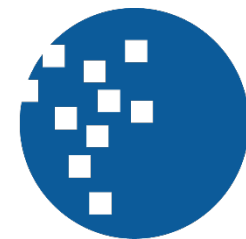


**Optimalisasi Prediksi Pertumbuhan Peserta dan Pendapatan
pada platform Labskill Menggunakan Machine Learning dan
Deep Learning**



UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

SKRIPSI

Marcello Roy

00000072593

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

2025

**Optimalisasi Prediksi Pertumbuhan Peserta dan Pendapatan
pada platform Labskill Menggunakan Machine Learning dan
Deep Learning**



Diajukan sebagai Salah Satu Syarat untuk Memperoleh

Gelar Sarjana Komputer

Marcello Roy

00000072593

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

2025

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Marcello Roy

Nomor Induk Mahasiswa : **00000072593**

Program Studi : Sistem Informasi

Skripsi dengan judul:

Optimalisasi Prediksi Pertumbuhan Peserta dan Pendapatan pada platform
Labskill Menggunakan Machine Learning dan Deep Learning

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, 27 November 2025



Marcello Roy

LEMBAR PERNYATAAN PENGGUNAAN BANTUAN KECERDASAN ARTIFISIAL (AI)

Saya yang bertanda tangan di bawah ini:

Nama Lengkap : Marcello Roy
NIM : 00000072593
Program Studi : Sistem Informasi
Judul Skripsi : Optimalisasi Prediksi Pertumbuhan Peserta dan
Pendapatan pada platform Labskill Menggunakan Machine Learning dan
Deep Learning

Dengan ini saya menyatakan secara jujur menggunakan bantuan Kecerdasan Artifisial (AI) dalam pengerjaan Tugas Akhir sebagai berikut:

- ☒ 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
- (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, 27 November 2025
Yang Menyatakan



Marcello Roy

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH

Yang bertanda tangan di bawah ini:

Nama : Marcello Roy

NIM : 00000072593

Program Studi : Sistem Informasi

Jenjang : S1

Judul Karya Ilmiah : Optimalisasi Prediksi Pertumbuhan Peserta dan Pendapatan pada platform Labskill Menggunakan Machine Learning dan Deep Learning

Menyatakan dengan sesungguhnya bahwa saya bersedia* (pilih s:

- ☐ 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, 27 November 2025



Marcello Roy

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas rahmat dan karunia-Nya sehingga skripsi berjudul “Optimasi Prediksi Pendapatan dan Pertumbuhan Peserta Platform Labskill dengan Pendekatan Machine Learning dan Deep Learning” dapat diselesaikan. Skripsi ini disusun sebagai salah satu syarat memperoleh gelar Sarjana Sistem Informasi di Universitas Multimedia Nusantara. Penulis mengucapkan terima kasih kepada dosen pembimbing, keluarga, dan rekan-rekan atas dukungan dan bimbingannya. Semoga karya ini bermanfaat bagi pengembangan ilmu dan praktik di bidang sistem informasi.

1. Dr. Ir. Andrey Andoko, M.Sc., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Universitas Multimedia Nusantara
3. Ibu Ririn Ikana Desanti, S.Kom., M.Kom., selaku Ketua Program Studi Universitas Multimedia Nusantara.
4. Ibu Wella, S.Kom., M.MSI., sebagai Pembimbing yang telah banyak meluangkan waktu untuk memberikan bimbingan dan arahan sehingga terselesainya tesis ini.
5. Bapak Arjun Roy, B.Sc., sebagai Pembimbing Lapangan yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya laporan MBKM ini.
6. Keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tugas akhir ini.

