

SKRIPSI

**PERBANDINGAN ALGORITMA *MACHINE LEARNING*
UNTUK MEMPREDIKSI TINGKAT PROMOSI KARYAWAN
PADA PT PELABUHAN TANJUNG PRIOK**



ANDI IRFAN DAENG MAPPA

1519620004

**PROGRAM STUDI
SISTEM DAN TEKNOLOGI INFORMASI
FAKULTAS TEKNIK
UNIVERSITAS NEGERI JAKARTA**


2025

LEMBAR PENGESAHAN SKRIPSI

Judul : Perbandingan Algoritma *Machine Learning* Untuk Memprediksi
Tingkat Promosi Karyawan Pada PT Pelabuhan Tanjung Priok
Penyusun : Andi Irfan Daeng Mappa
NIM : 1519620004

Skripsi telah didiskusikan dan diusulkan dari Dosen Pembimbing:

NAMA DOSEN	TANDA TANGAN	TANGGAL
Irma Permata Sari, M.Eng. (Dosen Pembimbing 1)		10 Desember 2024

Fuad Mumtas, M.T.I. (Dosen Pembimbing 2)		10 Desember 2024
---	--	------------------

Telah disetujui oleh:

NAMA DOSEN	TANDA TANGAN	TANGGAL
Lipur Sugiyanta, Ph.D (Dosen Penguji 1)		05 Februari 2025

Murien Nugraheni, S.T, M.Cs (Dosen Penguji 2)		05 Februari 2025
--	--	------------------

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa :

1. Skripsi ini merupakan karya asli dan belum pernah diajukan untuk mendapatkan gelar akademik sarjana, baik di Universitas Negeri Jakarta maupun di Perguruan Tinggi lain.
2. Skripsi ini belum dipublikasikan, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan dicantumkan dalam daftar pustaka.
3. Pernyataan ini saya buat dengan sesungguhnya dan apabila di kemudian hari terdapat penyimpangan, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Universitas Negeri Jakarta.

Jakarta, 05 Februari 2025

Yang membuat pernyataan



Andi Irfan Daeng Mappa

NIM. 1519620004



KEMENTERIAN PENDIDIKAN TINGGI, SAINS DAN TEKNOLOGI
UNIVERSITAS NEGERI JAKARTA
UPT PERPUSTAKAAN

Jalan Rawamangun Muka Jakarta 13220
Telepon/Faksimili: 021-4894221
Laman: lib.unj.ac.id

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademika Universitas Negeri Jakarta, yang bertanda tangan di bawah ini, saya:

Nama : Andi Irfan Daeng Mappa
NIM : 1519620004
Fakultas/Prodi : Fakultas Teknik/Sistem dan Teknologi Informasi
Alamat email : andiirfan23423@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada UPT Perpustakaan Universitas Negeri Jakarta, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah:

Skripsi Tesis Disertasi Lain-lain (.....)

yang berjudul :

Perbandingan Algoritma Machine Learning Untuk Memprediksi Tingkat Promosi Karyawan

Pada PT Pelabuhan Tanjung Priok

Dengan Hak Bebas Royalti Non-Eksklusif ini UPT Perpustakaan Universitas Negeri Jakarta berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya, dan menampilkan/mempublikasikannya di internet atau media lain secara *fulltext* untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan Universitas Negeri Jakarta, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 10 Februari 2025

Penulis

(Andi Irfan Daeng Mappa)
NIM. 1519620004

KATA PENGANTAR

Puji dan syukur penulis ucapkan ke hadirat Allah SWT yang telah melimpahkan rahmat serta karunia-Nya sehingga penulis dapat menyelesaikan Skripsi yang berjudul “Perbandingan Algoritma *Machine Learning* Untuk Memprediksi Tingkat Promosi Karyawan Pada PT Pelabuhan Tanjung Priok”. Skripsi ini disusun sebagai salah satu persyaratan meraih gelar Sarjana Komputer pada Program Studi S1 Sistem dan Teknologi Informasi Universitas Negeri Jakarta. Penyelesaian skripsi ini tidak lepas dari bantuan berbagai pihak, oleh karena itu penulis menyampaikan ucapan terima kasih kepada :

