

**ANALISIS PREDIKSI HASIL PANEN KELAPA SAWIT DI PT
SINAR PANDAWA MENGGUNAKAN METODE
ALGORITMA C4.5 DAN REGRESI LINEAR**

SKRIPSI

Diajukan Untuk Memperolah Gelar Sarjana (S1)
Pada Program Studi Sistem Informasi Fakultas Sains dan Teknologi
Universitas Labuhanbatu



OLEH:
YURI PINESTI TANTIA
2109100089

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS LABUHANBATU
RANTAUPRAPAT
TAHUN 2025**

LEMBAR PERSETUJUAN NASKAH SKRIPSI

Judul Skripsi : ANALISIS PREDIKSI HASIL PANEN KELAPA
SAWIT DI PT SINAR PANDAWA MENGGUNAKAN
METODE ALGORITMA C4.5 DAN REGRESI LINEAR

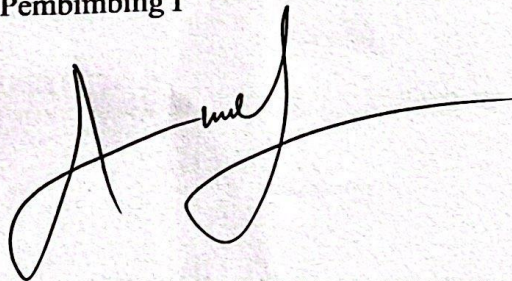
Nama : YURI PINESTI TANTIA

NPM : 2109100089

Prodi : SISTEM INFORMASI

Disetujui pada tanggal : 26 Agustus 2025

Pembimbing I



(Angga Putra Juledi, S.Kom., M.Kom)
NIDN. 119079401

Pembimbing II



(Rahma Muti'ah, S.Psi., M.Psi)
NIDN. 114068501

LEMBAR PENGESAHAN NASKAH SKRIPSI

JUDUL : ANALISIS PREDIKSI HASIL PANEN KELAPA SAWIT DI PT.
SINAR PANDAWA MENGGUNAKAN METODE ALGORITMA
C4.5 DAN REGRESI LINEAR

NAMA : YURI PINESTI TANTIA
NPM : 2109100089
PROGRAM STUDI : SISTEM INFORMASI

Telah Diuji Dan Dinyatakan Dalam Ujian Sarjana
Pada Tanggal 26 Agustus 2025

TIM PENGUJI

Penguji I (Ketua)

Nama : Sudi Suryadi, S. Kom., M. Kom
NIDN : 0128027903

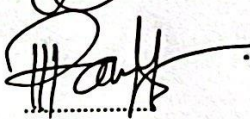
Penguji II (Anggota)

Nama : Angga Putra Juledi, S. Kom., M. Kom
NIDN : 0119079401

Penguji III (Anggota)

Nama : Rahma Muti'ah, S.Psi, M.Psi
NIDN : 0115028404

Tanda Tangan



Rantauprapat, 26 Agustus 2025

Dekan,
Fakultas Sains dan Teknologi



(Dr. Iwan Purnama, S. Kom., M. Kom)
NIDN. 0112029202

Ka. Prodi Studi,
Sistem Informasi



(Budianto Bangun, S. Kom., M. Kom)
NIDN. 0110058601

LEMBAR PERNYATAAN

Yang bertanda tangan dibawah ini :

Nama : YURI PINESTI TANTIA

NPM : 2109100089

Judul Skripsi : ANALISIS PREDIKSI HASIL PANEN KELAPA SAWIT DI PT. SINAR
PANDAWA MENGGUNAKAN METODE ALGORITMA C4.5 DAN
REGRESI LINEAR

Dengan ini penulis menyatakan bahwa skripsi ini disusun sebagai syarat untuk memperoleh gelar sarjana pada Program Studi Sistem Informasi Fakultas Sains dan Teknologi Universitas Labuhanbatu adalah hasil karya tulis penulis sendiri. Semua kutipan maupun rujukan dalam penulisan skripsi ini telah penulis cantumkan sumbernya dengan benar sesuai dengan ketentuan yang berlaku.

Jika di kemudian hari ternyata ditemukan seluruh atau sebagian skripsi ini bukan hasil karya penulis atau plagiat, penulis bersedia menerima sanksi pencabutan gelar akademik yang disandang dan sanksi-sanksi lainnya sesuai dengan peraturan perundang-undangan yang berlaku.

Rantauprapat, 26 Agustus 2025

Yang Membuat Pernyataan,



YURI PINESTI TANTIA

NPM. 2109100089

ABSTRAK

Perkebunan kelapa sawit memiliki peran penting dalam perekonomian Indonesia, sehingga prediksi hasil panen yang akurat sangat dibutuhkan untuk mendukung perencanaan dan efisiensi produksi. Saat ini, PT Sinar Pandawa masih menggunakan metode manual dalam memperkirakan hasil panen, yang seringkali kurang akurat dan subjektif. Penelitian ini bertujuan untuk membandingkan dua metode prediksi, yaitu algoritma C4.5 dan regresi linear. Data yang digunakan berupa data historis hasil panen bulanan selama satu tahun, dengan variabel curah hujan, jumlah pupuk, dan hasil panen. Analisis menggunakan algoritma C4.5 menghasilkan model pohon keputusan dengan kategori (rendah, sedang, tinggi) dan nilai evaluasi RMSE sebesar 60.779 serta MAE sebesar 56.425. Sementara itu, model regresi linear menghasilkan persamaan $Y = 1159.98 - 0.0702X_1 + 0.0104X_2$, dengan nilai RMSE sebesar 3.545, MAE sebesar 1.039, dan R^2 sebesar 0.519. Hasil penelitian menunjukkan bahwa regresi linear lebih akurat dalam memprediksi hasil panen kelapa sawit dibandingkan algoritma C4.5, meskipun C4.5 tetap memiliki keunggulan dalam interpretasi aturan klasifikasi. Dengan demikian, regresi linear direkomendasikan sebagai metode prediksi utama, sedangkan C4.5 dapat digunakan sebagai pendukung dalam pengambilan keputusan berbasis kategori.

Kata Kunci: *Prediksi Hasil Panen, Kelapa Sawit, Algoritma C4.5, Regresi Linear, Data Mining.*

ABSTRACT

Palm oil plantations play an important role in Indonesia's economy, making accurate yield prediction essential to support production planning and efficiency. Currently, PT Sinar Pandawa still relies on manual methods to estimate yields, which are often inaccurate and subjective. This study aims to compare two prediction methods, namely the C4.5 algorithm and linear regression. The dataset used consists of monthly historical harvest data for one year, with rainfall, fertilizer usage, and yield as variables. Analysis using the C4.5 algorithm produced a decision tree model with categorical outputs (low, medium, high), resulting in an RMSE of 60.779 and an MAE of 56.425. Meanwhile, the linear regression model generated the equation $Y = 1159.98 - 0.0702X_1 + 0.0104X_2$, with an RMSE of 3.545, MAE of 1.039, and R^2 of 0.519. The results indicate that linear regression provides more accurate predictions of palm oil yields compared to the C4.5 algorithm, although C4.5 remains advantageous in providing easily interpretable classification rules. Therefore, linear regression is recommended as the primary prediction method, while C4.5 can be applied as a complementary tool for decision-making based on categorical outputs.

