

**SKRIPSI**

**MODEL PREDIKSI BIAYA EKSPOR : STUDI PEMILIHAN FITUR DAN  
PERBANDINGAN ALGORITMA LINEAR REGRESSION DAN RANDOM  
FOREST REGRESSION**



**Disusun Oleh**

**Muhammad Rafli Pratama**

**2021230063**

**PROGRAM STUDI TEKNOLOGI INFORMASI**

**FAKULTAS TEKNIK**

**UNIVERSITAS DARMA PERSADA**

**JAKARTA**

**2025**

## LEMBAR BIMBINGAN



### UNIVERSITAS DARMA PERSADA

Jl. Taman Malaka Selatan, Pondok Kelapa, Jakarta Timur, Indonesia 13450  
Telp. (021) 8649051, 8649053, 8649057 Fax. (021) 8649052  
E-mail : [humas@unsada.ac.id](mailto:humas@unsada.ac.id) Home page : <http://www.unsada.ac.id>

#### Instrumen Monitoring Bimbingan Skripsi Program Studi Teknologi Informasi

Tahun Akademik : 2024/2025 Genap

NIM>Nama Mhs : 2021230063 / Muhammad Rafli Pratama  
Judul Skripsi : Model Prediksi Biaya Ekspor : Studi Pemulihan  
Fitur dan Perbandingan Algoritma Linear Regression dan Random Forest Regression  
Dosen Pembimbing : Ds. Linda Nur ARIFA

No	BAB Utama Skripsi dan BATAS WAKTU Bimbingan	Materi Yang dibahas saat Konsultasi	Tanggal Bimbingan	TTD Dosen
1				
2	BAB I PENDAHULUAN			
3	Paling lama upload: 9 Mei 2025	perlu revisi		
		Tanggal BAB I di ACC pembimbing =>	4/6/25	
4	BAB II LANDASAN TEORI			
5	Paling lama upload: 9 Mei 2025			
6		perlu revisi		
		Tanggal BAB II di ACC pembimbing =>	4/6/25	
7	BAB III METODOLOGI	Revisi (perbaikan dan)		
8		di bagian		
9	Paling lama upload : 16 Mei 2025	revisi perbaikan -		
		Tanggal BAB III di ACC pembimbing ->	16/6/25	

## LEMBAR PERBAIKAN REVISI



**UNIVERSITAS DARMA PERSADA**

Jl. Taman Malaka Selatan, Pondok Kelapa, Jakarta Timur, Indonesia 13450

Telp. (021) 8649051, 8649053, 8649057 Fax. (021) 8649052

E-mail : [humas@unsada.ac.id](mailto:humas@unsada.ac.id) Home page : <http://www.unsada.ac.id>

### LEMBAR REVISI - SIDANG SKRIPSI

NIM>Nama : 2021230063 – Muhammad Rafli Pratama  
Fakultas/Prodi : Teknik / Teknologi Informasi

No.	Keterangan Revisi	Dosen
	Perbaikan font disarabik	Yay
	Penulisan X Decision tree X class Diagram	20 Okt 2025

Mengetahui,

Ka Prodi Teknologi Informasi

Herianto, S.Pd., MT.

MONDUCUKU • TRILINGUAL • ENERGI TERBARUKAN



## LEMBAR PERNYATAAN

Saya yang bertanda tangan dibawah ini :

Nama : Muhammad Rafli Pratama

NIM : 2021230063

Fakultas : Teknik

Jurusan : Teknologi Informasi

Dengan ini saya menyatakan bahwa skripsi yang berjudul:

**“Model Prediksi Biaya Ekspor:Studi Pemilihan Fitur dan Perbandingan Algoritma Linear Regression dan Random Forest Regression”**

Menyatakan bahwa skripsi ini disusun secara mandiri, berdasarkan hasil survei, observasi, wawancara, serta merujuk pada berbagai referensi yang relevan dengan topik laporan ini.

Demikian pernyataan ini di buat dengan sesungguhnya.

