



TESIS

PENGEMBANGAN MODEL HIBRIDA UNTUK PREDIKSI MASA TUNGGU LULUSAN MENGUNAKAN ELASTICNET DAN LIGHTGBM STUDI KASUS PADA UNIVERSITAS XYZ

WIGANANDA FIRDAUS PUTRA ADITYA

NPM 23066020021

DOSEN PEMBIMBING

Dr. Eng. Agussalim, S.Pd., MT.

Dr. Rizky Parlika, S.Kom, M.Kom.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI TEKNOLOGI INFORMASI
SURABAYA
2026**

LAPORAN TESIS

PENGEMBANGAN MODEL HIBRIDA UNTUK PREDIKSI MASA TUNGGU LULUSAN MENGGUNAKAN ELASTICNET DAN LIGHTGBM STUDI KASUS PADA UNIVERSITAS XYZ

Tesis

Sebagai Salah Satu Syarat untuk Mencapai Gelar Magister Program
Studi Teknologi Informasi

Disusun dan diajukan oleh:

WIGANANDA FIRDAUS PUTRA ADITYA
23066020021

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN “VETERAN” JAWA TIMUR
SURABAYA
2026**

**PENGEMBANGAN MODEL HIBRIDA UNTUK PREDIKSI MASA
TUNGGU LULUSAN MENGGUNAKAN ELASTICNET DAN LIGHTGBM
STUDI KASUS PADA UNIVERSITAS XYZ**

Oleh :
WIGANANDA FIRDAUS PUTRA ADITYA
NPM 23066020021

Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Ujian Tesis
Prodi Magister Teknologi Informasi Fakultas Ilmu Komputer Universitas
Pembangunan Nasional Veteran Jawa Timur
Pada tanggal 22 Januari 2026

Menyetujui,

Dr. Eng. Agussalim, S.Pd., MT.
NIP: 19850811 2019031 005



(Pembimbing I)

Dr. Rizky Parlika, S.Kom, M.Kom
NIP: 19840518 2021211 003



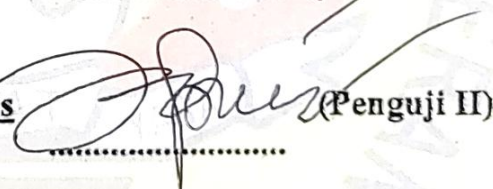
(Pembimbing II)

Dr. Basuki Rahmat, S.Si., MT.
NIP: 19690723 202121 1 002



(Penguji I)

Dr. Rr. Ani Dijah Rahajoe, ST, M.Cs
NIP: 19730512 2005012 003



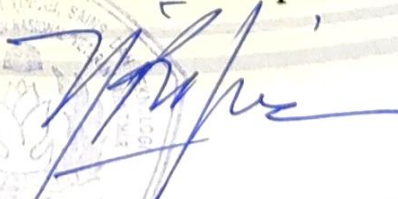
(Penguji II)

Dr. Faisal Muttaqin, S.Kom, M.T
NIP: 19851231 2021211 009



(Penguji III)

Mengetahui,
Dekan Fakultas Ilmu Komputer



Prof. Dr. Ir. Novirina Hendrasarie, MT.
NIP. 19681126 1994032 001

**PENGEMBANGAN MODEL HIBRIDA UNTUK PREDIKSI MASA
TUNGGU LULUSAN MENGGUNAKAN ELASTICNET DAN LIGHTGBM
STUDI KASUS PADA UNIVERSITAS XYZ**

**Disusun Oleh :
WIGANANDA FIRDAUS PUTRA ADITYA
NPM 23066020021**

Telah disetujui untuk mengikuti Ujian Tesis

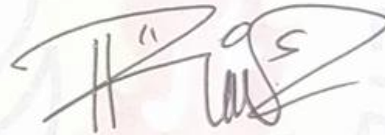
Menyetujui,

Pembimbing I



Dr. Eng. Agussalim, S.Pd., MT.
NIP. 198508112019031005

Pembimbing II



Dr. Rizky Parlika, S.Kom, M.Kom
NIP. 198405182021211003

**Mengetahui,
KOORDINATOR PRODI**



Dr. Eng. Agussalim, S.Pd., MT.
NIP. 198508112019031005

SURAT PERNYATAAN BEBAS PLAGIASI

Saya yang bertanda tangan dibawah ini:

Nama : Wigananda Firdaus Putra Aditya
NPM : 23066020021
Program : Magister (S2)
Program Studi : Magister Teknologi Informasi
Fakultas : Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah Tesis ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang/lembaga lain, kecuali yang secara tertulis disitasi dalam dokumen ini dan disebutkan secara lengkap dalam daftar pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabila dikemudian hari ditemulan indikasi plagiat pada Skripsi/Tesis/Desertasi ini, saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.

Surabaya, 22 Januari 2026
Yang Membuat pernyataan



Wigananda Firdaus Putra Aditya
23066020021

KATA PENGANTAR

Puji syukur Alhamdulillah penulis panjatkan kehadiran Allah SWT, yang telah melimpahkan berkah, rahmat, dan hidayah-Nya, sehingga penulis dapat menyelesaikan tesis ini dengan baik dan tepat waktu. Tesis ini disusun sebagai salah satu syarat untuk memperoleh gelar Magister Teknologi Informasi. Adapun penelitian ini berfokus pada Prediksi Masa Tunggu Lulusan Menggunakan ElasticNet dan LightGBM.

Gagasan yang diangkat dalam proposal ini adalah pergeseran dari praktik deskriptif menuju prediksi individual yang dapat ditindaklanjuti. Pendekatan yang digunakan adalah arsitektur hibrida ElasticNet dan LightGBM. Evaluasi dilakukan dengan K-Fold Cross-Validation dan pelaporan R^2 , RMSE, MAE, sementara penjelasan Variabel yang berpengaruh disajikan melalui koefisien ElasticNet dan feature importance dari LightGBM. Luaran yang diharapkan ialah prediksi masa tunggu lulusan serta daftar faktor paling berpengaruh yang dapat langsung dipakai untuk intervensi Pusat Karir di masa mendatang.

Bukan hal yang mudah untuk mewujudkan gagasan-gagasan tersebut ke dalam sebuah susunan tesis. Proses penyusunan ini membutuhkan pemikiran yang matang, usaha yang konsisten, serta dukungan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan terima kasih yang sebesar-besarnya kepada:

1. Dr. Eng. Agussalim, S.Pd., MT. Selaku Kaprodi Magister Teknologi Informasi dan sebagai Dosen pembimbing pertama
2. Dr. Rizky Parlika, S.Kom, M.Kom sebagai Dosen pembimbing kedua
3. Dr. Basuki Rahmat, S.Si., MT. Selaku Wakil Dekan 3 Fasilkom dan sebagai Dosen penguji pertama
4. Dr. Rr. Ani Dijah Rahajoe, ST, M.Cs. sebagai Dosen penguji kedua
5. Dr. Faisal Muttaqin, S.Kom, M.T. Sebagai Dosen penguji ketiga
6. Dr. I Gede Susrama Mas Diyasa, ST., MT. Sebagai Wakil Dekan 1 Fasilkom.
7. Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur dan Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur yang telah memfasilitasi saya menempuh program magister serta para dosen dan rekan-rekan yang telah membantu dalam melakukan penelitian ini.

8. Seluruh dosen dan staf administrasi prodi Magister Teknologi Informasi, yang telah membantu berbagai aspek akademik dan administrasi selama masa studi.
9. Istri dan Seluruh Keluarga yang telah memberikan semangat, dukungan, dan doa yang tulus selama masa studi serta proses penelitian ini.
10. Rekan-rekan kerja dan teman-teman yang telah memberikan dukungan, inspirasi, dan kerja sama selama masa studi dan proses penelitian ini.
11. Komunitas IT Dinamika serta Teman-teman mahasiswa Prodi MTI yang tanpa lelah berjuang bersama.
12. Keluarga WaowRek III, DLC, dan Clan ICS yang telah memberikan semangat dan dukungan kepada penulis.
13. Pihak-pihak lain yang tidak dapat disebutkan satu per satu, tetapi telah memberikan kontribusi dalam bentuk apapun demi terselesaikannya tesis ini.