1. Lipur Sugiyanta, Ph.D, selaku Koordinator Program Studi Sistem dan Teknologi Informasi, Fakultas Teknik, Universitas Negeri Jakarta;
2. Ibu Irma Permata Sari selaku Dosen Pembimbing I;
3. Bapak Fuad Mumtas selaku Dosen Pembimbing II;
4. Seluruh Dosen Sistem dan Teknologi Informasi Universitas Negeri Jakarta yang telah memberi bekal pengetahuan yang berharga;
5. Kepada orang tua saya yang selalu terus memberikan support serta doa;
6. Teman-teman seperjuangan STI yang setia saling membantu dan memberikan semangat dalam proses penyusunan Skripsi;
7. Terakhir saya ingin mengucapkan terima kasih kepada diri saya sendiri yang telah bersusah payah melawan rasa malas yang sangat merusakkan.

Penulis hanya dapat memanjatkan doa semoga semua pihak yang telah membantu penulis dalam penyusunan Skripsi ini mendapatkan pahala dari Allah SWT. Semoga Skripsi ini dapat bermanfaat dan memberikan sumbangan yang berarti bagi pihak yang membutuhkan.

Jakarta, 12 Desember 2024



Penulis

PERBANDINGAN ALGORITMA *MACHINE LEARNING* UNTUK MEMPREDIKSI TINGKAT PROMOSI KARYAWAN PADA PT PELABUHAN TANJUNG PRIOK

Andi Irfan Daeng Mappa

Dosen Pembimbing : Irma Permata Sari, M. Eng. & Fuad Mumtas, M.T.I.

ABSTRAK

Penelitian ini membandingkan empat algoritma Machine Learning untuk memprediksi tingkat promosi karyawan di PT Pelabuhan Tanjung Priok menggunakan metode *Knowledge Discovery in Database* (KDD). Dataset yang digunakan mencakup 15 variabel dengan total 1,107 data. Penelitian dilakukan melalui tahapan pengumpulan data, *pre-processing*, transformasi data, pemodelan data, dan evaluasi. Algoritma yang dibandingkan meliputi Random Forest, K-Nearest Neighbors, Decision Tree, dan XGBoost. Evaluasi model menggunakan *Confusion Matrix* dengan metrik *accuracy*, *precision*, *recall*, dan *f1-score*.

Hasil penelitian menunjukkan bahwa algoritma XGBoost memiliki performa terbaik setelah dilakukan optimasi parameter menggunakan *Hyperparameter Tuning*. Penelitian ini memberikan kontribusi praktis bagi PT Pelabuhan Tanjung Priok untuk meningkatkan efektivitas proses promosi jabatan serta menawarkan wawasan bagi perusahaan lain dalam mengembangkan sistem prediksi promosi yang efisien dan akurat.

Kata Kunci: Promosi Karyawan, Algoritma, *Machine Learning*.

**COMPARISON OF MACHINE LEARNING ALGORITHMS TO
PREDICT EMPLOYEE PROMOTION RATES AT PT
PELABUHAN TANJUNG PRIOK**

Andi Irfan Daeng Mappa

Advisor : Irma Permata Sari, M. Eng. & Fuad Mumtas, M.T.I.

ABSTRACT

This study compares four Machine Learning algorithms to predict employee promotion rates at PT Pelabuhan Tanjung Priok using the Knowledge Discovery in Database (KDD) method. The dataset consists of 15 variables and 1,107 records. The research process involves data collection, preprocessing, data transformation, data mining, and evaluation. The algorithms compared include Random Forest, K-Nearest Neighbors, Decision Tree, and XGBoost. Model evaluation uses a Confusion Matrix with metrics such as accuracy, precision, recall, and F1-score.

The results indicate that XGBoost algorithm outperforms other algorithms, achieving the highest accuracy after parameter optimization using Hyperparameter Tuning. This research provides practical insights for PT Pelabuhan Tanjung Priok to enhance the effectiveness of its promotion process and offers valuable guidance for other companies seeking to develop efficient and accurate promotion prediction systems.