Jakarta, 05 Agustus 2025



Muhammad Rafli pratama

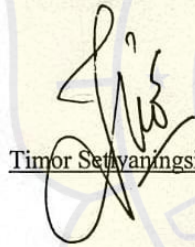
**LEMBAR PENGUJI SKRIPSI**

Laporan Skripsi yang berjudul:

**“Model Prediksi Biaya Ekspor: Studi Pemilihan Fitur dan Perbandingan  
Algoritma Linear Regression dan Random Forest Regression”**

ini telah diujikan pada tanggal “05 Agustus 2025”

Penguji I



Timor Setyaningsih, ST, MTI

Penguji II



Yan Sofyan Andhana Saputra, S.KOM., M.KOM

Penguji III



Dr. Linda Nur Afifa, ST, MT

**LEMBAR PENGESAHAN**

**“Model Prediksi Biaya Ekspor: Studi Pemilihan Fitur dan Perbandingan  
Algoritma Linear Regression dan Random Forest Regression”**

Disusun Oleh :

Nama : Muhammad Rafli Pratama

Nim : 2021230063

Pembimbing Skripsi



Dr.Linda Nur Afifa, ST,MT

Kajur Teknologi Informasi



Herianto, S.Pd., M.T

## KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas segala rahmat, karunia, dan kekuatan yang telah diberikan sehingga penulis dapat menyelesaikan penyusunan skripsi yang berjudul **“Model Prediksi Biaya Ekspor: Studi Pemilihan Fitur dan Perbandingan Algoritma Linear Regression dan Random Forest Regression”** ini dengan baik dan lancar.

Skripsi ini disusun sebagai salah satu syarat untuk menyelesaikan program Sarjana pada Program Studi Teknologi Informasi, Fakultas Teknik, Universitas Darma Persada. Penulisan skripsi ini tentu bukan hasil usaha sendiri, melainkan berkat dukungan, doa, bimbingan, dan semangat dari berbagai pihak yang sangat berarti bagi penulis.

Penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Tuhan Yang Maha Esa, atas segala petunjuk dan kekuatan dalam setiap proses yang penulis lalui.
2. Kedua orang tua tercinta, atas doa, kasih sayang, dukungan moral, dan semangat tanpa henti yang menjadi sumber motivasi terbesar dalam hidup penulis.
3. Bapak Dr. Ade Supriatna, M.T., selaku Dekan Fakultas Teknik Universitas Darma Persada.
4. Bapak Herianto, S.Pd., M.T. selaku Ketua Jurusan Teknologi Informasi Universitas Darma Persada.
5. Dosen pembimbing Dr.Linda Nur Afifa, ST,MT yang telah meluangkan waktu,

tenaga, dan pikiran untuk membimbing dan memberikan arahan selama proses penyusunan skripsi ini.

6. Dosen penguji yaitu Ibu Timor Setyaningsih,ST,MTI dan Bapak Yan Sofyan Andhana Saputra,S.KOM,M.KOM yang telah memberikan masukan dan koreksi demi penyempurnaan hasil penelitian ini.
7. Teman-teman seperjuangan di kampus yang telah menjadi bagian penting dalam proses pembelajaran dan penyusunan skripsi ini, serta memberikan semangat dan saling membantu selama proses berlangsung.
8. Terutama, untuk diriku sendiri yang tetap bertahan meski lelah, gelisah, dan seribu alasan untuk menyerah. Skripsi ini bukan sekadar ijazah, tapi bukti bahwa aku bisa. Semoga ini menjadi pengingat: tantangan apa pun bisa kuhadapi.

Penulis menyadari bahwa skripsi ini masih belum sempurna. Karena itu, saran dan masukan sangat diharapkan untuk perbaikan ke depannya. Semoga karya ini bermanfaat, terutama bagi yang tertarik pada penerapan machine learning di bidang logistik ekspor.

Jakarta, 05 Agustus 2025

Muhammad Rafli Pratama

# **Model Prediksi Biaya Ekspor: Studi Pemilihan Fitur Dan Perbandingan Algoritma Linear Regression Dan Random Forest Regression**

Muhammad Rafli Pratama, Linda Nur Afifa

Teknologi Informasi, Universitas Darma Persada

## **ABSTRAK**

Perhitungan biaya pengiriman ekspor yang masih dilakukan secara manual seringkali menyebabkan ketidakakuratan dan inefisiensi dalam proses logistik. Penelitian ini bertujuan membangun sistem prediksi biaya pengiriman ekspor berbasis machine learning untuk meningkatkan akurasi dan efisiensi perhitungan. Metode yang digunakan adalah pendekatan kuantitatif dengan membandingkan performa dua algoritma, yaitu Linear Regression dan Random Forest Regression, menggunakan data historis pengiriman dari PT Samudera Indonesia. Variabel yang dianalisis meliputi berat barang, volume, jenis komoditas, negara tujuan, jenis kapal, biaya asuransi, dan bea cukai. Evaluasi model dilakukan dengan metrik  $R^2$ , Mean Absolute Error (MAE), dan Root Mean Squared Error (RMSE). Hasil penelitian menunjukkan bahwa Random Forest Regression memberikan performa lebih baik dengan nilai  $R^2$  sebesar 0,7522, MAE 1.812, dan RMSE 2.277 dibandingkan Linear Regression ( $R^2$  0,6778, MAE 2.131, RMSE 2.597). Analisis feature importance mengungkapkan bahwa bea cukai dan biaya asuransi merupakan faktor paling berpengaruh terhadap biaya pengiriman. Sistem ini kemudian diimplementasikan dalam bentuk aplikasi berbasis web dengan antarmuka yang disesuaikan untuk tiga peran pengguna, yaitu admin, staf logistik, dan manajer. Dengan demikian, sistem prediksi ini dapat menjadi solusi efektif untuk meningkatkan akurasi perencanaan biaya ekspor dan mendukung pengambilan keputusan logistik yang lebih efisien.

**Kata Kunci:** Prediksi Biaya Ekspor, Machine Learning, Linear Regression, Random Forest, Sistem Berbasis Web.

## DAFTAR ISI

BAB I PENDAHULUAN.....	1
1.1 Latar Belakang dan Identifikasi Masalah.....	1
1.2 Rumusan Masalah.....	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	5
1.5 Manfaat Penelitian.....	6
1.6 Metode Penelitian.....	6
1.7 Sistematika Penulisan.....	10
BAB II LANDASAN TEORI.....	12
2.1 Tinjauan Pustaka.....	12
2.2.1 Biaya Pengiriman Ekspor.....	12
2.2.2 Machine Learning untuk Prediksi.....	15
2.2.3 Data Mining.....	15
2.2.4 Linear Regression.....	16
2.2.5 Random Forest Regression.....	18
2.2.6 Evaluasi Akurasi Perhitungan.....	20
2.2.7 CRISP-DM (Cross Industry Standard Process for Data Mining).....	22
2.2.8 Pemodelan UML.....	24
2.2.9 Database dan DBMS.....	27
2.2.10 Software dan Pemrograman Terkait.....	28
2.2 Kajian Penelitian Terdahulu.....	33
2.2.1 Paper 1.....	33
2.2.2 Paper 2.....	35
2.2.3 Paper 3.....	37
2.2.4 Paper 4.....	39
2.2.5 Paper 5.....	41
BAB III METODOLOGI PENELITIAN.....	44
3.1 Rancangan Dasar Penelitian.....	44
3.1.1 Bidang dan Jenis Penelitian.....	46

3.1.2 Lokasi Penelitian .....	47
3.1.3 Jadwal Tahapan Penelitian .....	48
3.2 Rancangan Metodologi Penelitian .....	49
3.2.1 Perancangan UML .....	50
3.2.3 Perancangan Antarmuka Aplikasi .....	64
3.2.4 Perancangan Flow Chart Algoritma Prediksi Biaya Ekspor .....	68
3.2.5 Perancangan Class Diagram Sistem Prediksi Biaya Pengiriman .....	70
3.2.6 Evaluasi Akurasi Model .....	71
3.2.7 Analisa Tahap Business understanding .....	73
3.2.8 Analisa Tahap Data understanding .....	79
3.2.9 Rancangan Tahap Data preparation .....	88
3.2.10 Rancangan Tahap Pemodelan .....	91
3.2.11 Rancangan Tahap Testing .....	94
3.2.12 Rancangan Tahap Deploy .....	96
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>	<b>98</b>
4.1 Hasil Penelitian .....	98
4.1.1 Spesifikasi Hardware dan Software Yang Digunakan .....	98
4.1.2 Tampilan Interface Hasil Deploy .....	99
4.1.3 Struktur Database .....	111
4.2 Analisa Hasil .....	113
4.2.1 Percobaan Input-Output Prediksi Biaya Ekspor .....	113
4.2.1 Percobaan Input-Output untuk Prediksi Durasi Keterlambatan .....	115
4.2.2 Testing Hasil .....	117
4.2.3 Hasil Pra-pemrosesan Data .....	122
4.2.4 Modifikasi atau Optimalisasi Dari Sistem Terdahulu .....	128
4.2.5 Proses Deploy Sistem Aplikasi .....	132
4.2.6 Perhitungan Manual Biaya Pengiriman Ekspor .....	135
4.2.7 Analisis Perbandingan antara Nilai Prediksi dan Nilai Aktual .....	138
<b>BAB V PENUTUP .....</b>	<b>142</b>
5.2 Kesimpulan .....	142



