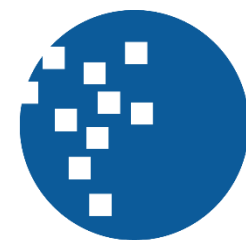


**ANALISIS PERBANDINGAN ALGORITMA *RANDOM*
FOREST DAN *GRADIENT BOOSTING* UNTUK
MEMPREDIKSI LAMA MENGINAP TAMU HOTEL XYZ:
STUDI KASUS PT INDONESIA DUNIA BERKREATIF**



UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

SKRIPSI

Kenny Adam

00000072683

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG**

2025

**ANALISIS PERBANDINGAN ALGORITMA *RANDOM FOREST* DAN *GRADIENT BOOSTING* UNTUK
MEMPREDIKSI LAMA MENGINAP TAMU HOTEL XYZ:
STUDI KASUS PT INDONESIA DUNIA BERKREATIF**



SKRIPSI

Diajukan sebagai Salah Satu Syarat untuk Memperoleh
Gelar Sarjana Komputer

Kenny Adam

00000072683

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG**

2025

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Kenny Adam

Nomor Induk Mahasiswa : 00000072683

Program Studi : Sistem Informasi

Skripsi dengan judul:

**ANALISIS PERBANDINGAN ALGORITMA *RANDOM FOREST* DAN
GRADIENT BOOSTING UNTUK MEMPREDIKSI LAMA MENGINAP
TAMU HOTEL XYZ: STUDI KASUS PT INDONESIA DUNIA
BERKREATIF**

Merupakan hasil karya saya sendiri bukan plagiat dari laporan karya tulis ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan maupun dalam penulisan laporan karya tulis ilmiah, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk mata kuliah yang telah saya tempuh.

Tangerang, 28 November 2025



Kenny Adam

LEMBAR PERNYATAAN PENGGUNAAN BANTUAN KECERDASAN ARTIFISIAL (AI)

Saya yang bertanda tangan di bawah ini:

Nama Lengkap : Kenny Adam
NIM : 00000072683
Program Studi : Sistem Informasi
Judul Skripsi : ANALISIS PERBANDINGAN ALGORITMA
RANDOM FOREST DAN *GRADIENT BOOSTING* UNTUK
MEMPREDIKSI LAMA MENGINAP TAMU HOTEL XYZ: STUDI
KASUS PT INDONESIA DUNIA BERKREATIF

Dengan ini saya menyatakan secara jujur menggunakan bantuan Kecerdasan Artifisial (AI) dalam pengerjaan Tugas/Project/Tugas Akhir*(coret salah satu) sebagai berikut (beri tanda centang yang sesuai):

- ☒ Menggunakan AI sebagaimana diizinkan untuk membantu dalam menghasilkan ide-ide utama saja
- ☒ Menggunakan AI sebagaimana diizinkan untuk membantu menghasilkan teks pertama saja
- ☒ Menggunakan AI untuk menyempurnakan sintaksis dan tata bahasa untuk pengumpulan tugas
- ☐ Karena tidak diizinkan: Tidak menggunakan bantuan AI dengan cara apa pun dalam pembuatan tugas

Saya juga menyatakan bahwa:

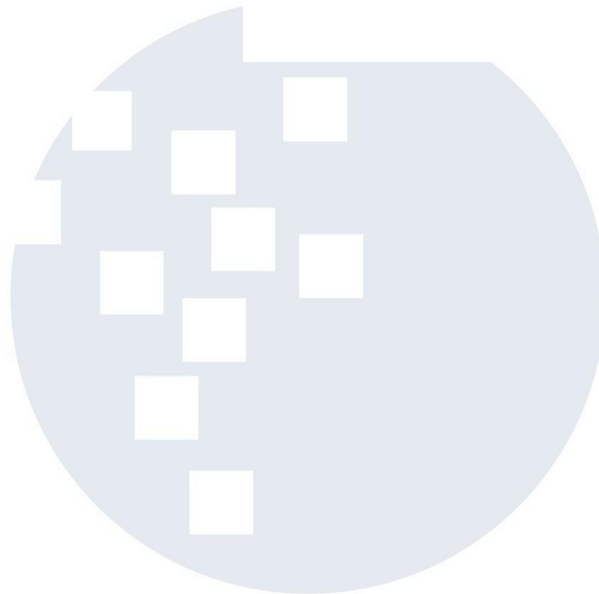
- (1) Menyerahkan secara lengkap dan jujur penggunaan perangkat AI yang diperlukan dalam tugas melalui Formulir Penggunaan Perangkat Kecerdasan Artifisial (AI)
- (2) Saya mengakui bahwa saya telah menggunakan bantuan AI dalam tugas saya baik dalam bentuk kata, paraphrase, penyertaan ide atau fakta penting yang disarankan oleh AI dan saya telah menyantumkan dalam sitasi serta referensi

