

**ANALISIS PENGARUH LUAS PANEN DAN FAKTOR IKLIM
TERHADAP PRODUKSI PADI DI PULAU SUMATRA
MENGUNAKAN ALGORITMA GRADIENT BOOSTING
REGRESSION**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana

Program Studi Informatika



disusun oleh

SITI WAHYUNI

21.11.4068

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**ANALISIS PENGARUH LUAS PANEN DAN FAKTOR IKLIM
TERHADAP PRODUKSI PADI DI PULAU SUMATRA
MENGUNAKAN ALGORITMA GRADIENT BOOSTING
REGRESSION
SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat

Sarjana Program Studi Informatika



disusun oleh

SITI WAHYUNI

21.11.4068

Kepada

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

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS PENGARUH LUAS PANEN DAN FAKTOR IKLIM
TERHADAP PRODUKSI PADI DI PULAU SUMATRA MENGGUNAKAN
ALGORITMA GRADIENT BOOSTING REGRESSION**

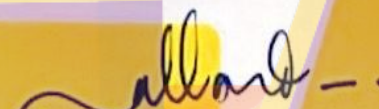
yang disusun dan diajukan oleh

Siti Wahyuni

21.11.4068

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 13 Maret 2025

Dosen Pembimbing,


Akhmad Dahlan, S.Kom., M.Kom

NIK. 190302174

HALAMAN PENGESAHAN

SKRIPSI

**ANALISIS PENGARUH LUAS PANEN DAN FAKTOR IKLIM
TERHADAP PRODUKSI PADI DI PULAU SUMATRA MENGGUNAKAN
ALGORITMA GRADIENT BOOSTING REGRESSION**

yang disusun dan diajukan oleh
Siti Wahyuni

21.11.4068

Telah dipertahankan di depan Dewan Penguji
pada tanggal 13 Maret 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Akhmad Dahlan, S.Kom., M.Kom
NIK. 190302174



M. Nuraminudin, M.Kom
NIK. 190302408



Arif Nur Rohman, M.Kom
NIK. 190302684



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 13 Maret 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusrini, S.Kom., M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Siti Wahyuni
NIM : 21.11.4068

Menyatakan bahwa Skripsi dengan judul berikut:

Analisis Pengaruh Luas Panen Dan Faktor Iklim Terhadap Produksi Padi Di Pulau Sumatra Menggunakan Algoritma Gradient Boosting Regression

Dosen Pembimbing : Akhmad Dahlan, 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, 13 Maret 2025
Yang Menyatakan,



Siti Wahyuni

MOTTO

“Direndahkan dimata manusia, ditinggikan dimata Tuhan, *Prove Them Wrong*”

“Gonna fight and don’t stop, until you are proud”

“Selalu ada harga dalam sebuah proses. Nikmati saja lelah-lelah itu. Lebarakan lagi kesabaran itu. Semua yang Bapak dan Ibu beri adalah untuk menjadikan diriku serupa yang aku impikan, mungkin tidak akan selalu berjalan lancar. Tapi gelombang-gelombang itu yang nanti akan bisa aku ceritakan”

-Yun



HALAMAN PERSEMBAHAN

Alhamdulillah, segala puji bagi Allah SWT dengan segala kuasa dan nikmat-Nya, sehingga penulis dapat menyelesaikan skripsi ini dengan lancar dan tepat pada waktunya dengan hasil yang terbaik. Terima kasih juga tak lupa saya sampaikan kepada semua orang yang ikut membantu secara moril dan dalam segala hal. Skripsi ini penulis persembahkan kepada:

1. Orang tua penulis (Ibu Dian Preticia Mokodompit, Bapak Ade Nurmansyah, dan Bapak Syaefuddin Arsyad) lalu Adik-adik Penulis yang senantiasa memberikan doa, nasihat dan kasih sayang, dukungan baik moril maupun materi sehingga dalam perkuliahan dan penyelesaian skripsi ini dapat terlaksana dengan baik.
2. Keluarga besar yang selalu memberikan support dan nasihat, serta dukungan dalam bentuk apapun.
3. Bapak Ibu dosen yang telah memberikan banyak ilmu dan pengalaman selama masa perkuliahan.
4. Bapak Akhmad Dahlan, S.Kom., M.Kom yang telah membimbing saya selama penelitian skripsi.
5. Partner Penulis (Siti Wahyuni) yang selalu memberikan motivasi dan kasih sayang dalam segala hal sehingga penulis dapat menyelesaikan skripsi ini dengan baik.
6. Sahabat- sahabat penulis (Faadhil Asyraf, Widi Widana Pandu Sunarya, Saddam Akbari Hasibuan, Hafiz, dan Rizal).
7. Teman-teman Camp Faadhil yang telah membantu dan berbagai kesenangan maupun suka duka selama perkuliahan.
8. Semua pihak yang tidak bisa penulis sebutkan satu per satu, terima kasih atas doa-doa dan dukungannya selama ini sehingga skripsi ini dapat terselesaikan.

KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Allah SWT, karena atas rohmat-Nya lah penulis dapat menyelesaikan penulisan skripsi ini dengan tepat pada waktunya yang berjudul “ANALISIS PENGARUH LUAS PANEN DAN FAKTOR IKLIM TERHADAP PADI DI PULAU SUMATRA MENGGUNAKAN ALGORITMA GRADIENT BOOSTING REGRESSION”. Shalawat serta salam tak lupa penulis hanturkan kepada baginda kita Nabi Besar Rasulullah Muhammad SAW, para keluarganya, sahabatnya, dan semoga syafaat beliau sampai kepada kita semua selaku umatnya.

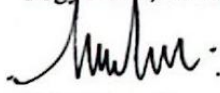
Penyusunan skripsi ini merupakan syarat kelulusan pendidikan pada program studi strata-1 (S1) Informatika Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.

Dalam penyusunan skripsi ini tentunya melibatkan banyak pihak yang telah membantu penulis sepenuh hati baik moril maupun materil. Oleh karena itu penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Bapak Prof. Dr. M. Suyanto, MM selaku Rektor Universitas Amikom Yogyakarta.
2. Bapak Hanif Al Fatta, S.Kom., M.Kom selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Ibu Windha Mega Pardnya Dhuhita, M.Kom selaku Kepala Program Studi Informatika Universitas Amikom Yogyakarta.
4. Bapak Akhmad Dahlan, S.Kom., M.Kom selaku Dosen Pembimbing Skripsi penulis yang selalu penuh perhatian dan kesabaran untuk memberikan masukan kepada penulis selama penyusunan skripsi ini.
5. Bapak M. Nuraminudin, M.Kom selaku Dosen Penguji 1 sidang skripsi dan Bapak Arif Nur Rohman, M.Kom selaku Dosen Penguji 2 sidang skripsi yang telah memberikan arahan sehingga skripsi ini menjadi lebih baik.

6. Seluruh Dosen dan Staf Universitas Amikom Yogyakarta yang selalu mengarahkan, mendampingi, dan memotivasi mahasiswa dalam proses penyelesaian skripsi ini.
7. Bapak Ferian Fauzi Abdulloh, S.Kom, M.Kom yang telah memberikan arahan, ilmu, dan pengalaman serta motivasi kepada saya, yang menjadi bekal saya selama perkuliahan.
8. Yang teristimewa kedua orang tuaku tercinta, Bapak Suyadi, Ibu Tarwiyah dan Adik saya Zani Salsabila yang senantiasa mendoakan, mencurahkan kasih sayang, serta memberikan dukungan dan semangat tanpa henti. Berkat doa serta pengorbanan Bapak dan Ibu, saya dapat menyelesaikan skripsi ini dengan baik. Semoga Bapak dan Ibu senantiasa diberi kesehatan rohani maupun jasmani serta seegala kebaikan yang telah diberikan dibalas dengan keberkahan dan limpahan rahmat dari Allah SWT.
9. Saya menyampaikan rasa terima kasih yang tulus kepada partner saya, Ari Valdy Pratama Arsyad, yang telah menemani dalam proses penyusunan skripsi ini. Dukungan, kerja sama, dan semangat yang diberikan sangat berarti dalam menghadapi berbagai tantangan selama penelitian. Apresiasi yang tinggi juga saya sampaikan atas dedikasi, motivasi, serta kebersamaan yang telah terjalin. Semoga segala kebaikan yang telah diberikan mendapatkan balasan yang berlimpah, dan hubungan baik ini tetap terjalin di masa yang akan datang.
10. Sahabat baik saya Ahita, Shera, Wawa, Tika, dan Idan yang telah membantu dan menemani susah, senang dalam keadaan saya selama masa kuliah.
11. Terimakasih untuk diri sendiri, karena telah bertahan sampai titik ini. Terimakasih karena telah begitu sabar serta tidak menyerah meskipun keadaan begitu sulit.