Keyword : *Employee Promotion, Algorithm, Machine Learning.*

DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI	ii
LEMBAR PERNYATAAN	iii
KATA PENGANTAR.....	iv
ABSTRAK	v
DAFTAR ISI.....	vii
DAFTAR TABEL	x
DAFTAR GAMBAR.....	xi
LAMPIRAN.....	xv
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah.....	1
1.2 Identifikasi Masalah	3
1.3 Pembatasan Masalah	3
1.4 Perumusan Masalah	4
1.5 Tujuan Penelitian	4
1.6 Manfaat Penelitian	4
BAB 2 TINJAUAN PUSTAKA.....	5
2.1 Kerangka Teoritik	5
2.1.1 Sumber Daya Manusia	5
2.1.2 PT. Pelabuhan Tanjung Priok	5
2.1.3 <i>Machine Learning</i>	6
2.1.4 <i>Supervised Learning</i>	7
2.1.5 SMOTE (<i>Synthetic Minority Over-sampling Technique</i>)	7
2.1.6 Algoritma Random Forest.....	8
2.1.7 Algoritma K-Nearest Neighbours	9

2.1.8	Algoritma Decision Tree.....	9
2.1.9	Algoritma XGBoost	11
2.1.10	<i>Confusion Matrix</i>	12
2.1.11	<i>Hyperparameter Tuning</i>	14
2.1.12	<i>Knowledge Discovery in Database</i>	19
2.2	Penelitian Relevan.....	21
2.3	<i>State of Art</i>	23
2.4	Kerangka Berpikir.....	25
BAB 3	METODE PENELITIAN.....	27
3.1	Alat dan Bahan Akhir Tugas.....	27
3.1.1	Alat.....	27
3.1.2	Bahan	27
3.2	Metodologi Penelitian	27
3.3	Alur Penelitian	28
3.3.1	<i>Data Selection</i>	28
3.3.2	<i>Data Preprocessing</i>	29
3.3.3	<i>Data Transformation</i>	31
3.3.4	<i>Data Mining</i>	32
3.3.5	Penerapan Teknik SMOTE.....	39
3.3.6	<i>Hyperparameter Tuning</i>	39
3.3.7	<i>Evaluation</i>	40
BAB 4	HASIL DAN PEMBAHASAN	41
4.1	Hasil Penelitian	41
4.1.1	<i>Data Selection</i>	41
4.1.2	<i>Data Preprocessing</i>	44
4.1.3	<i>Data Transformation</i>	54

4.1.4 <i>Data Mining</i>	60
4.1.5 <i>Evaluation</i>	66
4.1.6 Penerapan Teknik SMOTE	73
4.1.7 <i>Hyperparameter Tuning</i>	77
4.2 Pembahasan.....	88
4.2.1 Perbandingan Kinerja Algoritma <i>Machine Learning</i>	88
4.2.2 Wawasan dan Rekomendasi Praktis untuk Perusahaan	89
BAB 5 KESIMPULAN DAN SARAN	91
5.1 Kesimpulan	91
5.1.1 Hasil Evaluasi Metriks	91
5.1.2 <i>Hyperparameter Tuning</i>	92
5.1.3 Penerapan Teknik SMOTE	92
5.2 Saran.....	92
5.2.1 Eksplorasi Algoritma Lain	92
5.2.2 Optimasi Parameter	92
5.2.3 Penanganan Dataset	92
DAFTAR PUSTAKA	93
LAMPIRAN	99

DAFTAR TABEL

Tabel 2. 1 Tabel <i>Confusion Matrix</i>	12
Tabel 2. 2 Parameter Algoritma Random Forest	14
Tabel 2. 3 Parameter Algoritma KNN	15
Tabel 2. 4 Parameter Algoritma Decision Tree	17
Tabel 2. 5 Parameter Algoritma XGBoost.....	18
Tabel 2. 6 Tabel <i>State of Art</i>	24
Tabel 3. 1 Data Sampel	34
Tabel 4. 1 Kombinasi Parameter Random Forest	78
Tabel 4. 2 Kombinasi Parameter KNN	81
Tabel 4. 3 Kombinasi Parameter Decision Tree	83
Tabel 4. 4 Kombinasi Parameter XGBoost.....	86
Tabel 5. 1 Hasil Evaluasi Metriks	91