Akhirnya, kepada kedua orang tua tercinta saya mengucapkan terima kasih dan sembah sujud atas doa, pengorbanan dan memotivasi mereka selama saya menempuh pendidikan. Penghargaan yang besar juga saya sampaikan kepada kedua orang tua saya atas motivasi dan dukungan yang tak ternilai.

Penulis,

Wigananda Firdaus Putra Aditya

ABSTRAK

WIGANANDA FIRDAUS PUTRA ADITYA. PENGEMBANGAN MODEL HIBRIDA UNTUK PREDIKSI MASA TUNGGU LULUSAN MENGGUNAKAN ELASTICNET DAN LIGHTGBM STUDI KASUS PADA UNIVERSITAS XYZ (dibimbing oleh **Dr. Eng. Agussalim, S.Pd., MT.** dan **Dr. Rizky Parlika, S.Kom, M.Kom**)

Masa tunggu lulusan dalam memperoleh pekerjaan pertama merupakan salah satu indikator penting bagi perguruan tinggi. Hal tersebut menjadi salah satu penilaian dalam akreditasi dan IKU. Penelitian ini mengembangkan sebuah model hibrida *stacking* untuk memprediksi masa tunggu yang mengkombinasikan metode *ElasticNet* dan *LightGBM*. Data yang digunakan bersumber dari data *Tracer Study* dan data akademik Universitas XYZ, lulusan tahun 2020-2023 (N=493) dan setelah *data cleaning* digunakan data final sebanyak N=449. Variabel prediktor yang digunakan diantaranya masa studi, IPK, TOEFL, SSKM, nilai KP, Nilai TA, prodi, provinsi, jenis kelamin, masa mencari kerja, dan jumlah lamaran. Seluruh proses dibangun dalam *pipeline* untuk mencegah *data leakage*, mencakup imputasi nilai hilang, *standardization* fitur numerik, *ordinal encoding* untuk variabel berjenjang, serta *one-hot encoding* untuk variabel nominal. Tuning hiperparameter dilakukan dengan *RandomizedSearchCV*, model *stacking* dibentuk dari dua *base learner* *ElasticNet* dan *LightGBM* dengan *ElasticNet* sebagai meta-learner yang dilatih menggunakan prediksi *out-of-fold*. Evaluasi internal menggunakan *Cross Validation* menunjukkan model *stacking* memberikan kinerja terbaik dan lebih stabil dengan nilai R^2 mean=0,5884; RMSE mean=2,0168; MAE mean=1,6429. Uji eksternal pada cohort yang berbeda tahun lulus 2024 (N=62) menghasilkan R^2 =0,5304 dengan RMSE=2,5277 dan MAE=1,9387 hal tersebut mengindikasikan performa tetap terjaga meski galat meningkat pada nilai masa tunggu tinggi yang relatif jarang. Hasil *permutation importance* beberapa fitur teratas sebagai prediktor dominan. Model ini dapat digunakan sebagai alat bantu analitik untuk pemetaan risiko masa tunggu dan perumusan intervensi layanan karier secara lebih terarah.

Kata kunci: *tracer study*, masa tunggu lulusan, *regresi*, *ElasticNet*, *LightGBM*, Hibrida.

Halaman ini sengaja dikosongkan

ABSTRACT

WIGANANDA FIRDAUS PUTRA ADITYA. DEVELOPMENT OF A HYBRID MODEL TO PREDICTION OF GRADUATES WAITING PERIOD USING ELASTIC NET AND LIGHTGBM CASE STUDY AT XYZ UNIVERSITY (supervised by Dr. Eng. Agussalim, S.Pd., MT. and Dr. Rizky Parlika, S.Kom, M.Kom)