Yogyakarta, 10 Maret 2025


Siti Wahyuni

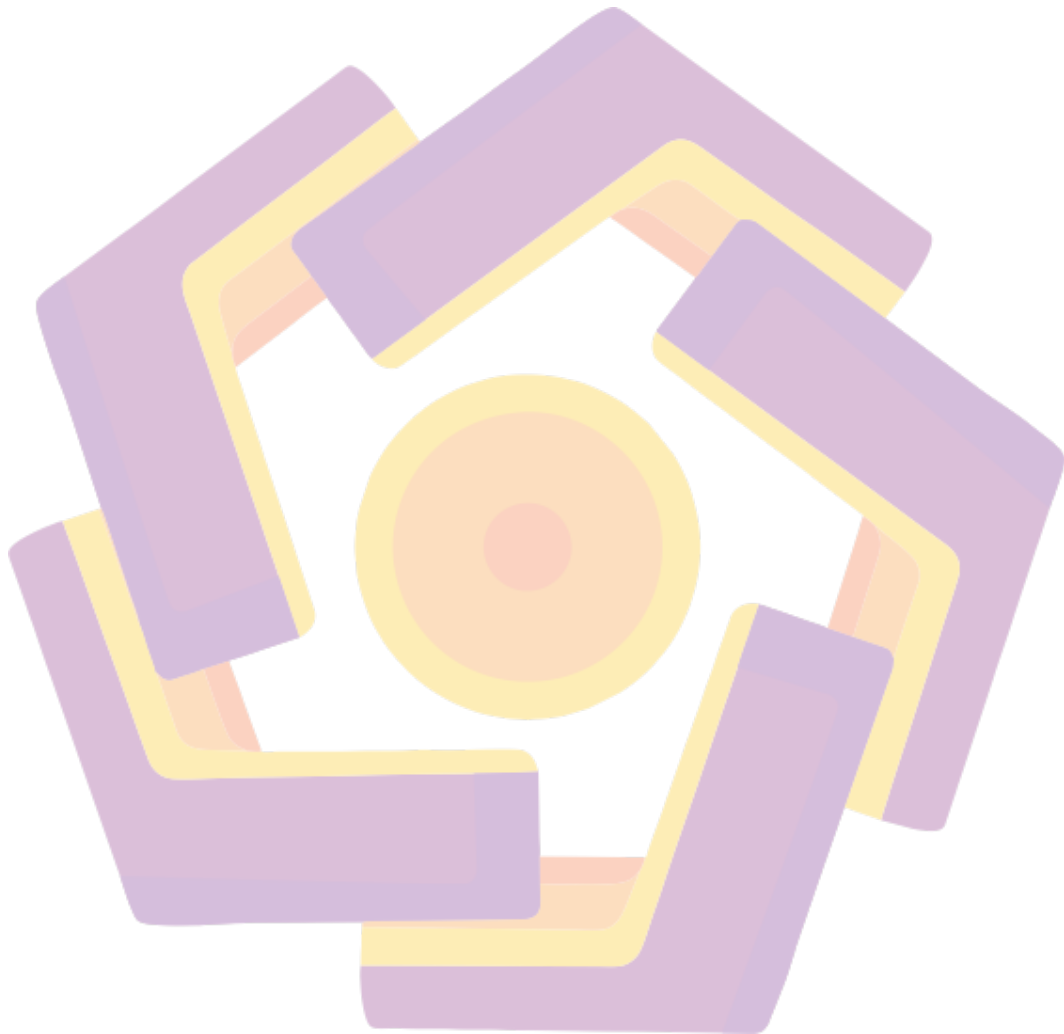
DAFTAR ISI

HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN.....	iv
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	v
MOTTO	vi
HALAMAN PERSEMBAHAN	vii
KATA PENGANTAR	viii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR.....	xv
DAFTAR TEXT CODE	xvii
DAFTAR LAMBANG DAN SINGKATAN	xix
DAFTAR ISTILAH.....	xx
INTISARI	xxi
<i>ABSTRACT</i>	xxii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	2
1.5 Manfaat Penelitian	2
1.6 Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5
2.2 Dasar Teori	11

2.2.1 Produksi Padi	11
2.2.2 Data Mining.....	11
2.2.3 Python	11
2.2.3.1 Pandas	12
2.2.3.2 Numpy.....	12
2.2.3.3 Scikit-Learn (SKlearn).....	12
2.2.4 Corelation Matrix.....	13
2.2.5 Regression.....	13
2.2.6 Gradient Boosting Regression	14
2.2.7 Preprocessing Data.....	14
2.2.7.1 Cleaning Data	15
2.2.7.2 One Hot Encoding	15
2.2.7.3 Feature Selection.....	15
2.2.7.4 Feature Scaling.....	15
2.2.8 Hyperparameter Tuning.....	15
2.2.8.1 GridSearch CV.....	16
2.2.8.2 Randomized Search CV	16
2.2.9 Evaluasi.....	17
2.2.9.1 Mean Absolute Error (MAE).....	17
2.2.9.2 Mean Squared Error (MSE).....	18
BAB III METODE PENELITIAN	19
3.1 Objek Penelitian	19
3.2 Alur Penelitian	19
3.2.1 Pengumpulan Data.....	20
3.2.2 Preprocessing	20
3.2.2.1 Cleaning	21