Keywords: *Yield Prediction, Palm Oil, C4.5 Algorithm, Linear Regression, Data Mining.*

KATA PENGANTAR

Puji dan syukur saya panjatkan kepada Allah SWT atas berkat Rahmat Hidayah dan Karunia-Nya memberikan kemudahan sehingga saya dapat menyelesaikan proposal dengan judul **Analisis Prediksi Hasil Panen Kelapa Sawit Di PT SINAR PANDAWA Menggunakan Metode Algoritma C4.5 Dan Regresi Linear**

Penyusunan proposal ini tidak lepas dari bantuan, bimbingan, dan dukungan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis mengucapkan terimakasih kepada:

1. Bapak Alm. Dr. H. Amarullah Nasution, SE., MBA. selaku Pendiri Yayasan Universitas Labuhanbatu.
2. Bapak Halomoan Nasution, S.H., M.H selaku Ketua Yayasan Universitas Labuhanbatu.
3. Bapak Assoc. Prof. Ade Parlaungan Nasution, S.E., M.SI., Ph.D. selaku Rektor Universitas Labuhanbatu.
4. Bapak Dr. Iwan Purnama, S.Kom., M.Kom. selaku Dekan Fakultas Sains dan Teknologi Universitas Labuhanbatu
5. Bapak Budianto Bangun, S.Sos., M.Kom. selaku Ka. Prodi Sistem Informasi Univesitas Labuhanbatu
6. Bapak Angga Putra Juledi, S.Kom., M.Kom. selaku pembimbing pertama yang telah meluangkan waktunya untuk memberikan arahan, saran, pentunjuk dan motivasi dalam penulisan proposal dan skripsi penulis.
7. Ibu Rahma Muti'ah, S.Psi., M.Psi. selaku pembimbing kedua yang telah meluangkan waktunya untuk memberikan arahan, saran, pentunjuk dan motivasi dalam penulisan proposal dan skripsi penulis.
8. Bapak Sudi Suryadi, S.Kom., M.Kom. yang telah menjadi dosen penguji dalam seminar proposal dan skripsi penulis.

Demikianlah yang dapat penulis sampaikan. Penulis menyadari bahwa penulisan ini masih jauh dari kata sempurna. Maka dari itu, penulis mengharapkan saran dan kritik yang bersifat membangun untuk menyempurnakan penyusunan proposal ini. Semoga dapat bermanfaat bagi penulis dan pembaca. Kebenaran datangnya dari Allah dan kesalahan datangnya dari diri penulis. Semoga Allah SWT senantiasa melimpahkan Rahmat dan Ridho-Nya kepada kita semua.

Rantauprapat, 25 Agustus 2025
Penulis



YURI PINESTI TANTIA

DAFTAR ISI

LEMBAR PERSETUJUAN NASKAH SKRIPSI	i
LEMBAR PENGESAHAN NASKAH SKRIPSI	ii
LEMBAR PERNYATAAN	iii
ABSTRAK	iv
<i>ABSTRACT</i>	v
KATA PENGANTAR.....	vi
DAFTAR ISI	vii
DAFTAR GAMBAR	ix
DAFTAR TABEL	x
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Perumusan Masalah.....	2
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	4
1.6 Tinjauan Umum Objek Penelitian	4
1.7 Sistematika Penulisan	5
BAB II LANDASAN TEORI	7
2.1 Knowledge Discovery In Database	7
2.2 Data Mining.....	9
2.3 Algoritma C4.5	9
2.3.1 Penjelasan Algoritma C4.5 dan Rumusnya.....	9
2.3.2 Tahapan Kerja Algoritma C4.5.....	9
2.3.3 Rumus Utama:	10
2.4 Regresi Linear	11
2.5 Prediksi Kelapa Sawit.....	13
2.6 RapidMiner.....	13
2.7 Metodologi Penelitian	14
2.8 Penelitian Terdahulu.....	14
BAB III METODOLOGI PENELITIAN.....	17
3.1 Kerangka Kerja Penelitian.....	17
3.2 Pengumpulan Data.....	19
3.3 Pra-pemrosesan Data	19
3.4 Analisa dengan Algoritma C4.5	20
3.5 Analisa Regresi Linear	24
3.6 Perbandingan metode	27
BAB IV HASIL DAN PEMBAHASAN	29
4.1 Pengujian Metode C4.5	29
4.1.1 Pengujian Menggunakan RapidMiner	29
4.1.2 Langkah Pengujian	32
4.1.3 Hasil Pengujian	35
4.2 Pengujian Metode Regresi Linear	37
4.2.1 Data Pelatihan (Training Data)	38
4.2.2 Data Pengujian (Testing Data).....	39
4.2.3 Langkah Pengujian.....	40
4.2.4 Hasil Prediksi Regresi Linear	42
4.2.5 Persamaan Model Regresi.....	42