DAFTAR GAMBAR

Gambar 2. 1 Tahapan KDD	20
Gambar 2. 2 Kerangka Berpikir	26
Gambar 3. 1 Metode KDD	27
Gambar 3. 2 Alur Penelitian.....	28
Gambar 3. 3 Syntax <i>Import</i> Data	28
Gambar 3. 4 Alur Tahapan Preprocessing Data.....	30
Gambar 3. 6 Syntax Untuk Pengecekan Nilai Duplikat.....	30
Gambar 3. 5 Syntax Untuk Pengecekan Nilai Hilang.....	30
Gambar 3. 8 Syntax Untuk Pengecekan Data Outlier.....	31
Gambar 3. 7 Syntax Untuk Pengecekan Data Skewed	31
Gambar 3. 9 Syntax Proses Logaritmik	32
Gambar 3. 10 Syntax Proses Normalisasi.....	32
Gambar 3. 11 Syntax Pembagian Rasio Data	33
Gambar 3. 12 Syntax Algoritma Random Forest.....	35
Gambar 3. 13 Syntax Algoritma KNN.....	35
Gambar 3. 14 Syntax Algoritma Decision Tree.....	37
Gambar 3. 15 Syntax Algoritma XGBoost	39
Gambar 4. 1 Memuat Dataset.....	41
Gambar 4. 2 Pengenalan Dataset	42
Gambar 4. 3 Penyesuaian Dataset.....	43
Gambar 4. 4 Tampilan Isi Dataset	43
Gambar 4. 5 Distribusi Dataset Setelah Pembatasan Masalah.....	44
Gambar 4. 6 Pemisahan Data Kategorikal dan Numerik	44
Gambar 4. 7 Identifikasi Data Numerik.....	45
Gambar 4. 8 Identifikasi Data Kategorikal	45
Gambar 4. 9 Pemeriksaan Nilai Hilang.....	45
Gambar 4. 10 Pemeriksaan Nilai Duplikat	46
Gambar 4. 11 Pengecekan Data Skewed.....	46
Gambar 4. 12 Nilai dan Grafik Skewed Data Numerikal	47
Gambar 4. 13 Pengecekan Data <i>Outlier</i>	47

Gambar 4. 14 Boxplot <i>Outlier</i>	47
Gambar 4. 15 Syntax Analisis Probabilitas Promosi Berdasarkan <i>awards_won</i> ..	48
Gambar 4. 16 Probabilitas Promosi Berdasarkan <i>awards_won</i>	49
Gambar 4. 17 Syntax Probabilitas Promosi Berdasarkan <i>KPI_met > 80%</i>	49
Gambar 4. 18 Syntax Probabilitas Promosi Berdasarkan <i>avg_training_score</i>	50
Gambar 4. 19 Syntax Probabilitas Promosi Berdasarkan <i>length_of_service</i>	51
Gambar 4. 20 Syntax Probabilitas Promosi Berdasarkan <i>previous_year_rating</i> ..	52
Gambar 4. 21 Syntax Probabilitas Promosi Berdasarkan <i>no_of_trainings</i>	53
Gambar 4. 22 Objek Data Clean	54
Gambar 4. 23 Probabilitas Promosi Berdasarkan Region	55
Gambar 4. 24 Klasifikasi Pontensial Region Setiap Karyawan	55
Gambar 4. 25 Probabilitas Promosi Berdasarkan <i>Potential Region</i>	55
Gambar 4. 26 Probabilitas Berdasarkan <i>previous_year_rating</i> , <i>KPIs_met > 80%</i> , dan <i>awards_won</i>	56
Gambar 4. 27 Fitur <i>performance_level</i>	57
Gambar 4. 28 <i>High_Avg_Tscore</i>	58
Gambar 4. 29 Fitur <i>Male</i>	58
Gambar 4. 30 Syntax Untuk Encoding Fitur <i>Department</i>	58
Gambar 4. 31 Syntax Untuk Encoding Fitur <i>Education</i>	59
Gambar 4. 32 Syntax Untuk Encoding Fitur <i>Performance_level</i>	59
Gambar 4. 33 Fitur Setelah Proses <i>one-hot encoding</i>	59
Gambar 4. 34 Tranformasi Logaritmik	60
Gambar 4. 35 Proses <i>Normalization</i>	60
Gambar 4. 36 Proses Pembagian Data	61
Gambar 4. 37 Ukuran Data Latih dan Data Uji Setelah Proses <i>Splitting</i>	61
Gambar 4. 38 Pemodelan Random Forest	61
Gambar 4. 39 Menyimpan Model dan Uji Akurasi Random Forest	62
Gambar 4. 40 Pemodelan KNN	63
Gambar 4. 41 Menyimpan Model dan Uji Akurasi	63
Gambar 4. 42 Pemodelan Decision Tree	64
Gambar 4. 43 Menyimpan Model dan Uji Akurasi	64
Gambar 4. 44 Pemodelan XGBoost	65