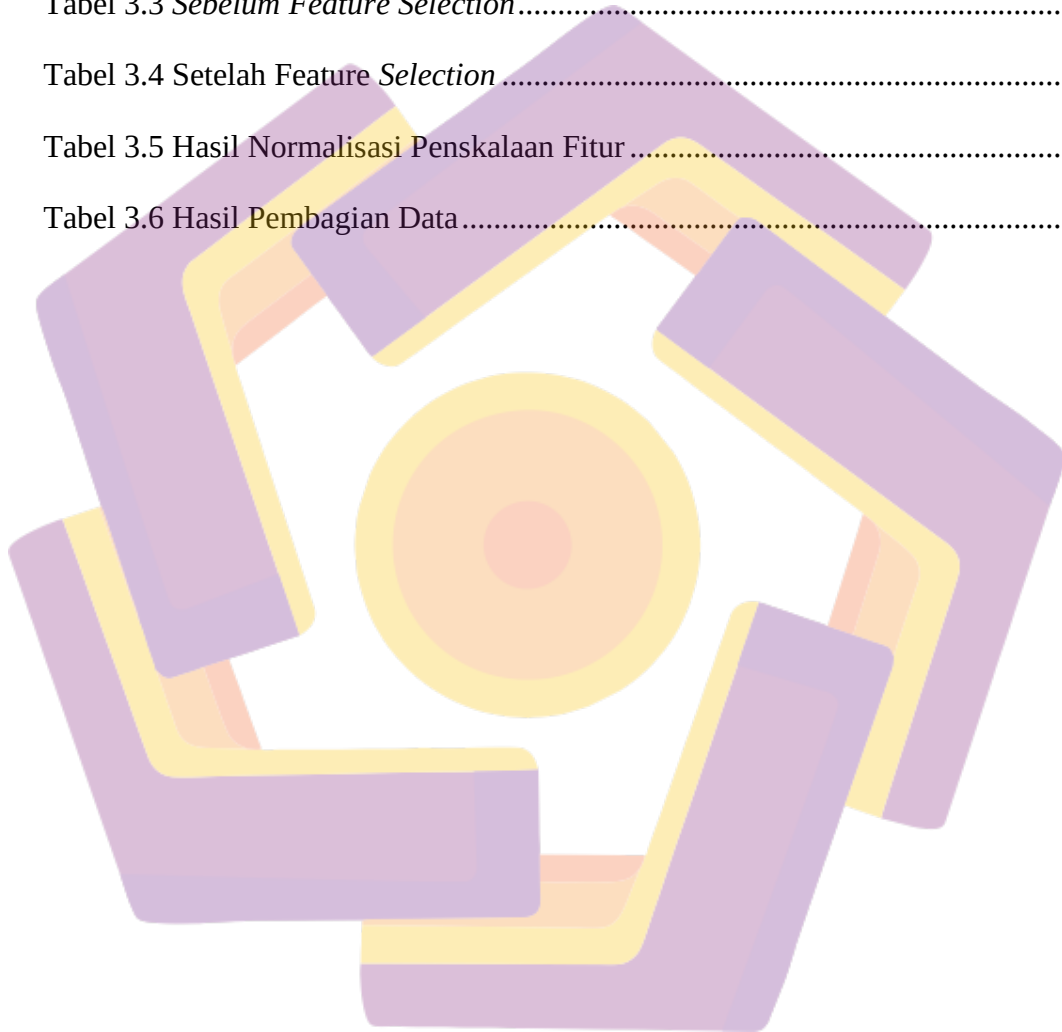
3.2.2.2 One Hot Encoding.....	21
3.2.2.3 <i>Feature Selection</i>	23
3.2.2.4 <i>Feature Scaling</i>	24
3.2.3 <i>Spilt Data</i>	25
3.2.4 <i>Gradient Boosting Regression</i>	26
3.2.5 Evaluasi Model	26
3.2.5.1 Evaluasi Model Mneggunakan <i>R2-Score</i>	26
3.2.5.2 Evaluasi Model Mneggunakan <i>Mean Absolute Error</i>	27
3.2.5.3 Evaluasi Model Mneggunakan <i>Mean Squared Error</i>	27
BAB IV HASIL DAN PEMBAHASAN	28
4.1 Implementasi Sistem.....	28
4.1.1 Hardware dan Software.....	28
4.1.2 Pengumpulan Data	28
4.1.3 Ekplorasi Data Analisis.....	30
4.1.4 <i>Preprocessing Data</i>	37
4.1.4.1 <i>Cleaning</i>	37
4.1.4.2 <i>One Hot Encoding</i>	38
4.1.4.3 <i>Feature Selection</i>	39
4.1.4.4 <i>Feature Scaling</i>	40
4.1.5 Split Data.....	43
4.1.6 Gradient Boosting Regression.....	45
4.1.7 Evaluasi Model.....	53
4.1.7.1 Evaluasi Model Menggunakan <i>R2-score</i>	54
4.1.7.2 Evaluasi Model Menggunakan <i>Mean Absolute Error</i>	54
4.1.7.3 Evaluasi Model Menggunakan <i>Mean Squared Error</i>	54
BAB V KESIMPULAN.....	55

