam bentuk garis lurus.
<i>Hyperparameter Tuning</i>	Proses pencarian nilai parameter terbaik untuk meningkatkan performa model machine learning
<i>GridSearchCV</i>	Teknik pencarian sistematis untuk menemukan kombinasi parameter optimal dalam machine learning.
<i>Cross-Validation</i>	Teknik evaluasi model dengan membagi dataset menjadi beberapa bagian untuk menghindari overfitting.
<i>R-Squared</i>	Metrik evaluasi yang mengukur seberapa baik model dapat menjelaskan variasi dalam data.
<i>Mean Squared Error</i>	Metrik evaluasi yang mengukur rata-rata kuadrat selisih antara nilai prediksi dan nilai aktual.
<i>Mean Absolute Error</i>	Metrik evaluasi yang menghitung rata-rata selisih absolut antara nilai prediksi dan nilai aktual.
<i>Root Mean Squared Error</i>	Akar dari MSE yang menunjukkan ukuran kesalahan dalam unit asli data.
<i>Overfitting</i>	Kondisi di mana model terlalu cocok dengan data latih sehingga kurang mampu memprediksi data baru.
<i>Streamlit Cloud</i>	Platform berbasis cloud yang digunakan untuk membuat dan membagikan aplikasi machine learning berbasis web.

INTISARI

Produksi padi di Provinsi Jawa Tengah mengalami fluktuasi setiap tahun, mempengaruhi ketahanan pangan dan distribusi hasil pertanian. Oleh karena itu, diperlukan metode prediksi yang akurat untuk membantu pemangku kepentingan dalam perencanaan sektor pertanian dan pengambilan keputusan strategis. Penelitian ini menerapkan algoritma *Linear Regression* untuk memprediksi produksi padi berdasarkan data historis 2014–2023 dari website resmi Dinas Pertanian dan Perkebunan Provinsi Jawa Tengah. Model dikembangkan dengan regresi linear berganda menggunakan variabel luas tanam, luas panen, dan produktivitas. Proses penelitian mencakup pengumpulan data, *preprocessing*, pemodelan, optimasi dengan *GridSearchCV*, evaluasi model, dan *deployment* dalam aplikasi berbasis website menggunakan *Streamlit Cloud*.

Hasil optimasi dengan *GridSearchCV* menunjukkan akurasi *cross-validation* sebesar 98,26%, menegaskan kemampuan prediksi yang sangat baik. Evaluasi model menunjukkan nilai R^2 sebesar 0,9754 dengan MAE 0,0957, MSE 0,0307, dan RMSE 0,1753, menandakan kesalahan prediksi yang rendah. Model ini diimplementasikan sebagai aplikasi berbasis website dengan *Streamlit Cloud*. Penelitian selanjutnya disarankan untuk menambahkan variabel seperti curah hujan, suhu, dan jenis varietas padi atau dapat membandingkan model dengan algoritma lain seperti *Random Forest*, *Support Vector Regression*, atau *Long Short-Term Memory* guna meningkatkan akurasi prediksi.

Kata kunci: Prediksi Produksi Padi, Linear Regression, Machine Learning, Evaluasi Model, Hyperparameter Tuning, Streamlit Cloud.

ABSTRACT

Rice production in Central Java Province fluctuates every year, affecting food security and the distribution of agricultural products. Therefore, an accurate prediction method is needed to assist stakeholders in agricultural sector planning and strategic decision-making. This research applies the Linear Regression algorithm to predict rice production based on historical data 2014-2023 from the official website of the Central Java Province Agriculture and Plantation Office. The model was developed with multiple linear regression using the variables of planting area, harvest area, and productivity. The research process includes data collection, preprocessing, modeling, optimization with GridSearchCV, model evaluation, and deployment in a web-based application using Streamlit Cloud.

Optimization results with GridSearchCV showed a cross-validation accuracy of 98.26%, confirming excellent predictive ability. Model evaluation showed an R^2 value of 0.9754 with MAE 0.0957, MSE 0.0307, and RMSE 0.1753, indicating low prediction error. The model was implemented as a web-based application with Streamlit Cloud. Future research is recommended to add variables such as rainfall, temperature, and type of rice varieties or can compare the model with other algorithms such as Random Forest, Support Vector Regression, or Long Short-Term Memory to improve prediction accuracy.

Keywords: *Rice Production Prediction, Linear Regression, Machine Learning, Model Evaluation, Hyperparameter Tuning, Streamlit Cloud.*