4.3 Perbandingan Hasil Prediksi.....	43
4.3.1 Perbandingan Jenis Output	44
4.3.2 Perbandingan Hasil Evaluasi	44
4.3.3 Kelebihan dan Kekurangan Masing-Masing Metode	45
4.3.4 Kesimpulan Perbandingan	45
BAB V KESIMPULAN DAN SARAN	47
5.1 Kesimpulan.....	47
5.2 Saran	48
DAFTAR PUSTAKA	49
LAMPIRAN	51
PERSEMBAHAN	56

DAFTAR GAMBAR

Gambar 2. 1 Tahapan Dalam KDD	7
Gambar 2. 2 Rapidminer	14
Gambar 3. 1 Kerangka Kerja	17
Gambar 3. 2 Pohon Keputusan.....	22
Gambar 4. 1 Aplikasi Altair RapidMiner 2025.1.0.....	30
Gambar 4. 2 Tampilan Awal Altair RapidMiner AI Studio	32
Gambar 4. 3 Import dan Pilih Data Excel	32
Gambar 4. 4 Pilih Rentang Kolom yang Digunakan.....	33
Gambar 4. 5 Tampilan <i>Change Rule</i>	34
Gambar 4. 6 Lokasi Penyimpanan Data Yang Akan Diproses	34
Gambar 4. 7 Tampilan Alur Proses Lengkap di RapidMiner (Decision Tree)	35
Gambar 4. 8 Hasil Prediksi dan Confidence Menggunakan Decision Tree.....	36
Gambar 4. 9 Visualisasi Pohon Keputusan (Decision Tree).....	37
Gambar 4. 10 <i>Import Data Training</i>	40
Gambar 4. 11 <i>Import Data Testing</i>	41
Gambar 4. 12 Tampilan <i>Design</i> di RapidMiner (Regresi Linear)	41
Gambar 4. 13 Hasil Prediksi Regresi Linear dengan Data Uji	42
Gambar 4. 14 Persamaan Regresi Linear yang Dibentuk oleh Sistem	43

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu	14
Tabel 2. 2 Penelitian Terdahulu	15
Tabel 2. 3 Penelitian Terdahulu	16
Tabel 3. 1 Pengumpulan Data	19
Tabel 3. 2 Klasifikasi Data	20
Tabel 3. 3 Nilai Entropy Curah Hujan	21
Tabel 3. 4 Nilai Entropy Total Dan Gain Curah Hujan	21
Tabel 3. 5 Nilai Entropy Pupuk	21
Tabel 3. 6 Nilai Entropy Total Dan Gain Pupuk.....	21
Tabel 3. 7 Prediksi Berdasarkan Pohon	22
Tabel 3. 8 Klasifikasi ke Nilai Numerik	23
Tabel 3. 9 Nilai AE, SE, MAE dan RMSE	24
Tabel 3. 10 Model Regresi	25
Tabel 3. 11 Model Regresi	25
Tabel 3. 12 Model Regresi	26
Tabel 3. 13 Evaluasi Model	27
Tabel 3. 14 Perbandingan Metode	28
Tabel 4. 1 Data yang telah dipraproses	30
Tabel 4. 2 Dataset Pelatihan Regresi Linear	38
Tabel 4. 3 Dataset Prediksi (Kolom Hasil Panen Kosong)	39
Tabel 4. 4 Perbandingan Jenis Output	44
Tabel 4. 5 Perbandingan Hasil Evaluasi	44
Tabel 4. 6 Kelebihan dan Kekurangan Metode.....	45