5.1 Kesimpulan	55
5.2 Saran	56
REFERENSI	57



DAFTAR TABEL

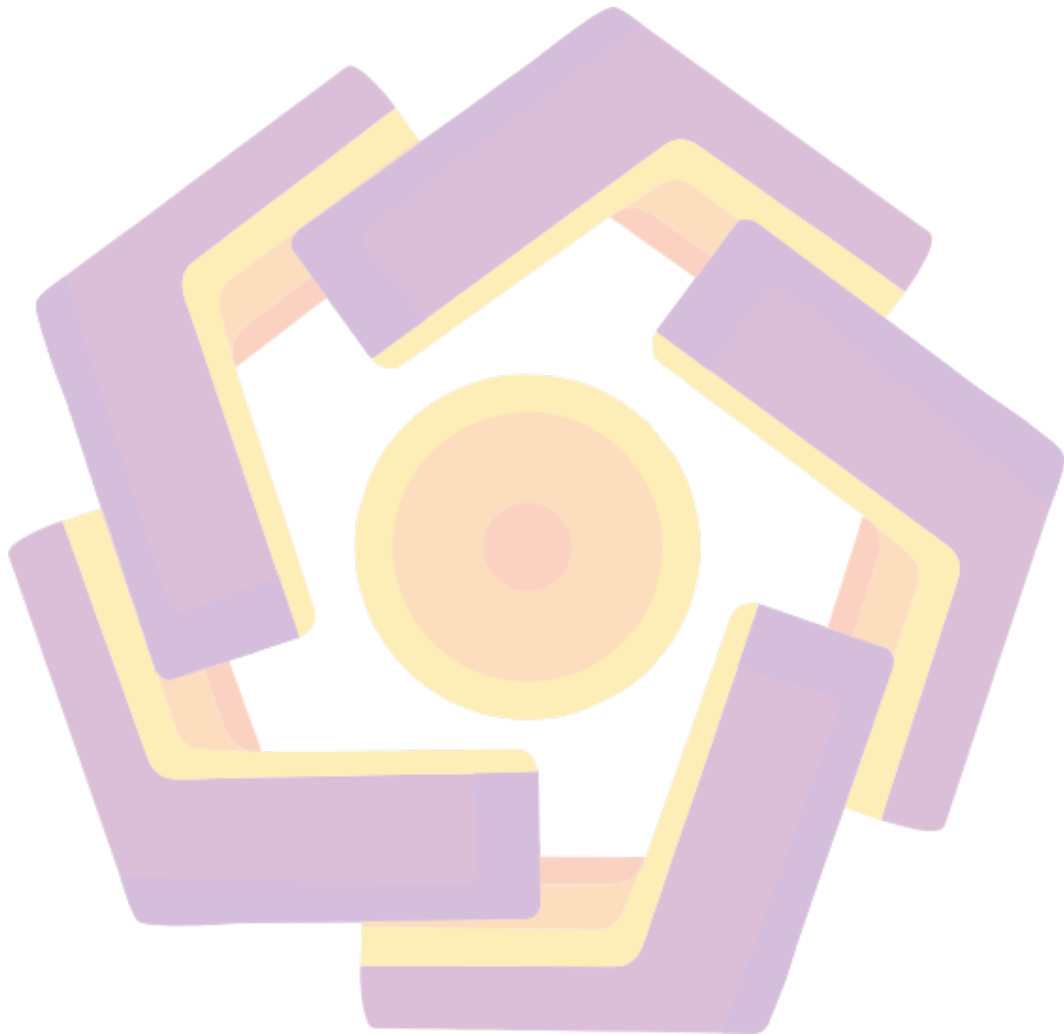
Tabel 2.1 Keaslian Penelitian.....	8
Tabel 3.1 Proses sebelum dilakukan <i>One Hot Encoding</i>	21
Tabel 3.2 Proses setelah dilakukan <i>One Hot Encoding</i>	21
Tabel 3.3 <i>Sebelum Feature Selection</i>	22
Tabel 3.4 Setelah <i>Feature Selection</i>	22
Tabel 3.5 Hasil Normalisasi Penskalaan Fitur	23
Tabel 3.6 Hasil Pembagian Data.....	24



DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian.....	18
Gambar 3.2 Hasil Pengumpulan Data.....	19
Gambar 3.3 Hasil <i>cleaning</i> data.....	20
Gambar 4.1 Proses Upload File	28
Gambar 4.2 Output Tampilan Data.....	29
Gambar 4.3 Statistik Deskriptif.....	30
Gambar 4.4 Output Produksi Padi Tahunan	31
Gambar 4.5 Suhu Rata-rata.....	32
Gambar 4.6 Output Suhu Rata-rata berdasarkan Tahun	32
Gambar 4.7 Output Curah Hujan	33
Gambar 4.8 Output Kelembapan	34
Gambar 4.9 Korelasi Matriks antara Variabel	35
Gambar 4.11 Output <i>One Hot Encoding</i>	37
Gambar 4.12 <i>Feature Selection</i>	38
Gambar 4.13 Output x_{train}	40
Gambar 4.14 Output Standarisasi x_{train} dan y_{train}	41
Gambar 4.15 Output Mencetak Nilai Min dan Max x_{train}	41
Gambar 4.16 Output Mencetak Nilai Min dan Max y_{train}	42
Gambar 4.17 Output Split Data.....	43
Gambar 4.18 Output <i>Gradient Boosting Regression</i>	45
Gambar 4.19 Output Visualisasi Prediksi	46
Gambar 4.20 Output Cek Akurasi Model	47
Gambar 4.21 Output <i>Hyperparameter Tuning</i>	49