Semoga karya ilmiah ini bermanfaat bagi pembaca mengenai kegiatan magang

Tangerang, 27 November 2025



Marcello Roy

Optimalisasi Prediksi Pertumbuhan Peserta dan Pendapatan pada platform Labskill Menggunakan Machine Learning dan Deep Learning

Marcello Roy

ABSTRAK

Pertumbuhan platform edukasi digital seperti LabSkill menuntut ketersediaan sistem prediksi yang akurat untuk memperkirakan jumlah peserta dan total revenue sebagai dasar pengambilan keputusan strategis. Penelitian ini bertujuan mengembangkan sistem forecasting berbasis machine learning dengan membandingkan algoritma Extreme Gradient Boosting dan Artificial Neural Network pada dua target, yaitu jumlah peserta dan total revenue bulanan. Tahapan penelitian meliputi data preprocessing, rekayasa fitur, pemodelan, serta optimasi hyperparameter menggunakan Optuna untuk Extreme Gradient Boosting dan Hyperband untuk Artificial Neural Network. Hasil evaluasi menunjukkan bahwa Extreme Gradient Boosting dengan tuning Optuna memberikan performa terbaik pada kedua target, dengan nilai R^2 sebesar 0.9543 untuk prediksi jumlah peserta dan 0.8972 untuk prediksi total revenue, disertai nilai MAE, MAPE, dan RMSE yang rendah dibandingkan model baseline. Sebaliknya, Artificial Neural Network tidak menunjukkan performa yang memadai akibat keterbatasan jumlah data dan kompleksitas pola deret waktu, sehingga menghasilkan error yang tinggi dan nilai R^2 negatif. Model terbaik kemudian diimplementasikan ke dalam sistem berbasis web menggunakan Streamlit untuk mendukung proses forecasting secara interaktif. Secara keseluruhan, penelitian ini menunjukkan bahwa pendekatan berbasis tree seperti Extreme Gradient Boosting lebih efektif dan stabil untuk forecasting data bisnis dengan jumlah sampel terbatas dan pola temporal yang kompleks.

Kata kunci: Deep Learning, Machine Learning, Prediksi Pendapatan, Prediksi Pertumbuhan

Optimizing Participant and Revenue Growth Prediction on the LabSkill Platform Using Machine Learning and Deep Learning

Marcello Roy

ABSTRACT (English)

The growth of digital education platforms such as LabSkill requires accurate prediction systems to estimate the number of participants and total revenue as a basis for strategic decision-making. This study aims to develop a machine learning-based forecasting system by comparing Extreme Gradient Boosting and Artificial Neural Network algorithms on two targets, namely the number of participants and monthly total revenue. The research stages include data preprocessing, feature engineering, model development, and hyperparameter optimization using Optuna for Extreme Gradient Boosting and Hyperband for Artificial Neural Network. The evaluation results show that Extreme Gradient Boosting with Optuna tuning achieves the best performance for both targets, with R^2 values of 0.9543 for participant prediction and 0.8972 for revenue prediction, accompanied by lower MAE, MAPE, and RMSE compared to baseline models. In contrast, the Artificial Neural Network fails to deliver satisfactory performance due to limited data size and the complexity of temporal patterns, resulting in high prediction errors and large negative R^2 values. The best-performing model is then deployed through a web-based system using Streamlit to enable interactive real-time forecasting. Overall, this study demonstrates that tree-based approaches such as Extreme Gradient Boosting are more effective and stable for business forecasting tasks involving limited datasets and complex temporal patterns.

Keywords: Deep Learning, Growth Prediction, Machine Learning, Revenue Prediction

DAFTAR ISI

HALAMAN PERNYATAAN TIDAK PLAGIAT	ii
LEMBAR PERNYATAAN PENGGUNAAN BANTUAN KECERDASAN ARTIFISIAL	iii
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	iv
KATA PENGANTAR.....	v
ABSTRAK	vi
<i>ABSTRACT (English)</i>	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	5
1.3 Batasan Masalah	5
1.4 Tujuan dan Manfaat Penelitian	5
1.4.1 Tujuan Penelitian.....	5
1.4.2 Manfaat Penelitian	6
1.5 Sistematika Penulisan.....	6
BAB II LANDASAN TEORI	8
2.1 Penelitian Terdahulu	8
2.2 Teori yang berkaitan / Konsep Dasar	12
2.2.1 Teori Prediksi	12
2.2.2 Deep Learning	13
2.2.3 Machine Learning	13
2.2.4 Hyperparameter Tuning	14
2.3 Framework/Algoritma yang digunakan	14
2.3.1 CRISP-DM.....	14
2.3.2 Extreme Gradient Boosting.....	16
2.3.3 Artificial Neural Network.....	17