Graduate waiting time to obtain a first job is an important indicator for higher education institutions. It is commonly used in accreditation assessments and key performance indicators (IKU). This study develops a hybrid stacking model to predict graduate waiting time by combining ElasticNet and LightGBM. The data were drawn from Universitas XYZ's Tracer Study and academic records for graduates from 2020-2023 (N=493, after data cleaning, the final dataset consisted of N=449 observations. The predictors included time to degree completion, GPA, TOEFL score, SSKM, internship grade (KP), thesis grade (TA), study program, province, gender, job-search duration, and number of job applications. The entire workflow was implemented within a pipeline to reduce the risk of data leakage, incorporating missing-value imputation, numerical feature standardization, ordinal encoding for ordered variables, and one-hot encoding for nominal variables. Hyperparameters were tuned using RandomizedSearchCV. The stacking model was constructed with ElasticNet and LightGBM as base learners and ElasticNet as the meta-learner trained on out-of-fold predictions. Internal evaluation using cross-validation showed that the stacking model achieved the best and most stable performance with mean $R^2=0.5884$, mean RMSE=2.0168, and mean MAE=1.6429. External testing on an independent 2024 graduate cohort (N=62) yielded $R^2=0.5304$ with RMSE=2.5277 and MAE=1.9387, indicating that performance was largely maintained although errors increased for higher waiting-time cases that were relatively rare. Permutation importance highlighted several top features as dominant predictors. This model can be used as an analytical decision-support tool for mapping waiting-time and designing more targeted career service interventions.

Keywords: tracer study, graduate waiting time, regression, ElasticNet, LightGBM, hybrid.

Halaman ini sengaja dikosongkan

DAFTAR ISI

HALAMAN JUDUL	i
PENGAJUAN TESIS	ii
PERSETUJUAN TESIS	iii
PERNYATAAN KEASLIAN TESIS.....	iv
KATA PENGANTAR	v
ABSTRAK.....	vii
<i>ABSTRACT</i>	viii
DAFTAR ISI	x
DAFTAR GAMBAR.....	xv
DAFTAR TABEL	xvii
DAFTAR NOTASI.....	xix
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan Penelitian	3
1.4 Manfaat Penelitian	4
1.5 Ruang Lingkup Penelitian	4
1.6 Batasan Permasalahan	5
1.7 Kontribusi Penelitian	5
1.7.1 Kontribusi Teoritis	5
1.7.2 Kontribusi Praktis	6
1.7.3 Kontribusi Metodologis	6
BAB II TINJAUAN PUSTAKA	7
2.1 Penelitian Terdahulu	7
2.2 <i>Tracer Study</i> dan Masa Tunggu Lulusan.....	20
2.3 Dasar Teori Pemodelan.....	21
2.3.1 ElasticNet	22
2.3.2 LightGBM	25
2.3.3 Hybrid Model (Model Hibrida).....	27
2.3.4 Stacking.....	27
2.3.5 Out-of-fold (OOF).....	29
2.4 <i>Data Preparation</i>	30

2.4.1 <i>Cleaning Data</i>	31
2.4.2 Preprocessing Data.....	31
2.5 Evaluasi Model	32
2.5.1 Cross Validation.....	32
2.5.2 R^2	33
2.5.3 Root Mean Squared Error (RMSE).....	34
2.5.4 Mean Absolute Error (MAE)	35
2.5.5 <i>Permutation Importance</i>	36
2.6 Penentuan Variabel Prediktor	37
2.7 Research Gap	38
2.7.1 Gap Tracer Study	39
2.7.2 Research Gap Metode	40
2.8 Research Environment (<i>Python</i>).....	42
2.8.1 Pandas	42
2.8.2 Numpy.....	43
2.8.3 Scikit-learn	44
2.8.4 Matplotlib.....	45
2.8.5 Jupyter Notebook	46
2.8.5 <i>LGBMRegressor</i>	46
BAB III METODOLOGI PENELITIAN	49
3.1 Alur Penelitian	49
3.1.1 Diagram Alir Penelitian	49
3.1.2 Alur Pemodelan.....	51
3.2 Data dan Variabel Penelitian	56
3.2.1 Variabel Penelitian.....	56
3.2.2 Variabel Target.....	57
3.2.3 Variabel Prediktor	58
3.2.4 Skala Data dan Tipe Variabel	59
3.3 Persiapan Data dan Pra Pemrosesan	60
3.3.1 Cleaning Data.....	60
3.3.2 Pra-Pemrosesan Data	62
3.4 Pengembangan Pemodelan	63
3.4.1 ElasticNet	63

3.4.2 LightGBM.....	65
3.4.3 Hibrida <i>Stacking</i>	67
3.4.4 Simulasi Manual Mekanisme Stacking.....	70
3.5 Evaluasi Model.....	73
3.5.1 Validasi Internal (<i>K-Fold Cross Validation</i>).....	73
3.5.2 Evaluasi Eksternal (Cohort 2024).....	74
3.5.3 Metrik Evaluasi (R^2 , RMSE, MAE).....	74
3.6 Hasil Prediksi.....	75
3.7 Perangkat dan Lingkungan Penelitian.....	75
3.7.1 Perangkat Lunak dan Eksekusi.....	76
3.7.2 Library Python.....	76
3.8 Skenario Pengujian Model.....	77
3.8.1 Luaran yang Dihasilkan.....	77
3.8.2 Rangkaian Skenario Pengujian.....	77
BAB IV PEMBAHASAN DAN SKENARIO PENGUJIAN.....	79
4.1 Dataset dan Variabel.....	79
4.2 Pra-Pemrosesan Data.....	80
4.2.1 Data Cleaning.....	82
4.2.2 Skema Pra-Pemrosesan Variabel.....	83
4.3 Implementasi Skenario Pengujian.....	86
4.3.1 Skenario 1 - Optimasi ElasticNet.....	86
4.3.2 Skenario 2 - Optimasi LightGBM.....	87
4.3.3 Skenario 3 - Pemodelan Hibrida Stacking (OOF).....	90
4.3.4 Skenario 4 - Evaluasi Eksternal Cohort 2024.....	97
4.4 Analisis Permutation Importance.....	100
BAB V KESIMPULAN DAN SARAN.....	105
5.1 Kesimpulan.....	105
5.2 Saran.....	105
DAFTAR PUSTAKA.....	107
Lampiran.....	111

Halaman ini sengaja dikosongkan

DAFTAR GAMBAR

Gambar 2. 1 Ilustrasi ElasticNet.....	23
Gambar 2. 2 Ilustrasi LightGBM.....	25
Gambar 2. 3 Ilustrasi Cross Validation	33
Gambar 2. 4 Penelitian Terdahulu tentang Tracer Study	39
Gambar 2. 5 Gap Research berdasarkan Metode	41
Gambar 2. 6 Logo Pandas.....	42
Gambar 2. 7 Logo NumPy.....	43
Gambar 2. 8 Logo Scikit-learn	44
Gambar 2. 9 Logo Matplotlib	45
Gambar 2. 10 Logo Jupyter	46
Gambar 3. 1 Alur Metodologi Penelitian	49
Gambar 3. 2 Diagram Alir Pemodelan <i>Hibrida Stacking</i>	51
Gambar 3. 3 Histogram variabel target	58
Gambar 3. 4 <i>Heatmap</i> Korelasi antar Variabel	59
Gambar 3. 5 Alur <i>Cleaning data</i>	60
Gambar 3. 6 Alur Pra-pemrosesan data.....	62
Gambar 3. 7 Alur ElasticNet	64
Gambar 3. 8 Alur <i>LightGBM</i>	66
Gambar 3. 9 Alur Hibrida <i>Stacking</i>	68
Gambar 4. 1 Alur pra-pemrosesan.....	84
Gambar 4. 2 Perbandingan R^2	94
Gambar 4. 3 <i>ScatterPlot</i> Aktual dan Prediksi.....	95
Gambar 4. 4 Perbandingan R^2 Data 2020-2023 dengan Data 2024	98
Gambar 4. 5 <i>Scatter Plot</i> Prediksi Aktual VS Prediksi Cohort 2024.....	98
Gambar 4. 6 Perbandingan R^2 model tunggal pada data 2024	99
Gambar 4. 7 Kurva <i>Cumulative Permutation Importance</i>	103

Halaman ini sengaja dikosongkan

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu.....	7
Tabel 3. 1 Tabel Variabel Penelitian	57
Tabel 3. 2 Data Variabel Simulasi.....	70
Tabel 3. 3 Simulasi pembagian $K=3$	71
Tabel 3. 4 Simulasi hasil prediksi <i>OOF base learner</i>	71
Tabel 3. 5 Simulasi hasil prediksi <i>OOF</i> seluruh data.....	72
Tabel 3. 6 Simulasi data meta (Z)	72
Tabel 3. 7 Simulasi hasil <i>Stacking</i>	73
Tabel 3. 8 Penggunaan <i>Library Python</i>	76
Tabel 3. 9 Rangkuman skenario pengujian dan luaran.....	78
Tabel 4. 1 Dataset Penelitian	79
Tabel 4. 2 Ringkasan Hasil Data Cleaning.....	83
Tabel 4. 3 Tabel Hasil Tuning <i>ElasticNet</i>	86
Tabel 4. 4 Ringkasan hasil tuning terbaik <i>ElasticNet</i>	87
Tabel 4. 5 Kandidat <i>hyperparameter</i> dan hasil tuning.....	88
Tabel 4. 6 Ringkasan hasil tuning terbaik LightGBM	89
Tabel 4. 7 Variabel Tuning <i>Stacking</i>	91
Tabel 4. 8 Perbandingan Model <i>Stacking</i> dengan Tunggal.....	92
Tabel 4. 9 Hasil Pemodelan Cohort 2024.....	97
Tabel 4. 10 Hasil <i>Permutation Importance</i>	100

Halaman ini sengaja dikosongkan

DAFTAR NOTASI

X	:	Matriks fitur / Variabel prediktor
y	:	Variabel target/dependen (Masa_Tunggu)
x	:	Vektor fitur untuk satu observasi
n	:	Jumlah observasi (ukuran data)
i	:	Indeks observasi ke- i
$(\bar{y})$	:	Rata-rata nilai aktual y
$(\widehat{y_i})$	:	Prediksi untuk observasi ke- i
$(\widehat{y}(x))$	:	Prediksi untuk input fitur x
(K)	:	Jumlah <i>fold</i> pada <i>K-Fold Cross Validation</i>
(k)	:	Indeks <i>fold</i> ke- k
(β)	:	Vektor koefisien regresi
$(\hat{\beta}_{EN})$	:	Koefisien hasil estimasi ElasticNet
(λ_1)	:	Parameter penalti L1 (Lasso)
(λ_2)	:	Parameter penalti L2 (Ridge)
$(\beta _1)$	:	Norma L1 dari beta (jumlah nilai absolut koefisien)
$(\beta^T \beta)$	:	Norma L2 kuadrat dari beta
(T)	:	Jumlah pohon/iterasi <i>boosting</i> pada <i>LightGBM</i>
(t)	:	Indeks pohon/iterasi ke- t
$(f_t(x))$	:	Fungsi prediksi pohon ke- t untuk input x
$(\widehat{y^{(t)}}(x))$	:	Prediksi pada iterasi ke- t untuk input x
(η)	:	<i>Learning rate</i> pada <i>boosting</i>
$(\widehat{y_{EN}^{OOF}})$	:	Prediksi <i>out-of-fold</i> dari <i>base learner ElasticNet</i>
$(\widehat{y_{LGBM}^{OOF}})$	:	Prediksi <i>out-of-fold</i> dari <i>base learner LightGBM</i>
(Z)	:	Data meta dari gabungan prediksi <i>OOF base learner</i>
$(g(\cdot))$	:	Fungsi <i>meta-learner</i> /penggabung pada <i>stacking</i>
(w_0)	:	Intersep/bias pada <i>meta-learner stacking</i>
(w_1, w_2)	:	Bobot/koefisien untuk prediksi <i>base learner</i> pada <i>meta-learner</i>
$(\widehat{y_{EN}}(x))$	:	Prediksi <i>base learner ElasticNet</i> untuk input x
$(\widehat{y_{LGBM}}(x))$	:	Prediksi <i>base learner LightGBM</i> untuk input x
$(\widehat{y_{final}}(x))$	:	Prediksi akhir model <i>stacking</i> untuk input x

Halaman ini sengaja dikosongkan