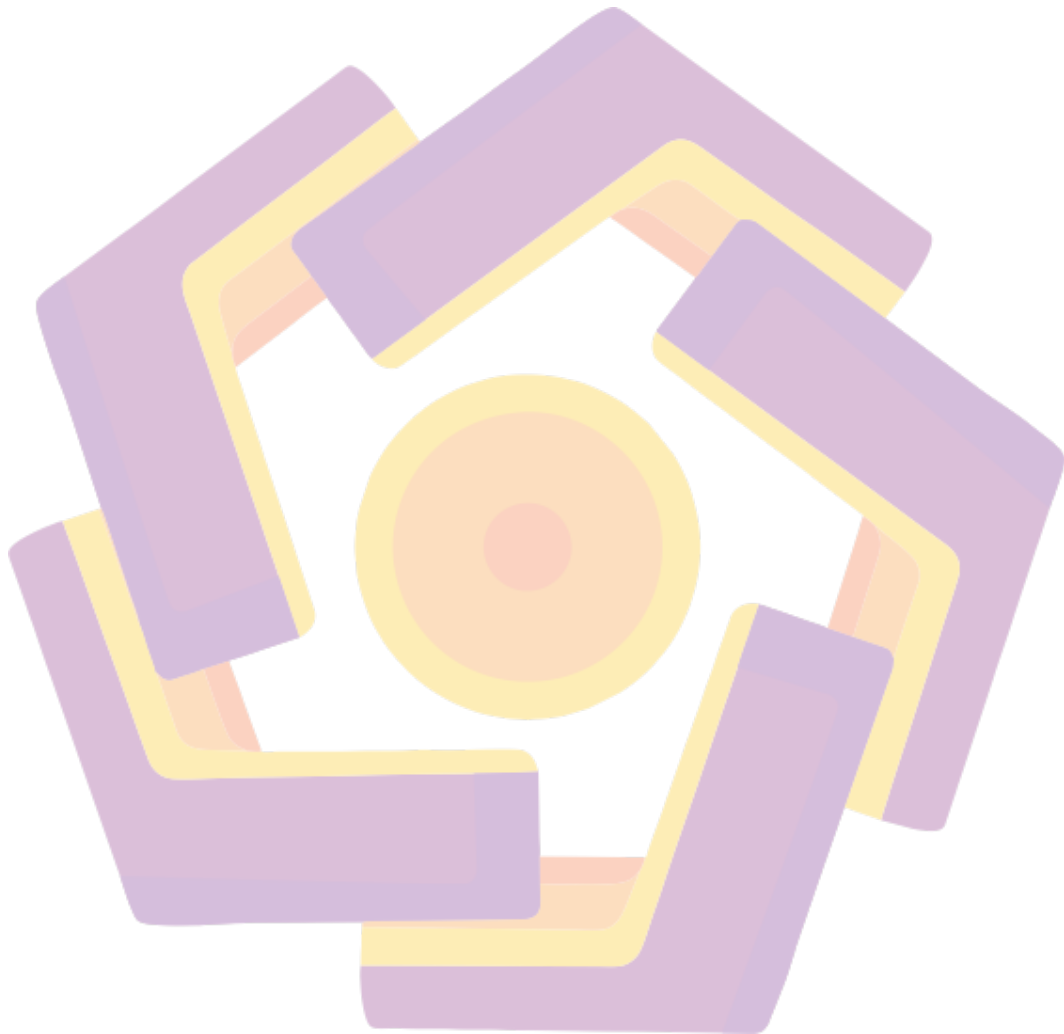
Gambar 4.22 Output <i>Fit New Model</i>	50
Gambar 4.23 Output Visualisasi Prediksi setelah <i>Fit New Model</i>	50
Gambar 4.24 Output <i>Hyperparameter</i> setelah <i>Fit New Model</i>	51
Gambar 4.25 Output Evaluasi Model.....	53



DAFTAR TEXT CODE

Text Code 4.1 Import Library	28
Text Code 4.2 Membaca Dataset	29
Text Code 4.3 Menampilkan Dalam Bentuk Diagram.....	30
Text Code 4.4 Script Code Produksi Padi Tahunan.....	31
Text Code 4.5 Suhu Rata-rata berdasarkan Tahun.....	32
Text Code 4.6 Curah hujan	33
Text Code 4.7 Kelembapan.....	34
Text Code 4.8 Mengecek <i>Missing Value</i>	36
Text Code 4.9 One Hot Encoding.....	37
Text Code 4.10 <i>Feature Selection</i>	38
Text Code 4.11 Feature Scaling.....	39
Text Code 4.12 x_train.....	39
Text Code 4.13 Setelah Standarisasi x_train dan y_train	40
Text Code 4.14 Mencetak Nilai Min dan Max x_train	41
Text Code 4.15 Mencetak Nilai Min dan Max y_train	42
Text Code 4.16 Split Data.....	43
Text Code 4.17 Library Python yang Diperlukan.....	44
Text Code 4.18 <i>Gradient Boosting Regression</i>	45
Text Code 4.19 Visualisasi Prediksi	45
Text Code 4.20 Mengecek Akurasi Model	47
Text Code 4.21 <i>Hyperparameter Tuning</i>	48
Text Code 4.22 <i>Fit New Model</i>	49
Text Code 4.23 Visualisasi Prediksi setelah <i>Fit New Model</i>	50

Text Code 4.25 <i>Hyperparameter Tunning</i> setelah <i>Fit New Model</i>	51
Text Code 4.26 Evaluasi Model.....	52



DAFTAR LAMBANG DAN SINGKATAN

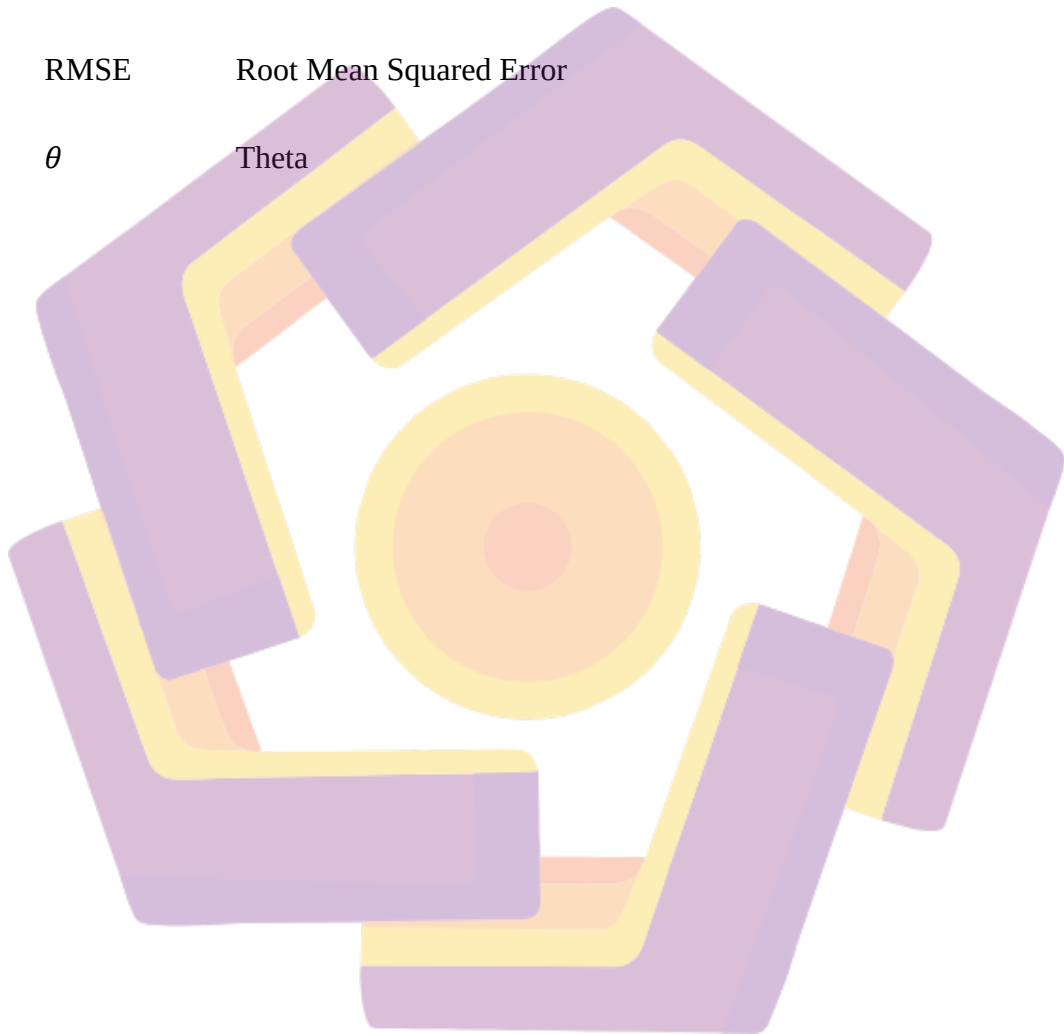
GBR Gradient Boosting Regression

MAE Mean Absolute Error

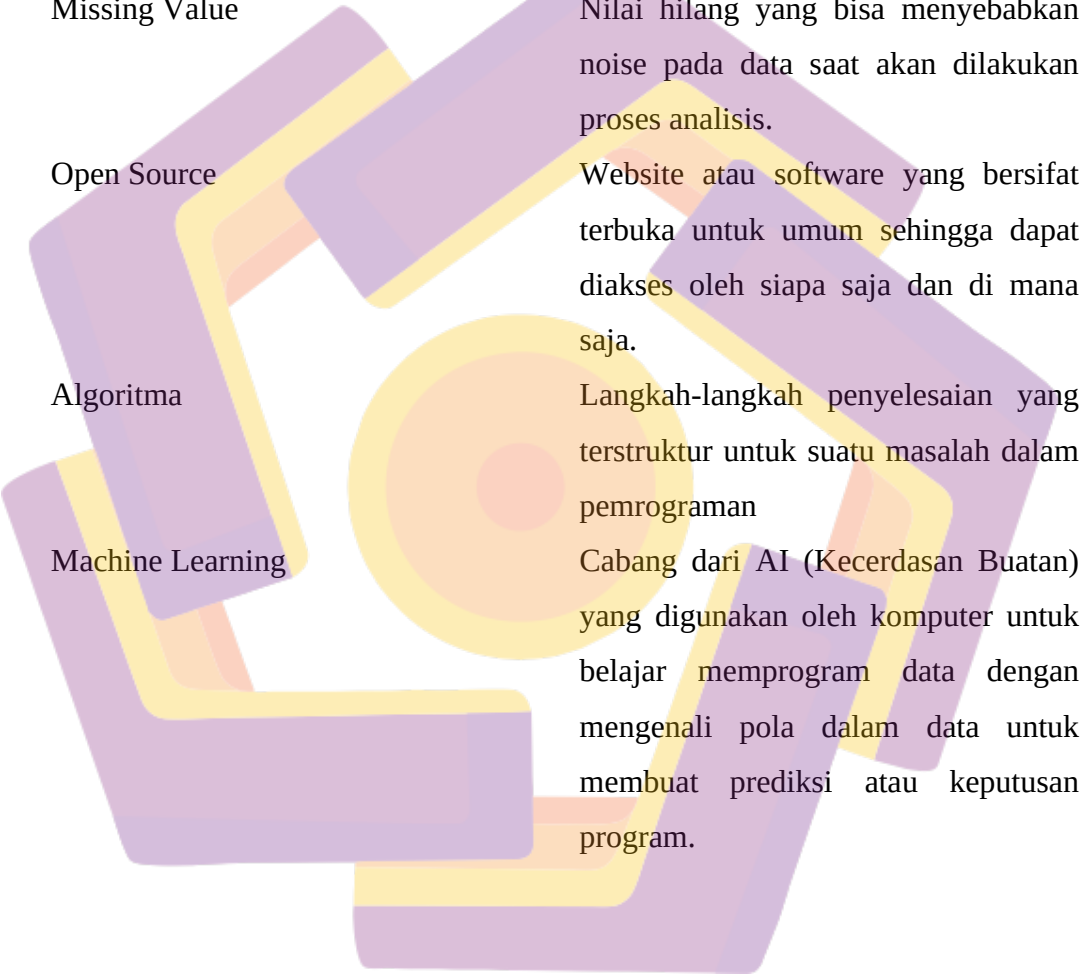
MSE Mean Squared Error

RMSE Root Mean Squared Error

θ Theta



DAFTAR ISTILAH



Library	Sekumpulan pustaka yang menyediakan berbagai macam modul dan fungsi untuk melakukan beberapa tugas pemrograman dan komputasi maupun keperluan analisis.
Missing Value	Nilai hilang yang bisa menyebabkan noise pada data saat akan dilakukan proses analisis.
Open Source	Website atau software yang bersifat terbuka untuk umum sehingga dapat diakses oleh siapa saja dan di mana saja.
Algoritma	Langkah-langkah penyelesaian yang terstruktur untuk suatu masalah dalam pemrograman
Machine Learning	Cabang dari AI (Kecerdasan Buatan) yang digunakan oleh komputer untuk belajar memprogram data dengan mengenali pola dalam data untuk membuat prediksi atau keputusan program.

INTISARI

Penelitian ini bertujuan untuk menganalisis pengaruh luas panen dan faktor iklim (seperti curah hujan, kelembapan, dan suhu rata-rata) terhadap produksi padi di Pulau Sumatra dengan menggunakan algoritma *Gradient Boosting Regression* (GBR). Data yang digunakan dalam penelitian ini diperoleh dari platform *Kaggle*, yang mencakup data produksi padi dari tahun 1993 hingga 2020 di delapan provinsi di Sumatra. Padi merupakan produksi tanaman di Indonesia yang menempati posisi ketiga di dunia, karena upaya untuk meningkatkan produksi pangan dan industri serta meningkatkan produksi ekspor, pertanian di Indonesia menjadi sektor yang sangat penting.

Untuk melakukan analisis produksi padi, peneliti menggunakan metode penelitian eksperimental. Pengumpulan data didapatkan melalui observasi terkait produksi padi dan faktor iklim yang memengaruhinya, berupa dataset yang didapatkan dari *kaggle*. Dataset yang didapatkan di analisis menggunakan algoritma *Gradient Boosting Regression*, karena kemampuannya untuk membuat model analisis yang akurat. Faktor produksi padi sangat dipengaruhi oleh sejumlah variabel, seperti luas panen dan kondisi iklim.