## DAFTAR GAMBAR

Gambar 2. 1 Metode CRISP_DM .....	23
Gambar 3. 1 Use Case Diagram (Admin).....	52.
Gambar 3. 2 Use Case Diagram (Staff Logistic) .....	53.
Gambar 3. 3 Use Case Diagram (Manager).....	54.
Gambar 3. 4 Activity Diagram (admin).....	54
Gambar 3. 5 Activity Diagram (Staff).....	55
Gambar 3. 6 Activity Diagram (Manager).....	56
Gambar 3. 7 Rancangan Halaman Login.....	65
Gambar 3. 8 Rancangan Dashboard Admin .....	66
Gambar 3. 9 Rancangan Dashboard Staff Logistic.....	66
Gambar 3. 10 Rancangan Dashboard Manager .....	67
Gambar 3. 11 Flowchart Prediksi Biaya Ekspor.....	69
Gambar 3. 12 histogram dan kde plot.....	82
Gambar 3. 13 Distribusi Negara Tujuan.....	83
Gambar 3. 14 Matriks korelasi fitur numerik dan target .....	84
Gambar 3. 15 Box Plot outlier.....	90
Gambar 3. 16 Tahapan Rancangan Pemodelan .....	91
Gambar 3. 17 Perbandingan visual antara nilai aktual dan prediksi .....	94
Gambar 4. 1 Halaman Login Setiap User.....	102
Gambar 4. 2 Halaman Manajemen Komoditas (Admin).....	103
Gambar 4. 3 Halaman Manajemen Negara Tujuan (Admin).....	104
Gambar 4. 4 Halaman Manajemen Jenis Kapal (Admin) .....	105
Gambar 4. 5 Halaman Manajemen Pengguna (Admin).....	105
Gambar 4. 6 Halaman Penanganan Khusus (Treatment).....	106
Gambar 4. 7 Halaman Dashboard (Staff Logistic) .....	107
Gambar 4. 8 Halaman Prediksi Biaya Pengiriman Ekspor .....	108
Gambar 4. 9 Halaman Riwayat Prediksi Biaya Pengiriman .....	109
Gambar 4. 10 Halaman Simulasi Multi-Skenario .....	110.
Gambar 4. 11 Performa Model .....	111
Gambar 4. 12 Agregat Lanjutan .....	111
Gambar 4. 13 Diagram Relasi antar Tabel .....	112
Gambar 4. 14 Scatter Plot Perbandingan Nilai Aktual vs. Prediksi.....	120
Gambar 4. 15 Grafik Feature Importance.....	123.

## DAFTAR TABEL

Table 3. 1 Struktur Tabel Pengiriman.....	62
Table 3. 2 Struktur Tabel Komoditas .....	63
Table 3. 3 Struktur Tabel Kapal .....	63
Table 3. 4 Struktur Tabel Negara .....	63
Table 3. 5 Struktur Tabel Biaya Tambahan.....	64
Table 3. 6 Deskripsi Data .....	80
Table 3. 7 Output dari df.info () .....	81
Table 3. 8 Output df.head() .....	81
Table 3. 9 missing values .....	85
Table 3. 10 Frekuensi kategori negara_tujuan sebelum koreksi.....	86
Table 3. 11 Frekuensi kategori negara_tujuan setelah koreksi.....	88
Table 4. 1 Analisis Kebutuhan Hardware.....	98
Table 4. 2 Analisis Kebutuhan Software .....	99
Table 4. 3 Detail Uji Coba Fungsional Input-Output Sistem .....	114
Table 4. 4 Ringkasan Rata-rata dan Standar Deviasi Fitur Numerik.....	125

## DAFTAR LAMPIRAN

Lampiran 1 Surat Keterangan Bebas Plagiat .....	148
Lampiran 2 Hasil Turnitin .....	149