Terlepas dari pernyataan di atas, tugas ini sepenuhnya merupakan karya saya sendiri

Tangerang, 28 November 2025
Yang Menyatakan,



Kenny Adam



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH

Yang bertanda tangan di bawah ini:

Nama : Kenny Adam

NIM : 00000072683

Program Studi : Sistem Informasi

Jenjang : S1

Judul Karya Ilmiah : ANALISIS PERBANDINGAN
ALGORITMA *RANDOM FOREST* DAN *GRADIENT BOOSTING*
UNTUK MEMPREDIKSI LAMA MENGINAP TAMU HOTEL XYZ:
STUDI KASUS PT INDONESIA DUNIA BERKREATIF

Menyatakan dengan sesungguhnya bahwa saya bersedia* (pilih salah satu):

- ☒ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.
- ☐ Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) **.
- ☐ Lainnya, pilih salah satu:
 - ☐ Hanya dapat diakses secara internal Universitas Multimedia Nusantara
 - ☐ Embargo publikasi karya ilmiah dalam kurun waktu 3 tahun.

Tangerang, 28 November 2025



Kenny Adam

KATA PENGANTAR

Puji Syukur kepada Tuhan Yang Maha Esa, karena berkat rahmat dan kasih karunia-Nya, penulis dapat mampu untuk menyelesaikan atas skripsi dengan judul: “Analisis Perbandingan Algoritma *Random Forest* dan *Gradient Boosting* untuk Memprediksi Lama Menginap Tamu Hotel XYZ: Studi Kasus PT Indonesia Dunia Berkreatif”. Penulisan skripsi ini dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar S1 Sarjana Komputer, Program Studi Sistem Informasi, Fakultas Teknik dan Informatika, Universitas Multimedia Nusantara. Penulis menyadari bahwa tanpa bantuan dan bimbingan dari berbagai pihak sangatlah sulit untuk menyelesaikan tugas akhir ini, oleh karena itu, saya mengucapkan terima kasih kepada:

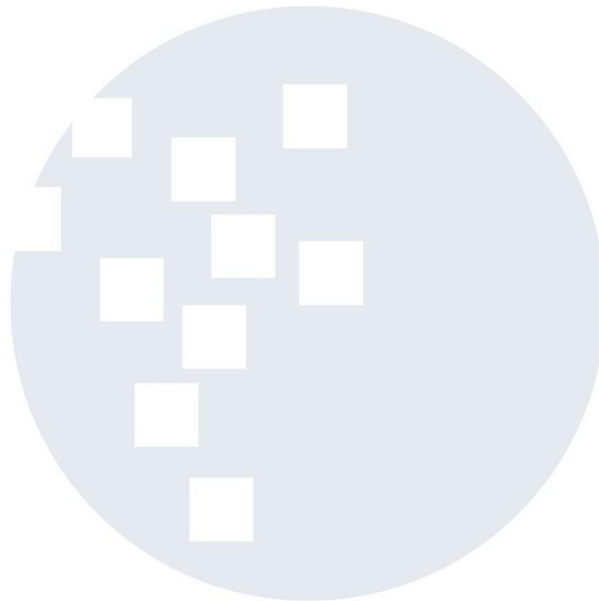
1. Ir. Andrey Andoko, M.Sc., Ph.D., selaku Rektor Universitas Multimedia Nusantara.
2. Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Ibu Ririn Ikana Desanti, S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi Universitas Multimedia Nusantara.
4. Ibu Melissa Indah Fianty, S.Kom., M.MSI, sebagai Dosen Pembimbing yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. PT Indonesia Dunia Berkreatif yang telah memberikan kesempatan untuk berkembang, serta memfasilitasi diskusi yang konstruktif dalam proses analisis, pengambilan data, serta menjadi bagian penting dalam penyusunan tugas akhir ini.
6. Keluarga saya, khususnya orang tua dan saudara, yang senantiasa memberikan dukungan materiil, moral, motivasi, doa, sehingga menjadi sumber semangat bagi penulis dalam menyelesaikan tugas akhir ini.
7. Felita Aileen Prillia, teman-teman saya yang telah memberikan bantuan dukungan moral, sehingga penulis dapat semangat untuk menyelesaikan tugas akhir ini.

Semoga karya ilmiah ini dapat berguna bagi para pembaca serta menjadi acuan penelitian di masa mendatang.

Tangerang, 28 November 2025



Kenny Adam



UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

ANALISIS PERBANDINGAN ALGORITMA *RANDOM*
***FOREST* DAN *GRADIENT BOOSTING* UNTUK**
MEMPREDIKSI LAMA MENGINAP TAMU HOTEL XYZ:
STUDI KASUS PT INDONESIA DUNIA BERKREATIF

Kenny Adam

ABSTRAK

Penelitian ini bertujuan untuk menganalisis dan membandingkan kinerja algoritma *Random Forest* dan *Gradient Boosting* dalam memprediksi lama menginap tamu (*length of stay*) pada Hotel XYZ yang dikelola oleh PT Indonesia Dunia Berkreatif. Penelitian ini dilatarbelakangi oleh adanya penurunan lama menginap tamu pada periode September hingga Desember 2024, sehingga diperlukan pendekatan prediktif berbasis *Machine Learning* untuk mendukung pengambilan keputusan manajemen operasional yang lebih terukur dan berbasis data. Metode penelitian menggunakan pendekatan *Knowledge Discovery in Database* (KDD) yang mencakup tahapan data *selection*, data *preprocessing*, data *transformation*, data *mining*, *evaluation*, dan *business recommendation*. Dataset yang digunakan terdiri dari 2.220 data transaksi tamu selama tahun 2024, dan kinerja model dievaluasi menggunakan metrik regresi berupa RMSE, MAE, MAPE, dan R^2 . Hasil penelitian menunjukkan bahwa model *Gradient Boosting* memiliki kinerja prediksi yang paling baik dan stabil dibandingkan *Random Forest*, dengan nilai R^2 sebesar 0,9921, RMSE sebesar 1,1827, MAE sebesar 0,2904, dan MAPE sebesar 3,50%, yang menunjukkan tingkat kesalahan prediksi yang relatif rendah. Sementara itu, model *Random Forest* menunjukkan kinerja prediksi yang lebih rendah pada data yang sama. Analisis *feature importance* menunjukkan bahwa variabel jumlah tamu merupakan faktor yang paling berpengaruh terhadap lama menginap tamu. Hasil penelitian ini menunjukkan bahwa model *Gradient Boosting* dapat digunakan sebagai dasar pendukung sistem pengambilan keputusan bagi manajemen Hotel XYZ dalam perencanaan operasional berbasis data.

Kata kunci: Analisis Perbandingan, *Gradient Boosting*, *Length of Stay*, *Machine Learning*, *Random Forest*.

**COMPARATIVE ANALYSIS OF RANDOM FOREST AND
GRADIENT BOOSTING ALGORITHMS FOR PREDICTING
HOTEL XYZ GUEST LENGTH OF STAY: A CASE STUDY OF
PT INDONESIA DUNIA BERKREATIF**

Kenny Adam

ABSTRACT (English)

This study aims to analyze and compare the performance of the Random Forest and Gradient Boosting algorithms in predicting guest length of stay at Hotel XYZ, which is managed by PT Indonesia Dunia Berkreatif. The study is motivated by a decline in guest length of stay during the period from September to December 2024, highlighting the need for a machine learning–based predictive approach to support more structured and data-driven operational decision-making. The research methodology follows the Knowledge Discovery in Databases (KDD) framework, which includes data selection, data preprocessing, data transformation, data mining, evaluation, and business recommendation stages. The dataset consists of 2,220 guest transaction records from 2024, and model performance is evaluated using regression metrics, namely RMSE, MAE, MAPE, and R^2 . The results show that the Gradient Boosting model demonstrates the best and most stable predictive performance compared to the Random Forest model, achieving an R^2 value of 0.9921, RMSE of 1.1827, MAE of 0.2904, and MAPE of 3.50%, indicating a relatively low prediction error. In contrast, the Random Forest model exhibits lower predictive performance on the same dataset. Feature importance analysis reveals that the number of guests is the most influential variable affecting length of stay. These findings indicate that the Gradient Boosting model can be used as a decision-support foundation for data-driven operational planning by the management of Hotel XYZ.

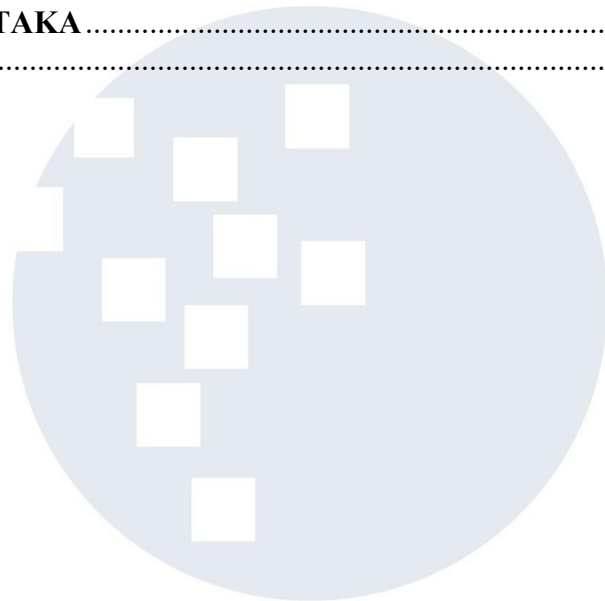
Keywords: *Comparative Analysis, Gradient Boosting, Length of Stay, Machine Learning, Random Forest.*

DAFTAR ISI

HALAMAN PERNYATAAN TIDAK PLAGIAT	ii
KATA PENGANTAR.....	vi
ABSTRAK	viii
<i>ABSTRACT (English)</i>	ix
DAFTAR ISI.....	x
DAFTAR TABEL	xiii
DAFTAR GAMBAR.....	xiv
DAFTAR RUMUS	xvi
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	9
1.3 Batasan Masalah	9
1.4 Tujuan dan Manfaat Penelitian.....	10
1.4.1 Tujuan Penelitian.....	10
1.4.2 Manfaat Penelitian	11
1.5 Sistematika Penulisan	11
BAB II LANDASAN TEORI	13
2.1 Penelitian Terdahulu	13
2.2 Tinjauan Teori.....	18
2.2.1 Machine Learning	18
2.2.2 Feature Engineering.....	19
2.2.3 Evaluasi Regresi	20
2.3 Framework dan Algoritma yang digunakan	20
2.3.1 Knowledge Discovery in Database (KDD).....	20
2.3.2 Algoritma Random Forest	23
2.3.2 Algoritma Gradient Boosting	24
2.3.3 Evaluasi Model Prediksi.....	26
2.4 Tools yang digunakan	29
2.4.1 Python	29
2.4.1 Jupyter Notebook.....	29
2.4.1 MariaDB	30
BAB III METODOLOGI PENELITIAN	31
3.1 Gambaran Umum Objek Penelitian.....	31
3.1.1 Sejarah Perusahaan	31
3.1.2 Struktur Organisasi Perusahaan	32

3.2	Metodologi Penelitian	33
3.3	Alur Penelitian.....	34
3.3.1	Business Understanding	36
3.3.2	Data Selection	36
3.3.3	Data Preprocessing	37
3.3.4	Data Transformation	37
3.3.5	Data Mining.....	38
3.3.6	Interpretation and Evaluation	38
3.3.7	Business Recommendation.....	39
3.4	Teknik Pengumpulan Data	39
3.4.1	Studi Literatur	39
3.4.2	Populasi Penelitian.....	40
3.4.3	Sampel Penelitian.....	40
3.4.4	Metode Pengumpulan Data	41
3.5	Teknik Analisis Data.....	41
BAB IV	ANALISIS DAN HASIL PENELITIAN	43
4.1	Business Understanding	43
4.2	Data Collection	44
4.3	Data Grouping.....	49
4.4	Data Cleaning	50
4.5	Data Transformation	52
4.5.1	Initial Transformation.....	52
4.5.2	Feature Engineering	53
4.5.3	Feature Selection.....	53
4.5.4	Normalization Data.....	54
4.6	Data Visualization	55
4.7	Data Mining.....	59
4.7.1	Train and Split Data	60
4.7.2	Cross Validation.....	61
4.7.3	Hyperparameter Tuning.....	62
4.8	Evaluation	63
4.8.1	Evaluasi Perbandingan.....	63
4.8.2	Feature Importance	66
4.8.3	Evaluasi Performa Model Terbaik.....	71

4.8.4 Temuan Prediksi Performa Model Terbaik.....	81
4.9 Business Recommendation	86
4.10 Hasil dan Pembahasan.....	95
BAB V SIMPULAN DAN SARAN.....	101
5.1 Simpulan.....	101
5.2 Saran.....	102
DAFTAR PUSTAKA.....	105
LAMPIRAN.....	110



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR TABEL

Tabel 1. 1 Permasalahan dan Dampak	3
Tabel 2. 1 Penelitian Terdahulu	13
Tabel 2. 2 Kriteria Evaluasi Metrik Regresi	28
Tabel 4. 1 Atribut Pada Dataset	45
Tabel 4. 2 Hasil Perbandingan Performa Model.....	64
Tabel 4. 4 Perbandingan Nilai Feature Importance	69
Tabel 4. 3 Ringkasan Evaluasi Performa Model.....	75
Tabel 4. 5 Temuan dan Rekomendasi Bisnis	87
Tabel 4. 6 Pembahasan Hasil Model Algoritma	96
Tabel 4. 7 Perbandingan Hasil Penelitian Ini dengan Penelitian Terdahulu.....	99



DAFTAR GAMBAR

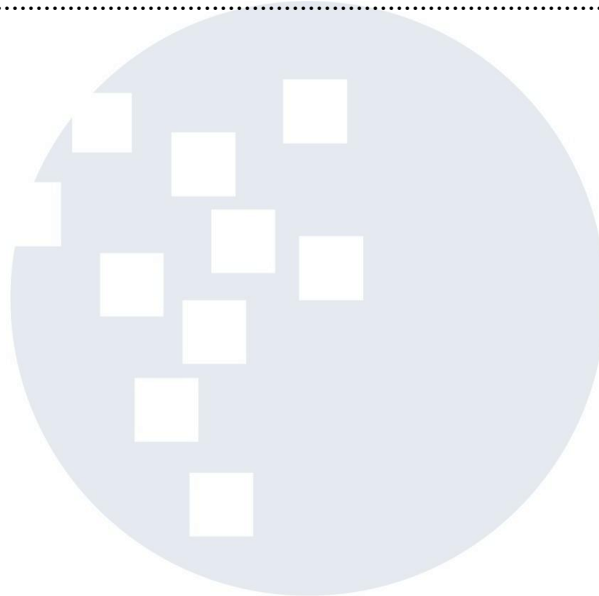
Gambar 1. 1 Jumlah Usaha Akomodasi Tahun 2020-2024	1
Gambar 1. 2 Data Internal Hotel XYZ	2
Gambar 2. 1 Tahapan Proses Knowledge Discovery in Database (KDD).....	21
Gambar 2. 2 Struktur Algoritma Random Forest.....	24
Gambar 2. 3 Struktur Algoritma Gradient Boosting	25
Gambar 2. 4 Proses Training Gradient Boosting	26
Gambar 3. 1 Logo Indonesia Dunia Berkreatif.....	31
Gambar 3. 2 Struktur Organisasi Perusahaan	32
Gambar 3. 3 Alur Penelitian.....	35
Gambar 4. 1 Proses Data Collection	44
Gambar 4. 2 Hasil Pengambilan Data	45
Gambar 4. 3 Script Proses Load Data	46
Gambar 4. 4 Script Dataset Information	47
Gambar 4. 5 Script Proses Filter Data.....	48
Gambar 4. 6 Script Data Grouping	49
Gambar 4. 7 Script Proses Cleaning Data.....	51
Gambar 4. 8 Script Proses Transformation Data	52
Gambar 4. 9 Script Proses Feature Engineering	53
Gambar 4. 10 Script Proses Pemilihan Feature.....	54
Gambar 4. 11 Script Proses Normalisasi Data	54
Gambar 4. 12 Interpretasi Visualisasi Data 1.....	55
Gambar 4. 13 Interpretasi Visualisasi Data 2.....	56
Gambar 4. 14 Interpretasi Visualisasi Data 3.....	57
Gambar 4. 15 Interpretasi Visualisasi Data 4.....	58
Gambar 4. 16 Heatmap Korelasi Data	59
Gambar 4. 17 Script Proses Train dan Split Data	60
Gambar 4. 18 Script Proses Cross Validation.....	61
Gambar 4. 19 Script Proses Hyperparameter Tuning	62
Gambar 4. 20 Script Evaluasi Perbandingan Performa Model	63
Gambar 4. 21 Visualisasi Perbandingan Model Terhadap Aktual dan Prediksi...	65
Gambar 4. 22 Feature Importance Algoritma Random Forest.....	66
Gambar 4. 23 Feature Importance Algoritma Gradient Boosting.....	67
Gambar 4. 24 Perbandingan Feature Importance Model	68
Gambar 4. 25 Script Sinkronisasi Index & Variabel Evaluasi.....	72
Gambar 4. 26 Evaluasi Performa Model Gradient Boosting Terhadap Lama Menginap Tamu – Actual dan Prediksi.....	73
Gambar 4. 27 Evaluasi Performa Model – Residual Distribution	74
Gambar 4. 28 Residual Terhadap Jumlah Tamu	76
Gambar 4. 29 Residual Terhadap Rata-rata per Booking	77
Gambar 4. 30 Residual Terhadap Pendapatan per Booking	78
Gambar 4. 31 Rata-rata Absolute Error Kategori	79

Gambar 4. 32 Korelasi Fitur Terhadap Absolute Error.....	80
Gambar 4. 33 Confusion Matrix Prediksi Lama Menginap Tamu – Durasi Fitur Bulanan	81
Gambar 4. 34 Confusion Matrix Prediksi Lama Menginap Tamu – Durasi Fitur Weekdays	82
Gambar 4. 35 Confusion Matrix Prediksi Lama Menginap Tamu – Durasi Fitur Weekend.....	83
Gambar 4. 36 Prediksi Lama Menginap Terhadap Segmentasi.....	84
Gambar 4. 37 Prediksi Rata-rata Lama Menginap Tamu per Segmen (Gradient Boosting).....	85
Gambar 4. 38 Histogram Error Pada Actual dan Predicted Terhadap Lama Menginap	86
Gambar 4. 39 Diagram Alur Keputusan Bisnis Hotel XYZ	89
Gambar 4. 40 Hasil Akhir Perbandingan Model Algoritma	98



DAFTAR RUMUS

Rumus (2.1).....	24
Rumus (2.2).....	24
Rumus (2.3).....	24
Rumus (2.4).....	27
Rumus (2.5).....	27
Rumus (2.6).....	28



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR LAMPIRAN

Lampiran A Turnitin Similarity Report	110
Lampiran B Form Konsultasi Bimbingan.....	123
Lampiran C Formulir Penggunaan Perangkat Kecerdasan Artifisial (AI).....	125