Gambar 4. 45 Menyimpan Model dan Uji Akurasi.....	65
Gambar 4. 46 Syntax <i>Classification Report</i> Random Forest	66
Gambar 4. 47 <i>Classification Report</i> Random Forest	67
Gambar 4. 48 Syntax Membuat <i>Confusion Matrix</i>	67
Gambar 4. 49 <i>Confusion Matrix</i> Random Forest	68
Gambar 4. 50 Syntax <i>Classification Report</i> KNN	68
Gambar 4. 51 <i>Classification Report</i> KNN	68
Gambar 4. 52 <i>Confusion Matrix</i> KNN.....	69
Gambar 4. 53 Syntax <i>Classification Report</i> Decision Tree	70
Gambar 4. 54 <i>Classification Report</i> Decision Tree	70
Gambar 4. 55 <i>Confusion Matrix</i> Decision Tree	71
Gambar 4. 56 Syntax <i>Classification Report</i> XGBoost	71
Gambar 4. 57 <i>Classification Report</i> XGBoost.....	72
Gambar 4. 58 <i>Confusion Matrix</i> XGBoost	72
Gambar 4. 59 Penerapan Teknik SMOTE	73
Gambar 4. 60 Data Latih dan Data Uji Setelah Penerapan Teknik SMOTE	73
Gambar 4. 61 Evaluasi Ulang Algoritma Random Forest	74
Gambar 4. 62 Evaluasi Ulang Algoritma KNN	75
Gambar 4. 63 Evaluasi Ulang Algoritma Decision Tree	76
Gambar 4. 64 Evaluasi Ulang Algoritma XGBoost.....	77
Gambar 4. 65 Syntax Pemanggilan <i>RandomizedSearchCV</i>	78
Gambar 4. 66 Definisi Parameter Random Forest	79
Gambar 4. 67 Penerapan <i>RandomizedSearchCV</i>	79
Gambar 4. 68 Menyimpan Model Parameter Terbaik Dalam Variabel Baru	79
Gambar 4. 70 Evaluasi Random Forest Setelah Penyetelan Parameter	80
Gambar 4. 69 Parameter Terbaik Random Forest.....	80
Gambar 4. 71 Definisi Parameter KNN	80
Gambar 4. 72 Penerapan <i>RandomizedSearchCV</i>	82
Gambar 4. 73 Menyimpan Model Parameter Dalam Variabel Baru.....	82
Gambar 4. 74 Parameter Terbaik KNN dan Hasil Evaluasi.....	82
Gambar 4. 75 Definisi Parameter Decision Tree	83
Gambar 4. 76 Penerapan <i>RandomizedSearchCV</i>	84

Gambar 4. 77 Menyimpan Model Parameter Terbaik Dalam Variabel Baru	84
Gambar 4. 78 Parameter Terbaik Decision Tree dan Hasil Evaluasi.....	85
Gambar 4. 79 Definisi Parameter XGBoost.....	85
Gambar 4. 80 Penerapan <i>RandomizedSearchCV</i>	86
Gambar 4. 82 Parameter Terbaik XGBoost dan Hasil Evaluasi	87
Gambar 4. 81 Menyimpan Model Parameter Terbaik Dalam Variabel Baru	87



LAMPIRAN

Lampiran 1. Surat Tugas Dosen Pembimbing	99
Lampiran 2. Lembar Konsultasi Dosen Pembimbing 1 dan 2	100
Lampiran 3. Surat Pernyataan Dosen Pembimbing 1	102
Lampiran 4. Surat Pernyataan Dosen Pembimbing 2	103
Lampiran 5. Lembar Persetujuan Dosen Pembimbing 1 dan 2.....	104
Lampiran 6. Surat Permohonan Penelitian Penulisan Skripsi.....	105
Lampiran 7. Surat Balasan Perusahaan	106
Lampiran 8. Instrumen Wawancara Penelitian	107