2.3.4	Tuning	19
2.4	Tools/software yang digunakan	21
2.4.1	Jupyter Notebook	21
2.4.2	Python	21
2.4.3	Github	22
2.4.4	Streamlit	23
2.5	Metrik Evaluasi	23
2.5.1	Mean Absolute Error	23
2.5.2	Mean Absolute Percentage Error	24
2.5.3	Root Mean Square Error	24
2.5.4	R^2	24
BAB III	METODOLOGI PENELITIAN	26
3.1	Gambaran Umum Objek Penelitian	26
3.2	Metode Penelitian	27
3.2.1	Alur Penelitian	27
3.2.2	Kerangka Kerja Data Mining dan Data Analysis	30
3.3	Teknik Pengumpulan Data	32
3.3.1	Populasi dan Sampel	32
3.3.2	Periode Pengambilan Data	33
3.4	Variabel Penelitian	33
3.4.1	Variabel Independen	33
3.4.2	Variabel dependen	34
3.5	Teknik Analisis Data	34
BAB IV	ANALISIS DAN HASIL PENELITIAN	37
4.1	Business Understanding	37
4.2	Data Understanding	38
4.3	Data Preparation	42
4.4	Modeling	57
4.4.1	Model Dasar - Target 1 Peserta	57
4.4.2	Model Dasar - Target 2 Revenue	60
4.4.3	Tuning Model Optuna dan Hyperband	64
4.5	Evaluation	80

4.5.1	Model Dasar Evaluasi Model Target 1 Peserta.....	81
4.5.2	Evaluasi Model Target 2 Revenue.....	81
4.5.3	Evaluasi Performa.....	81
4.5.4	Analisis Performa.....	82
4.5.5	Komparasi Penelitian Terdahulu dengan Model Terbaik	87
4.6	Deployment	89
BAB V SIMPULAN DAN SARAN		94
5.1	Simpulan.....	94
5.2	Saran	95
DAFTAR PUSTAKA		97
LAMPIRAN.....		102
Lampiran A Turnitin		102
Lampiran B Form Konsultasi Bimbingan		103
Lampiran C Formulir Penggunaan Perangkat Kecerdasan Artifisial (AI)		104



DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu	8
Tabel 3. 2 Kerangka Kerja Data Mining.....	30
Tabel 3. 3 Periode Data.....	33
Tabel 3. 4 Perbandingan Bahasa Pemrograman.....	35
Tabel 4. 1 Dataset Labskill.....	38
Tabel 4. 2 Extreme Gradient Boost Base - Target 1 Peserta.....	58
Tabel 4. 3 Artificial Neural Network Base - Target 1 Peserta	60
Tabel 4. 4 Revenue.....	62
Tabel 4. 5 Artificial Neural Network Base - Target 2 Revenue.....	64
Tabel 4. 6 Extreme Gradient Boost Optuna - Target 1 Peserta.....	67
Tabel 4. 7 Extreme Gradient Boost Optuna - Target 2 Peserta Revenue.....	73
Tabel 4. 8 Artificial Neural Network Hyperband - Target 1 Peserta	76
Tabel 4. 9 Artificial Neural Network Hyperband - Target 2 Revenue.....	80
Tabel 4. 10 Evaluasi Metric Target 1	81
Tabel 4. 11 Evaluasi Metric Target 2.....	81