Hasil penelitian menunjukkan bahwa model GBR yang dioptimalkan dengan *hyperparameter tuning* mampu memprediksi produksi padi dengan akurasi yang cukup baik. Nilai R^2 -score yang diperoleh adalah 86.21%, yang menunjukkan bahwa model mampu menjelaskan 86.21% variasi dalam data. Nilai MAE sebesar 265.742 dan MSE sebesar 168.603 menunjukkan adanya kesalahan prediksi yang cukup besar, terutama karena sensitivitas MSE terhadap *outlier*.

Kata Kunci : Machine Learning, Produksi Padi, Luas Panen, Faktor Iklim, Gradient Boosting Regression.

ABSTRACT

This study aims to analyze the effect of harvest area and climatic factors (such as rainfall, humidity, and average temperature) on rice production in Sumatra Island using the Gradient Boosting Regression (GBR) algorithm. The data used in this study was obtained from the Kaggle platform, which includes rice production data from 1993 to 2020 in eight provinces in Sumatra. Rice is a crop production in Indonesia that ranks third in the world, due to efforts to increase food and industrial production and increase export production, agriculture in Indonesia is a very important sector.

To analyze rice production, researchers used an experimental research method. Data collection is obtained through observations related to rice production and climatic factors that affect it, in the form of datasets obtained from kaggle. The dataset obtained is analyzed using the Gradient Boosting Regression algorithm, because of its ability to create an accurate analysis model. Rice production factors are strongly influenced by a number of variables, such as harvest area and climatic conditions.

The results showed that the GBR model optimized with hyperparameter tuning was able to predict rice production with fairly good accuracy. The R^2 -score value obtained was 86.21%, indicating that the model was able to explain 86.21% of the variation in the data. The MAE value of 265,742 and MSE of 168,603 indicate a considerable prediction error, mainly due to the sensitivity of MSE to outliers.

Keywords: *Machine Learning, Rice Production, Harvested Area, Climate Factors, Gradient Boosting Regression.*