DAFTAR GAMBAR

Gambar 2. 1 Metodologi CRISP DM - Project.....	15
Gambar 2. 2 Arsitektur Extreme Gradient Boost.....	16
Gambar 2. 3 Arsitektur Artificial Neural Network	18
Gambar 3. 1 Alur Penelitian.....	27
Gambar 3. 2 Tahap 1: Prediksi Model - Machine Learning	29
Gambar 3. 3 Tahap 2: Prediksi Model - Deep Learning	29
Gambar 4. 1 Data Understanding.....	40
Gambar 4. 2 Analisis Gender Customer	40
Gambar 4. 3 Distribusi Metode Pembayaran	41
Gambar 4. 4 Mengecek Missing Value.....	41
Gambar 4. 5 Validasi Diskon dan Rating Diberikan.....	42
Gambar 4. 6 Mengisi Missing Value	43
Gambar 4. 7 Mengoversi tipe data time	43
Gambar 4. 8 Memastikan Diskon dan Harga Program	43
Gambar 4. 9 Membuat Kolom Nominal _Dibayar dengan rumus	44
Gambar 4. 10 Melihat Hasil rumus Nominal Dibayar	44
Gambar 4. 11 Mengekstrak Tahun dan Bulan	45
Gambar 4. 12 Membuat Agregasi Bulan.....	46
Gambar 4. 13 Membuat Fitur Lag.....	46
Gambar 4. 14 Membuat Rolling Statistics	47
Gambar 4. 15 Membuat Fitur Growth dan Differences	48
Gambar 4. 16 Membuat fitur interaction.....	48
Gambar 4. 17 Membuat time index.....	49
Gambar 4. 18 Membersihkan Nan	50
Gambar 4. 19 Menangani Outlier.....	50
Gambar 4. 20 Melihat Outlier Jumlah Peserta dan Total Revenue.....	51
Gambar 4. 21 Korelasi Jumlah Peserta dengan Total Revenue	52
Gambar 4. 22 Korelasi Jumlah Peserta Bulan Sebelumnya dengan Bulan Berjalan	53
Gambar 4. 23 Pola Perubahan Jumlah Peserta Bulanan	54
Gambar 4. 24 Pola Perubahan Jumlah Revenue	55
Gambar 4. 25 Numerical Korelasi	56
Gambar 4. 26 Data Splitting.....	57
Gambar 4. 27 Model Extreme Gradient Boost Target 1	58
Gambar 4. 28 Model Artificial Neural Network Target 1	59
Gambar 4. 29 Model Extreme Gradient Boost Target 2	61
Gambar 4. 30 Model Artificial Neural Network Target 2	63
Gambar 4. 31 Mencari Tuning terbaik optuna Target 1.....	65

Gambar 4. 32 Hasil Model Extreme Gradient Boost Optuna Target 1	67
Gambar 4. 33 Mencari Tuning Terbaik optuna Target 2	69
Gambar 4. 34 Hasil Model Extreme Gradient Boost Optuna Target 2	72
Gambar 4. 35 Mencari Hasil Tuning Hyperband Target 1 Artificial Neural Network.....	74
Gambar 4. 36 Hasil Konfigurasi Model Artificial Neural Network Target 1	75
Gambar 4. 37 Hasil Model Artificial Neural Network Target 1	76
Gambar 4. 38 Pencarian Tuning Hyperband Artificial Neural Network Terbaik.	77
Gambar 4. 39 Hasil Konfigurasi Hyperband Target 2	79
Gambar 4. 40 Hasil Model Artificial Neural Network Target 2	79
Gambar 4. 41 Perbandingan Prediksi vs Aktual Target: Jumlah Peserta.....	84
Gambar 4. 42 Perbandingan Prediksi vs Aktual Target: Jumlah Revenue	84
Gambar 4. 43 Residual Analysis - Extreme Gradient Boosting + Optuna Model	85
Gambar 4. 44 Perbandingan R2 score untuk Semua Model	86
Gambar 4. 45 Perbandingan MAPE score untuk Semua Model.....	87
Gambar 4. 46 Menyimpan Hasil Model.....	90
Gambar 4. 47 Struktural Github untuk menyambung Streamlit	91
Gambar 4. 48 Homepage.....	91
Gambar 4. 49 Revenue Page	92
Gambar 4. 50 Forecast Peserta Page	93



DAFTAR LAMPIRAN

Lampiran A Form Konsultasi Bimbingan.....	103
Lampiran B Pernyataan Penggunaan AI.....	104
Lampiran C Pernyataan penggunaan AI	105



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA